Faculty Publications, Department of Psychology

Psychology, Department of

September 2003

Testing the Dynamic Field Theory: Working Memory for Locations Becomes More Spatially Precise Over Development

Anne R. Schutte
University of Nebraska-Lincoln, aschutte2@unl.edu

John P. Spencer
University of Iowa, john-spencer@uiowa.edu

Gregor Schöner
Institut für Neuroinformatik, Gregor.Schoener@neuroinformatik.ruhr-uni-bochum.de

Follow this and additional works at: http://digitalcommons.unl.edu/psychfacpub

Part of the Psychiatry and Psychology Commons

Testing the Dynamic Field Theory: Working Memory for Locations Becomes More Spatially Precise Over Development

Anne R. Schutte, John P. Spencer and Gregor Schöner

The dynamic field theory predicts that biases toward remembered locations depend on the separation between targets, and the spatial precision of interactions in working memory that become enhanced over development. This was tested by varying the separation between A and B locations in a sandbox. Children searched for an object 6 times at an A location, followed by 3 trials at a B location. Two- and 4-year-olds', but not 6-year-olds', responses were biased toward A when A and B were 9-in. and 6-in. apart. When A and B were separated by 2 in., however, 4- and 6-year-olds' responses were biased toward A. Thus, the separation at which responses were biased toward A decreased across age groups, supporting the predictions of the theory.

When trying to locate an object, several types of information can be used. One of the most common types of information people use is a longer term memory of where objects have been found in the past. Evidence suggests that such experience-dependent information is used throughout development. Use of experience-dependent information is well documented in infancy (e.g., Piaget, 1954; Smith, Thelen, Titzer, & McLin, 1999). For instance, infants who have found a hidden toy repeatedly at an A location will search back at A after seeing a toy hidden at a B location. Two- and 3-year-olds show similar experience-dependent effects in tasks that involve repeated hiding and finding events at a single location followed by a hiding event at a second location (Schutte & Spencer, 2002; Spencer, Smith, & Thelen, 2001; Zelazo, Reznick, & Spinnazzola, 1998). Experience-dependent effects have also been reported in several studies with older children and adults. For example, 6-year-olds, 11-year-olds, and adults make response errors in the direction of previously responded-to locations when recalling locations in the absence of salient location cues (Spencer & Hund, 2002, in press).

Although experience-dependent effects have been observed in infancy through adulthood, the situations in which such effects are apparent change. For example, at about 12 months of age, infants stop making A-not-B errors in the Piagetian task where the A and B locations are marked by lids (e.g., Diamond & Doar, 1989; Piaget, 1954; Smith et al., 1999). In contrast, children as old as 3 years of age make A-not-B-type errors if the A and B locations are not marked by lids but instead are two locations in a homogeneous task space such as a sandbox (Schutte & Spencer, 2002; Spencer et al., 2001; for related residual A-not-B errors with toddlers, see Piaget, 1954). Beyond 3 years of age, there are quantitative changes in the precision of children's memory for locations. In a series of studies, Spencer and colleagues (Schutte & Spencer, 2002; Spencer & Hund, 2002, in press) reported a quantitative reduction in the size of experience-dependent effects over development.

What is changing over development to produce these changes in experience-dependent effects? Several models have been proposed recently to account for the disappearance of the A-not-B error in infancy: the memory+inhibition model (Diamond, Cruttenden, & Neiderman, 1994), the response versus representation model (Marcovitch & Zelazo, 1999; Zelazo et al., 1998), the latent versus active memory model (Munakata, 1998; Munakata, McClelland, Johnson, & Siegler, 1997), and the dynamic field theory (DFT; Smith et al., 1999; Thelen, Schöner, Scheier, &

Smith, 2001). Moreover, a category-adjustment model has been proposed to explain recall biases with younger children, older children, and adults in tasks with homogeneous spaces (Huttenlocher, Hedges, & Duncan, 1991). Of these accounts, only the DFT has been used to span the gap between infancy and later development (see Schutte & Spencer, 2002; Spencer et al., 2001). Recent tests of this model suggest that the same memory processes that account for biases toward A in the A-not-B situation in infancy can account for experience- dependent biases in spatial recall tasks later in development.

The goal of the current study was to test a recent proposal by Spencer and colleagues (Spencer & Hund, in press; Spencer & Schöner, 2000) regarding how the DFT might account for the quantitative reduction in experience-dependent effects over development. Specifically, we tested the proposal that enhanced interactions in the model, that is, enhanced interactions in spatial working memory (SWM), underlie this change. Our data demonstrate that, consistent with the model's predictions, the magnitude of experience-dependent effects varies systematically with the age of the child and the spatial separation between target locations.

The DFT

The DFT is a neural network model that captures how location-related information is maintained via the activation of neurons in a working memory field (Schutte & Spencer, 2002; Spencer et al., 2001; Thelen et al., 2001). To explain the concepts of the DFT, we focus on 2-year-olds' performance in an A-not-B sandbox task (Spencer et al., 2001). In this task, children watch as the experimenter buries a toy at an A location in a large sandbox. Following a short delay, the children are allowed to search for the object. This sequence is repeated across six A trials, which are followed by three trials to a nearby B location. Spencer et al. (2001) reported that 2-year-olds' responses on the B trials were systematically biased toward A.

Figure 1 shows how the DFT captures children's biases toward A on the first B trial in the sandbox task. Figure 1d shows the evolution of activation in a SWM field over the course of the B trial. The range of possible reachable locations is captured along the *x* axis, with location A on the left and location B on the right. Time, from the start of the trial to the end of the trial, is on the *y* axis, and activation is on the *z* axis. Stronger activation at a particular location in the field reflects a stronger representation of the associated location in space. This, in turn, influ-

ences where the model (i.e., the child) will reach in the task: The model responds to the location associated with the highest activation at the end of the memory delay.

Activation profiles in the dynamic field depend, in part, on the three inputs to the field depicted in Figure 1. The task input captures the perceptual cues that specify the possible hiding locations in the task space. In the sandbox task, the task input is set to zero because the task space is a homogeneous field of sand (see Figure 1a). The target input captures the hiding event (Figure 1b). At the beginning of the trial when the toy is not visible, the target input is zero. Next, the experimenter places the toy in the sand at the B location. This is captured by the peak of activation at the B location between 2 s and 4 s. When the toy is no longer visible, the target input returns to zero. The final input, the memory input (Figure 1c), captures 2-year-olds' longer term memory of previous trials. In Figure 1c, activation is centered at the A location because all of the previous trials were to the A location. Note that this input is constructed from trial to trial by the model, even though this process is not depicted here (see Thelen et al., 2001).

The task, target, and memory inputs are integrated within the working memory field shown in Figure 1d. These inputs, in conjunction with the properties of the working memory field itself, determine how strongly different locations are represented from second to second during the trial. At the beginning of the trial, before the toy is visible, there is slightly stronger activation at location A than at location B in the working memory field. This is because of the memory input at A. At 2 s, the experimenter hides the toy in the sand. The target input that captures this event creates a peak of activation in the working memory field at the B location. At 4 s, the toy is hidden and the target input returns to zero. Nevertheless, target-related activation in working memory sustains itself during the memory delay. Moreover, this self-sustaining peak of activation drifts systematically toward the A location. As a consequence, the most strongly represented location in working memory after the 10-s delay is very close to A.

The simulation in Figure 1d raises two key questions regarding activation in working memory during the delay. First, how is activation sustained in the absence of the target input? Such sustained activation emerges because neurons in the network interact according to a local excitation– lateral inhibition function. According to this function, activated neurons excite other neurons that code similar spatial locations and inhibit neurons that code
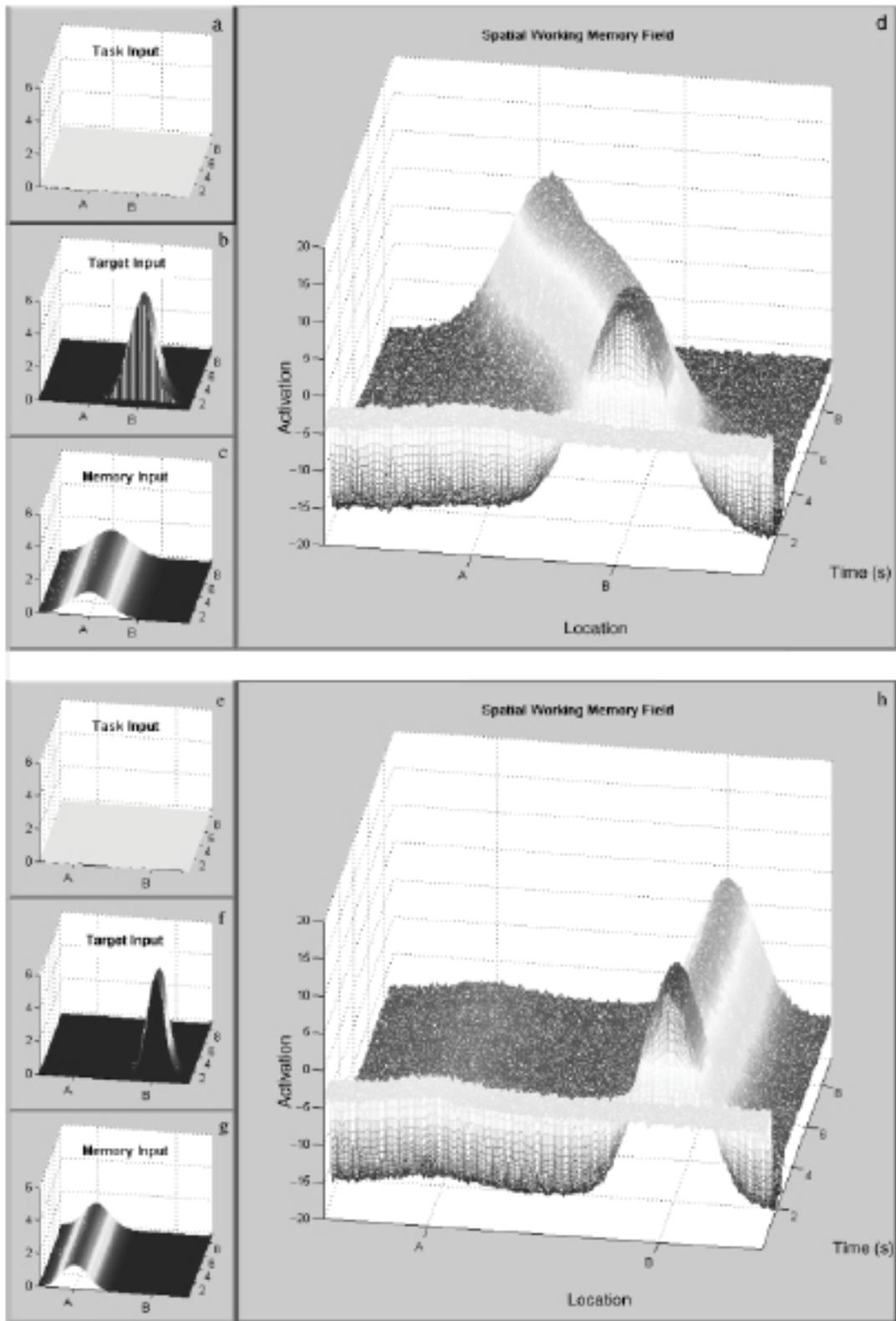
Figure 1. The dynamic field theory of the A-not-B error. The top panel shows time-dependent changes in a spatial working memory field (d) in the context of three inputs (a, b, c) during the first B trial in the sandbox task when the A and B locations are close together. The bottom panel shows changes in a spatial working memory field (h) in the context of three inputs (e, f, g) during the first B trial when the A and B locations are far apart. See text for further details.

locations far away. When neuronal interaction is strong (Figure 1d), a local group of neurons excited by input (e.g., target input) can enter a self-sustaining state in which they reciprocally activate one another through time while inhibiting other neurons far away (Figure 1d).

The second question is why does sustained activation in working memory drift toward A during the memory delay? This occurs because the left edge of the self-sustaining peak overlaps with the right edge of activation in the memory input (Figure 1c). This extra boost of input activation on the left edge of the self-sustaining peak recruits new neurons into the locally excitatory interaction and causes neurons on the right edge of the peak to drop out because of lateral inhibition. As this process continues during the memory delay, the peak systematically drifts to the left; the peak is attracted toward the site that receives the maximal input, in this case, the site associated with the A location. Where exactly the model reaches depends on the length of the delay. In this simulation, if the go signal is given at 5 s, the model would reach between A and B (the location of maximal activation). By contrast, if the go signal is given at 10 s, the model would reach very close to A. Thus, experience-dependent biases in the model are caused by a time-dependent process driven by the spatial overlap between the locally self-excitatory peak in working memory and the distribution of activation in the memory input.

Modifying Experience-Dependent Biases Through Changes in Spatial Overlap

Given that experience-dependent effects in the DFT arise because of the spatial overlap between self-sustaining peaks in working memory and activation in longer term memory, it should be possible to eliminate or create such effects by manipulating factors in the model—and associated factors in experiment—that modify the degree of spatial overlap. Here, we discuss two such factors: the separation between the A and B locations in the sandbox task, and the precision of the local excitation – lateral inhibition interaction function in working memory.

The Separation Between A and B

One way to modify the spatial overlap between the self-sustaining peak in Figure 1d and activation in the memory input is simply to move A and B farther apart. This is shown in the lower panel of Figure 1. All of the inputs in this simulation are the same as in the top panel except that the separation between A and B has been increased. Now, the model that made an A-not-B-type er-

ror in the top panel does not make the error in the bottom panel. A more formal analysis of the effect of target separation on experience-dependent biases is captured by the bold solid line in Figure 2. The *z* axis in this figure shows the amount of bias toward the A location (in units) for simulations of the model on the first B trial after a 10-s memory delay. The *y* axis shows the separation between A and B in these simulations. Separation was varied between 20 and 80 units. The *x* axis shows variations in the precision of the locally excitatory part of the interaction function in the working memory field. We discuss this manipulation later. Note that these simulations were conducted without noise to reveal the deterministic component of the model's performance. Later, we discuss how the model responds in the presence of stochastic perturbations provided by noise (see the Quantitative Modeling section).

As can be seen in Figure 2 (bold solid line), the model shows zero bias toward A at separations of 66 to 80 units (the small bias toward A in the figure is caused by numerical imprecision). At smaller separations between A and B, activation in working memory becomes more biased toward A, such that there is a very strong bias toward A at a separation of 40 to 50 units. As separation decreases even further, the bias toward A decreases. This reflects a type of ceiling effect: At a separation of 40 units, the model makes the maximum possible error (i.e., 40 units); similarly, at a separation of 30 units, the model makes the maximum possible error (i.e., 30 units). Thus, as the separation decreases below 40 units, the model makes the maximum possible error, which necessarily decreases as separation decreases. As the separation between A and B decreases, therefore, a bias toward A emerges and becomes stronger, ultimately reducing slightly as A and B are moved quite close.

Spatial Precision of Interactions in Working Memory

In addition to changing the separation between A and B, one can also modify experience-dependent effects in the DFT by changing the precision of the local excitation– lateral inhibition function. The interaction function in the model has a Gaussian form (see the Quantitative Modeling section). Thus, enhanced spatial precision indicates an interaction function with a narrow region of local excitation (a Gaussian with a narrow width) and an associated broad region of lateral inhibition. This results in a narrower peak. Moreover, narrow peaks tend to be more stable (i.e., less subject to drift) because sites to the left and right of the peak are more effectively inhibited. Final-

ly, dynamic fields with narrow peaks show weaker experience-dependent effects for the same reason described earlier: A narrower self-sustaining peak is less likely to overlap with input from longer term memory. This is captured by the bold dashed line in Figure 2. When the interaction width is narrow (10 – 12 units), the model shows no bias toward the A location after a 10-s delay. As the interaction width increases, however, a bias toward A emerges (e.g., 16 units) and becomes systematically larger, peaking at an interaction width of 22 to 30 units.

These simulation results are consistent with a recent proposal by Spencer and colleagues (Spencer & Hund, in press; Spencer & Schöner, 2000) that interactions in SWM become more spatially precise (narrower local excitation) over development. Thus, in Figure 2, broad widths (e.g., 28 units) would capture younger children's abilities, whereas narrow widths (e.g., 12 units) would capture older children's abilities. As the interaction widths narrow, activation peaks become more stable and less subject to spatial drift (for related effects, see Compte, Brunel, Goldman-Rakic, & Wang, 2000). As the

bold dashed line in Figure 2 demonstrates, this can explain the reduction in the magnitude of experience-dependent biases over development (Schutte & Spencer, 2002; Spencer & Hund, 2002, in press).

The solid and dashed lines in Figure 2 demonstrate that target separation and interaction width in the DFT can modify experience-dependent effects. The complete surface in this figure demonstrates that these two factors interact. As can be seen in Figure 2, the separation at which biases toward A first emerge varies systematically as the interaction width is manipulated. For instance, at an interaction width of 28, biases toward A are present across most of the target separations examined, whereas at an interaction width of 12, biases toward A are only present at separations smaller than 50 units. This characteristic of the model, combined with the proposal that interactions in working memory become more spatially precise over development, leads to the prediction that the presence of A-not-B-type biases in the sandbox task should depend on separation and age. The goal of this study was to test this prediction.
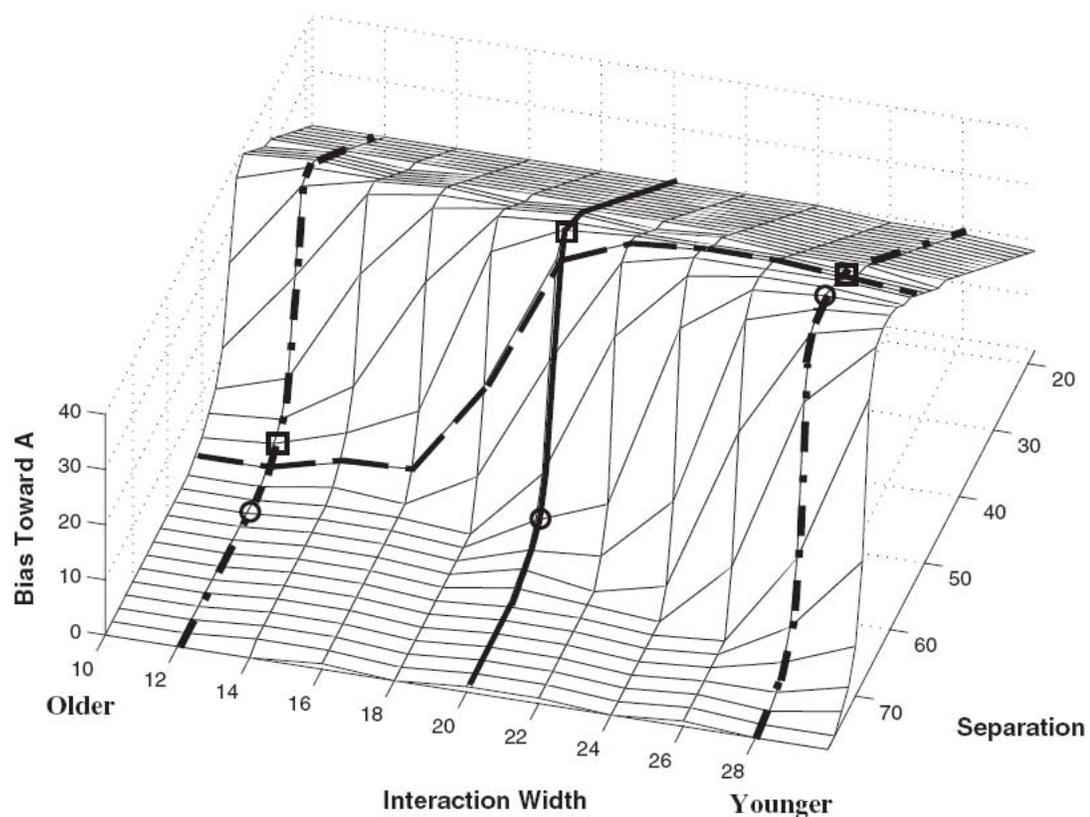

Figure 2. The amount of bias toward A (*z* axis) on the first B trial for different widths of neuronal interaction (*x* axis) and separations between A and B locations (*y* axis). Note that wider interaction widths representing younger children are to the right and narrower interaction widths representing older children are to the left. See text for further details.

Specific Goals

The specific predictions we tested are captured by the bold solid line in Figure 2 and the two dashed-dotted lines. According to the proposal by Spencer and colleagues (Spencer & Hund, in press; Spencer & Schöner, 2000), these three lines represent three different points in development, with the youngest age to the right and the oldest age to the left. If these three ages are tested in the A-not-B sandbox task with A and B far apart (see circles in Figure 2), the oldest group would not make A-not-B-type errors, the middle age group would make small A-not-B-type errors, and the youngest age group would make large A-not-B-type errors. The analysis of the DFT predicts, however, that all three of these age groups would make A-not-B-type errors when tested at a smaller separation (see squares in Figure 2). It is important to note that the prediction does not assume a linear relationship between age and interaction width. It is likely there are more dramatic changes in interaction width at certain points in development than others.

Experiment 1 tested 2-, 4-, and 6-year-olds in the A-not-B sandbox task used by Spencer et al. (2001), with A and B separated by 9 in. Experiment 2 tested these same age groups with A and B separated by 6 in. Finally, Experiment 3 tested 4-, 6-, and 11-year-olds and adults in the sandbox task with A and B separated by only 2 in. Results of these experiments show the separation by age interaction predicted by the analysis of the DFT in Figure 2.

As with any test of a theoretical model, it is important to consider the nature of the predictions tested. Given that the dynamic field model is in a class of neural network models, one question is whether the predictions are behavioral or neurophysiological in nature. The DFT is a functional model of behavior, and the predictions tested here are behavioral predictions based on how the model behaves in different situations. The DFT, however, has several components that are inspired by neurophysiology. For example, the representation of particular locations in space by a population of spatially tuned neurons is inspired by studies of cortical neurophysiology (e.g., in motor cortex: Georgopoulos, Kettner, & Schwartz, 1988; Georgopoulos, Taira, & Lukashin, 1993; in premotor cortex: di Pellegrino & Wise, 1993; in prefrontal cortex: di Pellegrino & Wise, 1993; Wilson, Scalaidhe, & Goldman-Rakic, 1993). The local excitation – lateral inhibition function is also inspired by neurophysiology (e.g., Georgopoulos et al., 1993; for network models

that use related functions, see Compte et al., 2000; Tanaka, 2000).

Recently, several studies have demonstrated that DFT models can effectively move from being neurally inspired to being neurally constrained. In particular, dynamic fields can be directly estimated by measuring the responses of neurons to a stimulus and then lining up the neurons according to their preferred location rather than their location in cortex (e.g., Bastian, Riehle, Erlhagen, & Schöner, 1998; Erlhagen, Bastian, Jancke, Riehle, & Schöner, 1999). For instance, several reaction time predictions of a motor planning model have been recently tested both behaviorally (McDowell, Jeka, Schöner, & Hatfield, 2002) and neurophysiologically (Bastian et al., 1998; Erlhagen et al., 1999), providing strong support for a DFT of movement preparation (Erlhagen & Schöner, 2002; Schöner, Kopecz, & Erlhagen, 1997). Although these ties to neurophysiology are exciting, developmental work in this area is limited by the challenges of conducting developmental primate neurophysiology. Thus, the predictions tested here are necessarily behavioral in nature.

This raises an important issue. We have made a developmental hypothesis that interactions in working memory become more spatially precise over development. If we cannot test this hypothesis directly using neurophysiological techniques, why bother? We see progress in this area proceeding as follows. First, it is necessary to test behavioral predictions that emerge from the implementation of this hypothesis in the model. In this regard, it is important to note that the predictions tested here are strong predictions. Simply put, if we do not find evidence supporting the interaction between age and separation captured in Figure 2, the developmental hypothesis that links age to spatial precision is incorrect.

A second step in this progression toward theory testing and development is to examine how spatial precision changes over development. Recent work in our laboratory has started to examine this issue; however, we view this as a second-order problem: Before one can investigate how spatial precision changes over development, one must first have evidence that it changes at all. This is the purpose of the present study. A third step in theory development is to examine neurophysiological correlates of the spatial precision hypothesis, ultimately testing the spatial precision hypothesis directly using developmental primate neurophysiology. We view behavioral studies as a critical predecessor of these efforts.

Experiment 1

Spencer et al. (2001) reported that 2-year-olds made A-not-B-type errors in a sandbox task when A and B were separated by 8 to 10 in. The first goal of Experiment 1 was to replicate this finding. Toward this end, 2-year-olds participated in an A-not-B sandbox task with A and B separated by 9 in.—the average separation used by Spencer et al. The second goal was to examine developmental changes in A-not-B-type errors. Thus, 4- and 6-year-olds also completed an A-not-B sandbox task with A and B separated by 9 in. We included 6-year-olds in the present experiment because these children make qualitatively different errors than younger children in tasks similar to the sandbox task (Huttenlocher, Newcombe, & Sandberg, 1994; Spencer & Hund, in press). In particular, in tasks with a homogeneous task space, 2- and 3-year-olds show biases toward the midline of the task space. Six-year-olds are biased toward midline in the sandbox task (Huttenlocher et al., 1994) but are biased away from midline in other similar tasks (Huttenlocher et al., 1994; Spencer & Hund, in press). Thus, SWM abilities are changing between 2 and 6 years of age, but will such changes be reflected in the predicted effects? Note that we selected 4-year-olds because they are halfway between the 2-year-olds studied by Spencer et al. (2001) and the 6-year-olds studied by Spencer and Hund (in press).

The third goal of this experiment was to examine whether the memory delay on the A trials is a critical contributor to A-not-B-type errors. Spencer et al. (2001) used a 3-s delay on the A trials and a 10-s delay on the B trials. According to the DFT, the shorter delay on the A trials should create a strong memory of A. Consequently, when SWM is challenged by the long delay on the B trials, there should be a strong bias toward A. We used a 10-s delay on both the A and B trials. This might create a weaker memory of A, raising an important question: Will a weaker memory of A still be strong enough to cause A-not-B-type errors?

One of the challenges of measuring experience-dependent effects in the sandbox task is the presence of a second type of bias: Young children's responses are biased toward the midline axis of the sandbox (Huttenlocher et al., 1994; Spencer et al., 2001). We handled this bias using the experimental design from Spencer et al. (2001) where the A and B locations were always in the same half of the sandbox and the locations of A and B were counterbalanced across conditions.

Method

Participants

Participants were twenty 2-year-olds (M = 2 years, 4.5 months; SD = .82 months; 9 males, 11 females), sixteen 4-year-olds (M = 4 years, 6.6 months; SD = .85 months; 5 males, 11 females), and seventeen 6-year-olds (M = 6 years, 10 months; SD = 1.71 months; 10 males, 7 females). Data were collected from 10 additional participants; however, these data were not analyzed. One 2-year-old did not complete the task, one 4-year-old did not wait for the delay to reach, and data from eight children (two 2-year-olds, one 4-year-old, and five 6-year-olds) were excluded because of experimenter error (e.g., the session was not videotaped, the experimenter included an extra A trial, or the object was hidden in the wrong location). Participants were recruited from a database at the University of Iowa and were given a small gift for participating. The parents of all participants gave informed consent. The majority of the participants were White from middle-class families.

Materials

The sandbox was 60-in. long, 16-in. wide, and 20-in. high. There was a false bottom 6 in. from the top of the sandbox. The sandbox was constructed from 1-in. thick plywood and painted dark blue. It was filled with approximately 4 in. of sterile play sand. White markings were painted every 6 in. on the outside wall of the experimenter's side of the sandbox to help the experimenter identify the hiding locations. The markings could not be seen by the participant. A metal bar ran along the length of the sandbox underneath the sand. The bar was used to keep the hiding locations exact. Holes were drilled through the bar at 1-in. intervals, and wooden pegs were fitted into the holes at the hiding locations. Small toys (approximately 1 in. tall by 1 in. wide by 1 in. deep) that fit over the tops of the pegs were hidden for the 2- and 4-year-olds. Dimes that fit into slots in the top of the pegs were hidden for the 6-year-olds. Note that we used dimes for the 6-year-olds because they presented a very precise target input—the dime could be placed in an exact spatial position along the length of the sandbox. Moreover, pilot testing indicated that this age group was very motivated to perform the task regardless of the characteristics of the target object. Nevertheless, we addressed the possible role of the target objects in Experiment 1b.

A video camera was mounted on the ceiling directly over the sandbox to record each session. To facilitate coding

of data from each session, a grid was placed over the top of the VCR monitor. Before each session, the numbers on the grid were aligned to prespecified locations within the sandbox. Participants' responses were scored from the videotapes following each session using the aligned grid. Curtains hung from the ceiling around the sandbox to eliminate any external landmarks. Four- and 6-year-olds sat on a child-sized chair centered along the length of the sandbox directly opposite the experimenter. The chair was 12 in. from the sandbox. Two-year-olds stood on an outline of feet taped on the floor with the parent seated in a chair directly behind them. The feet were 12 in. from the sandbox.

Procedure

Each trial proceeded as follows: The experimenter placed the object on the peg at the correct location (the A location for the first six trials and the B location for the final three trials), the object was buried such that its location was not marked in any way, the experimenter and child counted to 10 while maintaining eye contact, and then the child was allowed to search for the object. For the 2-year-olds, the first two A trials were training trials. These trials proceeded in the same manner as the others except the toys were not completely hidden. On the first training trial, the toy was placed on the peg at the A location and only half buried. On the second training trial, the toy was buried at the A location so that only the very tip of the toy was visible. For the 4- and 6-year-olds, the object was completely buried on all of the trials.

Experimental Design

Objects were hidden at one of two possible locations in the left half of the sandbox: 15 in. or 24 in. from the left edge (from the child's perspective). The location of A and B was counterbalanced across children such that A was at 15 in. for half of the children and A was at 24 in. for the other half of the children. Children were randomly assigned to each condition.

Behavioral Scoring

All sessions were videotaped and scored at a later time using the grid affixed to the monitor. Scorers coded the first contact with the sand. If a 2-year-old reached with both hands, the hand that touched the sand first was scored. If both hands touched the sand at the same time, the hand closest to the hiding location was scored. This was the most conservative approach because it resulted in the smallest response error. Responses were coded to the nearest .5 in. A second scorer scored 15% of the sessions. The mean deviation (absolute value) between the two scorers was .25 in. ($SD = .55$ in.). The initial scorer's values were used in all analyses.

Method of Analysis

Given that our goal was to test a set of specific predictions in the present study, we did not use the typical analysis of variance (ANOVA) approach to hypothesis testing. Rather, we conducted a set of focused statistical tests designed to answer specific questions. First, we tested whether each age group individually made A-not-B-type errors; that is, did each age group show biases toward the A location on the B trials? We investigated this issue both in terms of mean responses on the B trials and responses on the first B trial. Next, given the potential presence of midline effects, we examined whether each age group individually showed biases toward or away from midline. This issue was particularly relevant on the A trials before there were potentially two factors influencing children's responses—a bias toward or away from midline and a bias toward A. Third, we examined whether there were age-related changes in A-not-B-type biases by conducting pairwise comparisons. The standard approach to investigating an age effect at a given separation would be to conduct an overall ANOVA with all three age groups included. If a significant age effect was found, follow-up pairwise comparisons would be conducted to evaluate the nature of the effect. Given that we predicted a reduction in A-not-B-type biases over age, we opted to conduct the pairwise comparisons directly, thereby reducing the overall number of tests conducted. Although unconventional, these tests addressed the details of the model predictions more directly than the omnibus ANOVA approach. Finally, in Experiments 2 and 3, we conducted pairwise cross-experiment analyses to examine whether the performance of children within a given age group changed across manipulations of target separation.

Results

Analyses of 2-Year-Olds' Performance

Midline biases on the A trials. We began our analyses by examining whether 2-year-olds showed a bias toward midline on the A trials. Data from the final four A trials are shown in Figure 3. As can be seen in the figure, 2-year-olds' responses were biased toward the midline of the sandbox on the A trials in both conditions. A *t* test comparing mean error on the A trials across conditions

(A at 15 in., A at 24 in.) revealed no significant effects, $t(18) = -0.59$, *ns*, $d = .26$. An additional *t* test revealed that 2-year-olds' responses differed significantly from zero error in the direction of the midline of the sandbox, $t(19) = 2.70$, $p < .01$, $d = .61$, ($M = 1.58$ in.).

A-not-B-type biases. Next, we examined whether 2-year-olds' mean errors on the B trials were biased toward A. As can be seen in Figure 3, 2-year-olds made errors toward A on the B trials in both conditions. A *t* test comparing mean errors on the B trials across conditions (A at 15 in., A at 24 in.) revealed a significant difference, $t(18) = -2.55$, $p < .025$, $d = 1.14$. Thus, 2-year-olds made A-not-B-type errors.

To examine A-not-B-type effects at a trial-by-trial level, we compared errors on the last A trial with errors on the first B trial. According to the DFT, children should construct a longer term memory of the A location during the six A trials. Consequently, by trial A6 performance should be relatively accurate. On the next trial (B1), however, responses should be strongly biased toward A. As can be seen in Figure 3, 2-year-olds' errors on the last A trial were biased toward the midline of the sandbox. On the first B trial, however, responses were strongly biased toward A. Two-year-olds' errors on the last A trial

Figure 3. Mean distance errors in Experiment 1 across the A trials (gray bars) and B trials (black bars), and mean distance errors on the last A trial (white bars) and the first B trial (striped bars) for 2-, 4-, and 6-year-olds for each layout of A and B locations. The absolute location in the sandbox is marked along the *x* axis with the hiding locations at 15 and 24 in. Error bars indicate standard deviation.

and first B trial were analyzed in a Condition (A at 15 in., A at 24 in.) × Trial (A6, B1) ANOVA with condition as a between-subjects variable and trial as a within-subjects variable. The ANOVA revealed a significant Condition × Trial interaction, $F(1, 18) = 5.19$, $p = .035$. Additional tests of simple effects indicated that performance across conditions on the last A trial did not differ, $t(18) = -.39$, *ns*, $d = .17$. However, there was a significant effect of condition on the first B trial, $t(18) = -1.93$, $p < .05$, $d = .86$. Thus, as with the mean analyses earlier, 2-year-olds showed A-not-B-type errors on the first B trial, and these errors differed significantly from errors toward midline on the last A trial.

Analyses of 4-Year-Olds' Performance

Midline biases on the A trials. As can be seen in Figure 3, 4-year-olds were biased toward midline on the A trials in both conditions. A *t* test comparing mean error on the A trials across conditions revealed no significant effects, $t(14) = -1.17$, *ns*, $d = .59$, although the relatively large effect size indicates that there was a trend for responses to be more strongly biased toward midline at 24 in. (see Figure 3). An additional *t* test revealed that 4-year-olds' responses differed significantly from zero error in the direction of midline, $t(15) = 2.15$, $p < .05$, $d = .54$, ($M = .43$ in.). Thus, 4-year-olds, like 2-year-olds, showed a significant bias toward the midline of the sandbox.

A-not-B-type biases. Four-year-olds made large rightward errors toward A on the B trials when A was toward the center of the sandbox, and made smaller leftward errors toward A when A was toward the left edge of the sandbox (see Figure 3). A *t* test comparing mean error on the B trials across conditions revealed a significant difference, $t(14) = -2.66$, $p < .025$, $d = 1.33$. Thus, like the 2-year-olds, 4-year-olds made A-not-B-type errors. We compared errors on the last A trial with errors on the first B trial (see Figure 3) by conducting an ANOVA with condition as a between-subjects variable and trial as a within-subjects variable. This analysis revealed a significant Trial × Condition interaction, $F(1, 18) = 11.08$, $p = .005$. Tests of simple effects revealed that errors on the last A trial did not vary significantly across conditions, $t(14) = .24$, *ns*, $d = .12$. Errors on the first B trial, however, did vary significantly across conditions, $t(14) = -3.30$, $p = .005$, $d = 1.65$. Thus, as with the 2-year-olds, 4-year-olds showed A-not-B-type errors on the first B trial, and these errors differed significantly from errors on the last A trial.
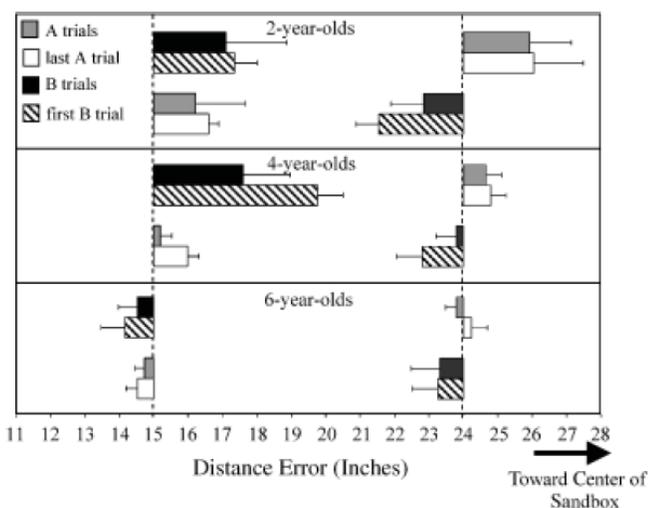
Analyses of 6-Year-Olds' Performance

Midline biases on the A trials. In contrast to the 2- and 4-year-olds, 6-year-olds' responses were biased slightly away from midline on the A trials in both conditions. A *t* test comparing mean errors on the A trials across conditions revealed no significant effects, $t(15) = -.21$, *ns, d* = .10. In addition, a *t* test comparing mean error on the A trials with zero error also was not significant, $t(16) = -1.74$, *ns, d* = .42, ($M = -.24$ in.). Thus, 6-year-olds' responses to A were not significantly biased toward or away from midline.

A-not-B-type biases. As can be seen in Figure 3, 6-year-olds made small errors away from the midline of the sandbox in both conditions. A *t* test comparing mean error on the B trials across conditions was not significant, $t(14) = -.32$, *ns, d* = .16. We also compared errors on the last A trial and first B trial (see Figure 3). A Condition × Trial ANOVA revealed no significant effects. Thus, 6-year-olds did not make A-not-B-type errors.

Across-Age Comparisons

In our final set of analyses, we examined how A-not-B-type errors changed across age groups. First, we compared 2-year-olds' and 4-year-olds' performance on the B trials by conducting an Age (2, 4) Condition ANOVA on the mean error on the B trials. We only report age effects, of which there were none. Thus, 2- and 4-year-olds showed comparable A-not-B-type errors on the B trials. Next, we compared the 2- and 6-year-olds' performance on the B trials. An Age (2, 6) × Condition ANOVA revealed a marginal Condition × Age interaction, $F(1, 33) = 3.99$, $p = .054$. Although marginal, this effect is consistent with the pattern of results reported earlier, that is, the significant A-not-B-type errors for the 2-year-olds and the nonsignificant findings for the 6-year-olds. In our final planned analysis, we compared 4- and 6-year-olds' performance in an Age (4, 6) × Condition ANOVA. There was a significant age main effect, $F(1, 29) = 8.28$, $p = .01$, and a significant Condition × Age interaction, $F(1, 29) = 4.38$, $p < .05$. As with the *t* tests, 4-year-olds showed a significant bias toward A on the B trials across conditions whereas 6-year-olds did not. In summary, pairwise comparisons across age groups indicated that the two youngest age groups made A-not-B-type errors that differed from the performance of the 6-year-olds.

Discussion

Two- and 4-year-olds' responses were significantly biased toward the midline of the sandbox on the A trials, whereas 6-year-olds' responses were not. The 2- and 4-year-olds' results are consistent with Huttenlocher et al.'s (1994) finding that young children are biased toward the midline of the sandbox (see also Spencer et al., 2001). Huttenlocher et al. also reported that 6-year-olds showed a bias toward midline. It is likely that we did not find such a bias here because 6-year-olds responded to the same location six times in a row rather than only once (see Huttenlocher et al., 1994). In contrast to the current results, Spencer et al. (2001) did not find a systematic bias toward midline on the A trials with 2-year-olds. However, Spencer et al. used a 3-s delay whereas we used a 10-s delay. The longer delay used here would allow more time for activation in working memory to drift toward midline, leading to larger biases on the A trials (see Schutte and Spencer, 2002, for similar results).

The primary goal of this experiment was to examine whether there was a reduction in A-not-B-type biases on the B trials across age groups. This was indeed the case. Two- and 4-year-olds' responses were significantly biased toward the A location on the B trials, whereas 6-year-olds' responses were not. Pairwise comparisons across age groups were consistent with these results. The significant A-not-B-type effects with the 2-year-olds replicates results from Spencer et al. (2001) and demonstrates that such effects generalize to the performance of children as old as 4 years of age. The absence of significant A-not-B-type biases with the 6-year-olds is consistent with predictions of the DFT that older children would not make A-not-B-type errors at large separations between A and B. It is possible, however, that 6-year-olds did not make A-not-B-type errors because of a difference in the experimental procedure across age groups; we hid dimes instead of toys for this age group. Results with infants have shown that A-not-B errors depend on the characteristics of the hidden object. The more attractive or salient the object, the less likely it is that infants will make the error (see Thelen et al., 2001). We examined this issue in the next experiment.

Experiment 1b

This experiment was similar to Experiment 1 with two exceptions: We only examined the performance of 6-year-olds and we hid the toys used with the 2- and 4-year-olds instead of dimes. We predicted that this change in object characteristics would not influence 6-year-olds' performance because the dynamic field model used here is not sensitive to such factors (see also Spencer et al., 2001). Specifically, when the DFT operates in the self-

sustaining mode, the target input has little effect during memory delays because: (a) this input is no longer present, and (b) self-sustaining peaks relax to a size dictated by the local excitation–lateral inhibition function.

Method

Participants

Participants were sixteen 6-year-olds (M = 6 years, 10.2 months; SD = 2.73 months; 9 males, 7 females). All other participant details were identical to Experiment 1a.

Materials and Procedure

The materials were the same as in Experiment 1 except that we hid the toys that were used with the 2- and 4-year-olds instead of dimes. The procedures, experimental design, method of analysis, and behavioral scoring were the same as in Experiment 1. All responses were coded to the nearest .5 in., and 15% of the data were scored by a second scorer. The mean deviation (absolute value) between the two scorers was .10 in. (SD = .33 in.).

Results and Discussion

Six-year-olds' errors on the A trials were small. A t test comparing mean errors on the A trials across conditions was not significant, $t(14)$ =.39, ns, d = .20 (A at 15 in.: M = .25, A at 24 in.: M = .08). In addition, errors on the A trials did not differ significantly from zero error, $t(15)$ = .78, ns, d = .20. These data replicate results from Experiment 1. Next, we examined responses on the B trials that were consistently small (A at 15 in.: M = −.38, A at 24 in.: M = .13). A t test comparing mean error on the B trials across conditions was not significant, $t(14)$ = −.84, ns, d = .42. Again, this replicates results from Experiment 1. Finally, we compared errors on the last A trial (A at 15 in.: M = .20, A at 24 in.: M = .20) with errors on the first B trial (A at 15 in.: M = −.99, A at 24 in.: M = −.63). As in Experiment 1, a Condition × Trial ANOVA revealed no significant effects. Therefore, as predicted by the dynamic field model, 6-year-olds did not make A-not-B-type errors when a toy was hidden instead of a dime. Note that this model also predicted that 6-year-olds—like 2- and 4-year-olds—would make A-not-B-type errors at a target separation smaller than 9 in. (see Figure 2). We turn to this prediction in Experiment 2.

Experiment 2

In this experiment we tested the prediction of the DFT that 6-year-olds would make A-not-B-type errors in the sandbox task when the A and B locations were closer than 9 in., in particular, at 6 in. A 6-in. separation was selected to be a large enough change relative to Experiment 1 that we might detect bias but not so large that gradual increases in error across age groups would go undetected. Two- and 4-year-olds also participated in this experiment. We expected these children to make A-not-B-type errors at the smaller separation.

It is possible that 6-year-olds will not make A-not-B-type errors at any separation. Six-year-olds' location memory abilities might differ qualitatively from the abilities of younger children; thus, this age group might no longer be susceptible to such errors. For instance, Marcovitch and Zelazo (1999) proposed that infants make A-not-B errors because they rely on a response system that encodes locations relative to the actions required to reach to these locations. Older children, by contrast, encode locations using an accurate representational system that encodes space allocentrically. Therefore, older children are able to override the tendency to repeat a previous action in A-not-B-type situations (Marcovitch & Zelazo, 1999; Zelazo et al., 1998; for a different account, see Spencer & Schutte, in press). By this view, then, 6-year-olds should not make A-not-B-type errors even at a small spatial separation.

Method

Participants

Participants were twenty-two 2-year-olds (M = 2 years, 4.75 months; SD = 1.11 months; 10 males, 12 females), twenty-seven 4-year-olds (M = 4 years, 4.67 months; SD = 1.03 months; 12 males, 15 females), and twenty 6-year-olds (M = 6 years, 9.75 months; SD = 1.20 months; 10 males, 10 females). Data were collected from nine additional participants; however, these data were not analyzed. Two 2-year-olds and one 6-year-old did not complete the task, one 2-year-old did not wait for the delay to reach for the object, one 4-year-old reached so far from the hidden object most of his reaches were not within the range of the video, and data for four participants (three 2-year-olds and one 6-year-old) were excluded because of experimenter error. All other participant details were identical to Experiment 1.

Materials, Procedure, and Experimental Design

The materials, procedures, method of analysis, and behavioral scoring were the same as in Experiment 1. All

responses were coded to the nearest .5 in., and 15% of the data were scored by a second scorer. The mean deviation between the two scorers was .48 in. (SD = .75). Children were randomly assigned to one of two conditions, with A and B separated by 6 in. In one condition, A was at 18 in. and B was at 24 in.; in the other condition, A was at 24 in. and B was at 18 in.

Results

Analyses of 2-Year-Olds' Performance

Midline biases on the A trials. As can be seen in Figure 4, 2-year-olds' responses on the A trials were biased toward midline. A t test comparing 2-year-olds' mean error on the A trials across conditions was not significant, $t(20)$ = −.49, *ns, d* = .21. An additional t test comparing mean error on the A trials with zero error was significant, $t(21)$ = 1.94, $p < .05, d$ = .41. Therefore, 2-year-olds' responses on the A trials were significantly biased toward midline, although the modest effect size indicates that this was a relatively weak effect.

A-not-B-type biases. As can be seen in Figure 4, when A was toward midline relative to B, 2-year-olds' mean responses on the B trials were strongly biased toward A. When A was toward the left edge of the sandbox, 2-year-olds' mean responses were still slightly biased toward midline, but the bias was much weaker than when A was toward midline. A t test comparing mean er-
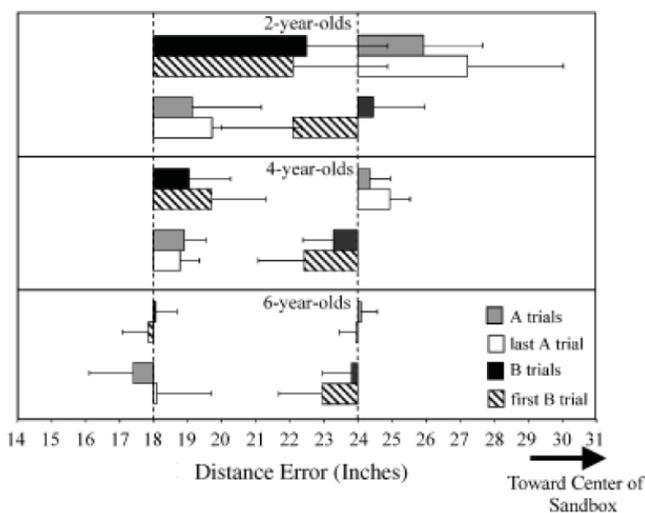

Figure 4. Mean distance errors in Experiment 2 across the A trials (gray bars) and B trials (black bars), and mean distance errors on the last A trial (white bars) and the first B trial (striped bars) for 2-, 4-, and 6-year-olds for each layout of A and B locations. The hiding locations were at 18 and 24 in. Error bars indicate standard deviation.

ror on the B trials across conditions was significant, $t(20)$ = −2.38, $p < .025, d$ = 1.01, replicating the results from Experiment 1.

Two-year-olds' responses on the first B trial were biased toward A in both conditions (see Figure 4). An ANOVA with condition (A at 18 in., A at 24 in.) as a between-subjects variables and trial (A6, B1) as a within-subjects variable revealed a significant main effect of condition, $F(1, 20)$ = 4.54, $p < .05$. Although this effect was clearly driven by the large errors toward A on the first B trial, errors on the last A trial also contributed. In particular, 2-year-olds' bias toward midline was smaller when A was at 18 in. than when A was at 24 in. It is not clear why this is the case.

Analyses of 4-Year-Olds' Performance

Midline biases on the A trials. As in Experiment 1, 4-year-olds' responses on the A trials were biased toward the midline of the sandbox (see Figure 4). A t test comparing errors across conditions was not significant, $t(25)$ = 1.15, *ns, d* = .44, whereas a t test comparing mean error on the A trials with zero error was, $t(26)$ = 2.63, $p < .025, d$ = .51.

A-not-B-type biases. Four-year-olds' mean responses on the B trials were biased toward A in both conditions (see Figure 4). A t test comparing mean error on the B trials across conditions revealed a significant difference, $t(25)$ = −2.14, $p < .05, d$ = .82. Four-year-olds' responses on the last B trial were also biased toward A in both conditions. A Condition × Trial ANOVA revealed a significant condition main effect, $F(1, 25)$ = 8.37, $p < .01$, and a significant Condition × Trial interaction, $F(1, 25)$ = 6.20, $p < .025$. Tests of simple effects showed a nonsignificant condition effect on trial A6, $t(25)$ = −.35, *ns, d* = .13, and a significant condition effect on trial B1, $t(25)$ = −2.89, $p < .01, d$ = 1.11. Thus, as in Experiment 1, 4-year-olds made significant A-not-B-type errors.

Analyses of 6-Year-Olds' Performance

Midline biases on the A trials. As can be seen in Figure 4, 6-year-olds' errors on the A trials were small in both conditions. A t test comparing errors across conditions was not significant, $t(18)$ = −.81, *ns, d* = .36, nor was an additional t test comparing mean error on the A trials with zero error, $t(19)$ = −.58, *ns, d* = .13. Thus, as in Experiment 1, 6-year-olds' responses were not significantly biased toward or away from midline.

A-not-B-type biases. The critical question was whether 6-year-olds would show A-not-B-type biases. As can

be seen in Figure 4, 6-year-olds' responses on the B trials were generally accurate, although there was a slight bias toward A on the first B trial in the A at 18 in. condition. A *t* test comparing mean error on the B trials across conditions was nonsignificant, $t(18) = -.37$, *ns, d = .17*, and a Condition × Trial ANOVA did not reveal any significant effects. Therefore, as in Experiment 1, 6-year-olds' failed to show A-not-B-type biases.

Across-Age Comparisons

As in Experiment 1, we compared errors on the B trials across ages to determine whether there were age-related changes in A-not-B-type effects. We began by comparing 2- and 4-year-olds' mean performance on the B trials. An Age (2, 4) × Condition ANOVA revealed a significant age main effect, $F(1, 45) = 6.71$, $p = .01$. As can be seen in Figure 4, 2-year-olds' errors were generally larger than 4-year-olds' errors, and 2-year-olds' responses were more strongly biased toward midline on the B trials. Two-year-olds' errors also differed significantly from 6-year-olds' errors. An Age (2, 6) × Condition ANOVA revealed a significant age main effect, $F(1, 38) = 7.15$, $p = .01$, and a marginal Age × Condition interaction, $F(1, 38) = 4.01$, $p = .052$. As can be seen in Figure 4, 2-year-olds' errors on the B trials were generally larger than 6-year-olds' errors, particularly in the A at 24 in. condition. The final age comparison between 4- and 6-year-olds' errors—an Age (4, 6) × Condition ANOVA—revealed no significant effects. Although the 4-year-olds made significant A-not-B-type errors and the 6-year-olds did not, the difference across age groups was not robust.

Cross-Experiment Comparisons

In our final set of planned analyses, we examined whether there were significant differences in the performance of children within each age group as the separation between A and B was reduced from 9 in. (Experiment 1) to 6 in. (Experiment 2). Given that we found comparable results for each age group across experiments, we did not expect these comparisons to be very informative. An Experiment (1, 2) × Condition (A to the right of B, A to the left of B) ANOVA of the 2-year-olds' mean error on the B trials revealed no significant effects of experiment. Similarly, an Experiment × Condition ANOVA for the 4-year-olds' responses showed no significant effects of experiment. Finally, an Experiment × Condition ANOVA for the 6-year-olds' mean

errors on the B trials revealed no significant effects of experiment. Therefore, across all three age groups, response errors did not change significantly when the separation changed from 9 in. to 6 in.

Discussion

The main question of interest in the present experiment was whether 6-year-olds would make A-not-B-type errors at a small separation between A and B as predicted by the DFT, or whether the 6-year-olds' responses would be immune from such effects as predicted by some accounts of the A-not-B error (e.g., Marcovitch & Zelazo, 1999; Zelazo et al., 1998). As in Experiment 1, the 6-year-olds' responses on the B trials were not biased toward A. These effects are generally consistent with the response-to-representation shift described by Marcovitch and Zelazo (1999; see also, Zelazo et al., 1998). Nevertheless, it is possible that a 6-in. separation is still too large for 6-year-olds to make A-not-B-type errors. Thus, in the next experiment we once again decreased the separation between A and B.

Experiment 3

To test the prediction that 6-year-olds would make A-not-B-type errors at a small separation, we decreased the separation between A and B to 2 in. Given that 4- and 6-year-olds' errors in the previous experiments were generally around 1 in. to 2 in., we reasoned that this was the smallest separation that would still yield reliable effects. Four-year-olds also participated in this experiment. We expected these children to show robust A-not-B-type errors. We did not include 2-year-olds, however, because we thought a 2-in. separation might be too challenging for this age group.

In addition to investigating 4- and 6-year-olds' performance, we further explored the developmental trajectory of A-not-B-type errors by examining whether older children (11-year-olds) and adults show similar errors in the sandbox task. Although we predicted that SWM is becoming more precise over development, it is possible that these older age groups will make A-not-B-type errors at a very small separation. Indeed, based on our analysis of the DFT in Figure 2, we expected that there would be an age-related trend at this separation similar to what we found in Experiment 1. Specifically, 4- and 6-year-olds should show robust A-not-B-type errors, 11-year-olds might show weaker biases, and adults might be consistently accurate in this task.

Method

Participants

   Participants were twenty 4-year-olds (*M* = 4 years, 6.6 months; *SD* = 2.0 months; 12 males, 8 females), twenty-one 6-year-olds (*M* = 6 years, 6.0 months; *SD* = 2.8 months; 13 males, 8 females), twenty 10- to 11-year-olds (*M* = 10 years, 10.1 months; *SD* = 6.5 months; 12 males, 8 females), and twenty adults (*M* = 35 years, 8.5 months; *SD* = 9.4 years; 5 males, 15 females). Data were collected from 11 additional participants; however, these data were not analyzed. Data from 5 participants (three 6-year-olds, one 11-year-old, and one adult) were excluded from the final analyses because of experimenter error. In addition, data from one 4-year-old were excluded because of a video equipment failure, one 4-year-old and two 6-year-olds reached out of the range of the video camera, one 6-year-old's responses were uncodable (the child's head blocked the view of the camera), and one 11-year-old's session was interrupted by the parent. Adult participants were either the parents of children who were participating in various studies in the lab or from the undergraduate population. All adults gave informed consent. All other participant details were as in the previous experiments.
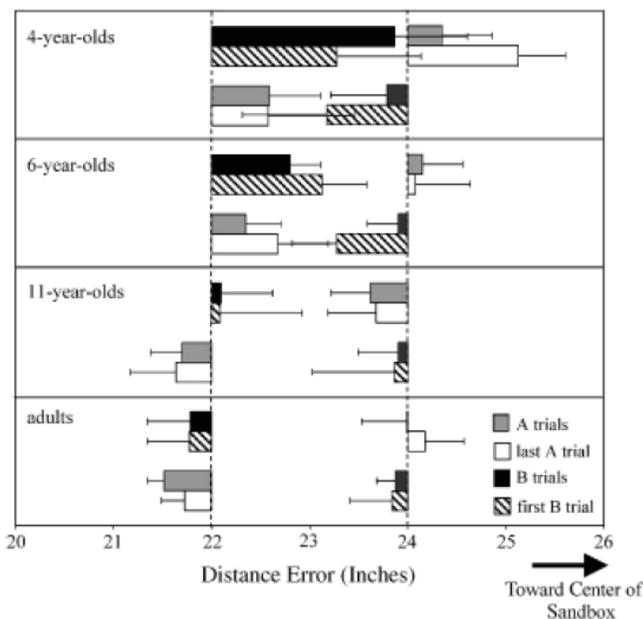

Figure 5. Mean distance errors in Experiment 3 across the A trials (gray bars) and B trials (black bars), and mean distance errors on the last A trial (white bars) and the first B trial (striped bars) for 4-, 6-, and 11-year-olds and adults for each layout of A and B locations. The hiding locations were at 22 and 24 in. Error bars indicate standard deviation.

Materials, Procedure, and Experimental Design

   The materials and procedure were the same as in Experiment 1. The behavioral scoring was the same except that responses were coded to the nearest .25 in. A second coder coded 15% of the data. The mean deviation between the two scorers was .03 in. (*SD* = .08). Participants were randomly assigned to one of two conditions, with A and B separated by 2 in. In one condition, A was at 22 in. and B was at 24 in., and in the other condition, A was at 24 in. and B was at 22 in.

Results

Analyses of 4-Year-Olds' Performance

   Midline biases on the A trials. As can be seen in Figure 5, 4-year-olds' responses on the A trials were biased toward the midline of the sandbox. A *t* test comparing errors across conditions was not significant, *t*(18) = .52, *ns, d* = .23. A second *t* test comparing mean error on the A trials with zero error established that 4-year-olds' responses were significantly biased toward midline, *t*(19) = 2.08, *p* < .05, *d* = .47, although the modest effect size indicates that this was a relatively weak effect. Nevertheless, these results replicate the results of Experiments 1 and 2.

   A-not-B-type biases. Four-year-olds made errors toward A on the B trials in both conditions (see Figure 5). A *t* test comparing mean error on the B trials across conditions confirmed that there was a significant effect of condition, *t*(18) = −3.53, *p* < .01, *d* = 1.58. Thus, as in Experiments 1 and 2, 4-year-olds made significant A-not-B-type errors. Four-year-olds also made large errors toward A on the first B trial in both conditions. A Condition (A at 22 in., A at 24 in.) × Trial (A6, B1) ANOVA revealed a significant main effect of condition, *F*(1, 18) = 6.58, *p* < .025. Although this effect was clearly driven by the large errors toward A on the first B trial, errors on the last A trial also contributed. In particular, as with the 2-year-olds in Experiment 2, the bias on A6 was smaller when A was at 22 in. than when A was at 24 in. It is possible that these effects on Trial A6 across experiments are related to the distance from midline: Locations farther from midline might be pulled toward this location less strongly. Regardless, the key issue for the purposes of the present experiment is the clear A-not-B-type effect on Trial B1.

Analyses of 6-Year-Olds' Performance

   Midline biases on the A trials. As can be seen in Figure 5, 6-year-olds made small errors toward the mid-

line of the sandbox on the A trials. A *t* test comparing the mean error on the A trials across conditions was not significant, $t(18) = .59$, *ns, d* = .23. A second *t* test revealed that 6-year-olds' errors toward the midline of the sandbox were not significantly different from zero error, $t(19) = 1.47$, *ns, d* = .33. These data replicate the results of Experiments 1 and 2.

A-not-B-type biases. Six-year-olds made errors toward A on the B trials in both conditions; however, these errors were largest when A was toward the midline of the sandbox relative to B. A *t* test revealed that mean errors on the B trials varied significantly depending on condition, $t(18) = -3.22$, *p* = .005, *d* = 1.44. Six-year-olds' responses on the first B trial were also biased toward A in both conditions. A Condition × Trial ANOVA revealed a significant Condition × Trial interaction, $F(1, 18) = 12.64$, *p* < .01. Tests of simple effects indicated that there was a not a significant effect of condition on the last A trial, $t(18) = 1.25$, *ns, d* = .56, although the relatively large effect size indicates that the larger biases toward midline at 22 in. approached significance (see Figure 5). In contrast to responses on the A trials, there was a significant effect of condition on the first B trial, $t(18) = -3.48$, *p* < .005, *d* = 1.56. In both conditions, errors on the first B trial were biased toward A. Thus, in contrast to Experiments 1 and 2, 6-year-olds made significant A-not-B-type errors when A and B were separated by 2 in.

Analyses of 11-Year-Olds' and Adults' Performance

Midline biases on the A trials. Eleven-year-olds' errors on the A trials were biased slightly away from the midline of the sandbox (see Figure 5). A *t* test comparing mean error across conditions established that errors on the A trials did not depend on the location of A in the sandbox, $t(18) = .26$, *ns, d* = .12. A second *t* test comparing mean error with zero error revealed that 11-year-olds' responses on the A trials were significantly biased away from midline, $t(19) = -2.17$, *p* < .025, *d* = .48. These data replicate results from Huttenlocher et al. (1994) with 10-year-olds. Adults' responses on the A trials were generally accurate (see Figure 5). A *t* test comparing mean error on the A trials across conditions was not significant, $t(18) = -1.53$, *ns, d* = .68, although the relatively large effect size indicated that there was a trend toward larger biases away from midline at 22 in. (see Figure 5). Similarly, a *t* test comparing adults' mean error with zero error was not significant, $t(19) = -1.56$, *ns, d* = .35.

A-not-B-type biases. Eleven-year-olds' responses on the B trials were generally accurate, although there was a slight bias toward A in each condition (see Figure 5). A *t* test comparing errors on the B trials across conditions was not significant, $t(18) = -.47$, *ns, d* = .21. Furthermore, a Condition × Trial ANOVA did not reveal any significant effects on the last A trial and first B trial (see Figure 5). Adults' responses on the B trials were also accurate (see Figure 5). A *t* test comparing mean error on the B trials across conditions was not significant, $t(18) = .29$, *ns, d* = .13. Similarly, a Condition × Trial ANOVA showed no significant effects. Thus, neither the 11-year-olds nor the adults made significant A-not-B-type errors.

Across-Age Comparisons

In our next set of analyses, we conducted pairwise comparisons across age groups to determine whether A-not-B-type biases became significantly weaker over age. A Condition × Age (4, 6) ANOVA showed no significant age-related differences in the performance of 4- and 6-year-olds. However, a second Condition × Age (4, 11) ANOVA revealed a significant effect of age, $F(1, 36) = 5.20$, *p* < .05, and a significant Condition × Age interaction, $F(1, 36) = 6.78$, *p* < .025. Thus, 4-year-olds' A-not-B-type biases differed significantly from the performance of 11-year-olds. This was also the case in comparisons with adults' performance. A third Condition × Age (4, adults) ANOVA revealed a significant age main effect, $F(1, 36) = 9.06$, *p* = .005, and a significant Condition × Age interaction, $F(1, 36) = 10.72$, *p* < .01.

We also compared 6-year-olds' responses on the B trials with 11-year-olds' and adults' responses. A Condition × Age (6, 11) ANOVA indicated that 6-year-olds' mean errors on the B trials did not differ significantly from 11-year-olds' errors. However, an additional Condition × Age (6, adults) ANOVA revealed a significant age main effect, $F(1, 36) = 6.50$, *p* < .025, and a significant Condition × Age interaction, $F(1, 36) = 5.79$, *p* < .025. As can be seen in Figure 5, 6-year-olds' errors on the B trials were biased toward A, whereas adults' responses were biased away from the midline of the sandbox. A final Condition × Age (11, adults) ANOVA revealed no significant differences between 11-year-olds' and adults' responses.

Taken together, results from these age comparisons reveal a gradual change in A-not-B-type effects over development. Specifically, 4- and 6-year-olds' responses did not differ, but the 4-year-olds' responses differed

significantly from the responses of 11-year-olds and adults. Likewise, 6-year-olds' responses did not differ significantly from the responses of the next closest age group—the 11-year-olds. Their responses did differ significantly, however, from the responses of the adult participants.

Cross-Experiment Comparisons

In a final set of analyses, we examined whether 4- and 6-year-olds' responses differed significantly across manipulations of the separation between A and B (i.e., across experiments). Four-year-olds' mean error on the B trials were analyzed in an Experiment (1, 3) × Layout (A to the left, A to the right) ANOVA. There were no significant effects. We also compared 4-year-olds' errors on the B trials across Experiments 2 and 3. Once again, an Experiment (2, 3) × Layout ANOVA revealed no significant effects. Thus, 4-year-olds made robust A-not-B-type errors across separations of 2 in. to 9 in.

The more critical comparisons examined 6-year-olds' performance across experiments. An Experiment (1, 3) × Layout ANOVA revealed a significant main effect of experiment, $F(1, 33) = 7.24$, $p = .01$. Six-year-olds did not make significant A-not-B-type errors when A and B were separated by 9 in., but they did make significant errors when A and B were separated by 2 in. We also compared results from the present experiment with those of Experiment 2. This Experiment (2, 3) × Layout ANOVA revealed no significant effects, suggesting that A-not-B-type biases changed gradually as separation was varied from 2 in. to 6 in.

Discussion

The primary goal of this experiment was to determine whether 6-year-olds would make robust A-not-B-type errors at a 2-in. separation. This was indeed the case. In contrast to accounts that suggest that older children's spatial memory abilities differ qualitatively from the abilities of toddlers (e.g., Marcovitch & Zelazo, 1999; Zelazo et al., 1998), both 4- and 6-year-olds showed significant biases toward the A location on the B trials. These effects were prevalent in analyses of mean error on the B trials and in analyses of error on Trial B1. Moreover, there was a systematic reduction in A-not-B-type biases across age. Finally, of critical importance to tests of the DFT, 6-year-olds' mean errors on the B trials differed significantly across experiments. In particular, 6-year-olds failed to make A-not-B-type errors at a large separation (Experiment 1), but did make A-not-B-type errors at a small sep-

aration (Experiment 3). These results are consistent with recent data showing that this age group is influenced by a longer term memory of locations (Hund & Spencer, 2003; Spencer & Hund, in press).

Quantitative Modeling

To date, our tests of the DFT have generally emphasized qualitative tests of the model (e.g., Schutte & Spencer, 2002; for exceptions to this trend, see Erlhagen & Schöner, 2002; Thelen et al., 2001). Although this can provide strong tests of specific predictions, it is also useful to examine quantitative fits. Such fits can demonstrate that there is nothing in the model to preclude such a fit, that is, that the model can behave—in detail—like children do. Moreover, quantitative fits place some constraints on parameter settings for future tests of the model. For instance, we would not want to find one parameter setting that fit the data effectively for the present circumstances and, then, in the next study, start from scratch. Rather, the longer term goal would be to use the set of parameters identified here to generate new predictions.

In this section we describe simulations of the model that quantitatively fit the results of Experiments 1 through 3. Our specific goal was to simulate children's responses on the first B trial (B1). We also modeled the variability (standard deviations) of children's responses. Conceptually, the variability of the model's performance represents within-subject variability. Given that we only had measures of between-subjects variability in the present study, we were forced to assume that between-subjects variability was a reasonable approximation of within-subject variability. Although this placed limits on how strongly we could push the model, fitting the variability of children's responses provided a useful index of how realistically the model behaved.

It is worth noting two other caveats to our quantitative modeling efforts. First, we did not attempt to simulate the trial-to-trial details of children's performance. This simplified the modeling task, allowing us to use a static longer term memory of the A location rather than the dynamically constructed longer term memory used by Thelen et al. (2001). Nevertheless, a static longer term memory was sufficient to capture the critical results germane to the prediction tested here. Second, we did not simulate midline effects. Simulations by Schutte and Spencer (2002) demonstrated that the inclusion of a midline input can effectively capture variations in A-not-B-type biases depending on the spatial layout of A and B in the task space.

Method

To simulate the developmental differences reported here, we varied the spatial precision of the interaction function that governs neural activation in the DFT (for a complete description of the model equations, see the Appendix). This function had a Gaussian form:

$$w(x - x') = \frac{w_e}{\sqrt{2\pi}} \left[\exp\left(-\frac{(x - x')^2}{2\sigma_w^2} \right) - w_i \right].$$

The dotted curve in Figure 6 shows one example of an interaction function used in the present simulations. As can be seen in this figure, the spatial precision of local excitatory interactions was given by σ_w, the strength of lateral inhibition was specified by the inhibitory parameter w_i, and the overall strength of interaction was given by the excitatory scaling parameter w_e. To capture developmental changes in A-not-B-type effects, we varied σ_w, the spatial precision of interaction. According to the hypothesis proposed by Spencer and Hund (in press), interactions should be more precise (i.e., narrower) later in development. However, σ_w not only affects the spatial range across which neurons interact, it also affects the stability of activation in the self-sustaining state. In particular, as σ_w becomes smaller and peaks become narrower, the total excitatory energy in a self-sustaining peak is reduced. This can lead to less stable peaks at small values of σ_w, which is not consistent with the proposal that activation is more stable later in development (see Spencer & Hund, in press).
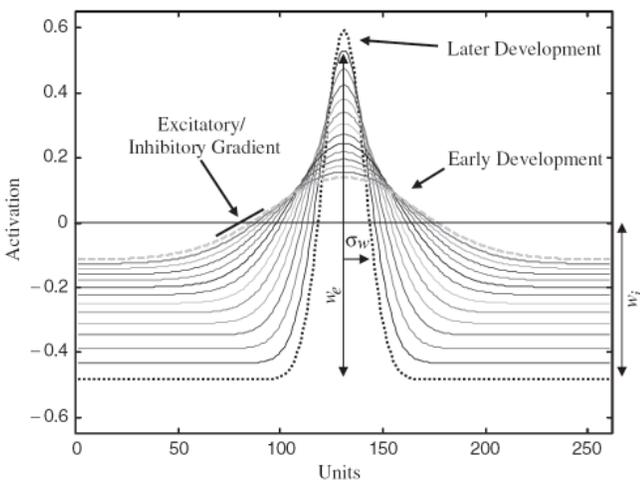

Figure 6. Modulation of interaction function over development.

What factors influence the stability of self-sustaining peaks? A key contributor to stability is the gradient as the interaction function shifts from excitatory to inhibitory (i.e., when this function crosses zero). The sharper the gradient, that is, the stronger the differential between excitation and inhibition, the more stable the peak. The gradient for the dashed curve in Figure 6—an interaction function from early development—is captured by the excitatory–inhibitory gradient line. As can be seen in the figure, the slope of this line is shallow. Consequently, self-sustaining peaks arising from this interaction function will be relatively unstable. By contrast, the gradient at the zero crossing for the dotted curve—an interaction function later in development—is sharper. Consequently, self-sustaining peaks arising from this form of interaction will be more stable. Although the excitatory – inhibitory gradient can be influenced by a variety of parameters, we modulated this gradient by manipulating the scaling parameter w_e.

To bring together developmental changes in spatial precision and stability, we coupled variations in the spatial precision parameter σ_w to changes in the scaling parameter w_e. In particular, we was an exponentially decaying function of σ_w according to the following equation:

$$w_e = \frac{S_e}{\exp\left[\frac{\sigma_w}{\alpha_e}\right]},$$

where S_e specifies the overall strength of the developmental modulation in the excitatory – inhibitory gradient, and α_e specifies the steepness of the exponential modulation over development. Note that S_e and α_e were constants across the set of simulations reported here. Thus, to capture developmental changes in SWM, we changed only a single parameter, σ_w.

Figure 6 shows the interaction functions resulting from variations in σ_w across the range of parameter values examined in the following simulations. Two critical changes in interaction are apparent. As the interaction functions move from early development (dashed line) to later development (dotted line), the spatial precision of interaction narrows, and the excitatory–inhibitory gradient becomes steeper. This results in relatively unstable self-sustaining peaks early in development that are sensitive to A-not-B-type effects across a broad range of separations between A and B, as well as stable self-sustaining peaks later in development that are only sensitive to A-not-B-type effects at narrow separations.

To capture the variability of children's responses, spatially correlated noise was added to the working memory field. Noise is needed for three reasons. First, noise causes variability in activation patterns in working memory. Second, the addition of noise mirrors the noisy nature of the neural system we are trying to model—the brain of a young child. Third, the stochastic perturbations generated by noise probe the stability of activation patterns in working memory. This allowed us to determine whether the model shows the same developmental change in the stability of working memory evident in children's responses. It is important to note that we used the same noise parameters in all of the simulations. Thus, any changes in variability over development produced by the model resulted from changes in the stability of working memory rather than from an explicit increase in the amount of noise.

Model Parameters

Model parameters were selected based on five constraints. First, we selected parameters such that a self-sustaining peak of activation was generated when the target input was presented. Second, we varied only a single parameter over development (σ_w). Thus, the only difference between the age groups was the value of this parameter. Third, we used a single parameter setting for each age group across manipulations of the separation between A and B with the exception of the position of the target input. Fourth, perceptual inputs (target input) were stronger and more spatially precise than inputs from remembered information (memory input). Finally, several general constraints were provided by the model. The dynamic field equations capture strong theoretical assumptions, such as localized input, a homogeneous and symmetrical interaction function, and superposition of inputs. Thus, the internal structure of the DFT precludes simulation of any arbitrary input–output relationship. It is not the case, therefore, that the DFT can account for any pattern of results. This is why the predictions tested in the present manuscript are strong predictions. Moreover, this is why quantitative fits of the model to data provide a useful check of the relationship between the theory and our empirical results.

General model parameters. The resting level of the spatial working memory field, h, was set to –4.0 (see the Appendix for model equations). This value must be negative and sufficiently close to zero to allow for the generation of self-sustaining peaks (see Amari, 1977; Thelen et al., 2001). The noise strength, q, was set to 0.25, and the

spatial spread of the noise kernel, σ_{noise}, was set to 4. The time scale of the working memory field, τ, was set to 75. This value, combined with the temporal sampling factor, $\Delta T = 15$ (see the following), determined the time scale of the simulations: one time step = 15 ms. Finally, there were 261 neurons or units in the working memory field, which resulted in a spatial scale of four units to 1 in.

Inputs to the model. Given the spatial scale of the simulations, x_{tar}—the location of the target on the first B trial—varied from 8 to 36 units (2 in. to 9 in.) in increments of two units (.5 in.). The strength of the target input, S_{tar}, was set to 8.0. This strength was sufficient to establish a self-sustaining peak at all values of the interaction function. Moreover, $S_{tar} \gg S_{mem}$ to reflect the greater salience of perceptually specified information relative to remembered information. The third parameter, σ_{tar}, was set to 10, the lower limit of spatial precision given the field size (261), the spatial scale of the simulations, and numerical estimation constraints. Finally, the target input was turned on at 2,100 ms and turned off at 4,100 ms to reflect the presentation and hiding of the target object. For the longer term memory input, x_{mem} was fixed at 0—the A location in the simulations. The strength of the memory input, S_{mem}, was set to 0.065, and the spatial precision, σ_{mem}, was set to 12. Thus, $\sigma_{mem} > \sigma_{tar}$ to reflect the poorer spatial resolution of remembered information relative to perceptually specified information.

Interaction function. The slope of the threshold function, β, was set to 5.0. This value creates a relatively sharp activation threshold, allowing for the strong nonlinearities needed to enter the self-sustaining state. The spatial precision of the interaction function, σ_w, varied from 10 to 32 units in increments of 1 unit to capture changes in performance over development. As mentioned earlier, 10 units is the lower limit of spatial precision given the field size, and 32 units was sufficiently large to: (a) remain stable within the self-sustaining state, and (b) simulate the desired developmental effects. The inhibitory parameter, w_i, was set to 0.45. Finally, the two parameters that modulated the excitatory scaling parameter w_e as a function of σ_w were as follows: $S_{e} = 4.7$ and $\alpha_{e} = 18$. The range of interaction functions used is shown in Figure 6.

Simulation Details

Simulations were conducted in MATLAB (Mathworks Inc.) on a PC with a 1.7GHz Pentium 4 processor (the MATLAB code is available from the authors on request). The dynamic field equation was integrated using the Euler method with one time step = 15 ms. This

time step was sufficiently small so that numerical solutions approximated real solutions of the dynamical system given a target presentation of 2 s and a memory delay of 10 s.

Each run of the model simulated a single B1 trial. The trial began with a 100-ms relaxation period to allow the field to reach a stable resting state. This was followed by 2-s prestimulus interval, a 2-s target presentation, and a 10-s memory delay. At the end of the delay, we identified the spatial position associated with maximal activation. This position was converted to inches to indicate the model's response error on that trial. We conducted 50 repetitions of Trial B1 for each value of σ_w and for each separation between A and B. This allowed us to examine the model's mean response error and the standard deviation of responses. These values were compared with the empirical results of Experiments 1 through 3. Note that although 50 simulations are far fewer than the number used by Thelen et al. (2001; 500 simulations), we verified that the model converged on similar values across sets of simulations.

Results and Discussion

Table 1 shows the mean response error of children on Trial B1 across Experiments 1 through 3, as well as the standard deviation of their responses. In addition, Table 1 shows the means and standard deviations of the model on Trial B1 when A and B were separated by 9 in. (Experiment 1), 6 in. (Experiment 2), and 2 in. (Experiment 3) for three values of σ_w that fit children's data relatively well. The three values of the spatial precision variable (σ_w) selected were 32 units for the 2-year-olds, 29 units for the 4-year-olds, and 11 units for the 6-year-olds. Note

that these values decrease over development as dictated by the spatial precision hypothesis. It is also important to note that the values decrease nonlinearly. This suggests that there is a larger change in SWM between 4 and 6 years of age than between 2 and 4 years of age.

As can be seen in Table 1, the model provided a good quantitative fit to children's performance across ages and experiments. The largest deviation between the mean responses of the children and the model was 0.4 in., with a root mean square error of 0.3 in. The model generally underestimated the variability of children's responses, particularly for the 2-year-olds. The largest deviation between the children's standard deviations and the *SD*s of the model was 2.2 in., with a root mean square error of 1 in. Despite this underestimation, the model is clearly within a range of acceptable performance, particularly given that we did not have a direct measure of within-subject variability. Moreover, the model effectively captured developmental changes in variability even though we used the same noise parameters in all of our simulations. These developmental changes arise from the changing interaction function: As interactions become more spatially precise, they also become more stable (see Figure 6).

It is important to emphasize that the results in Table 1 reflect the manipulation of a single developmental parameter. The only other factor manipulated in these simulations was the separation between A and B. Moreover, we obtained the results in Table 1 while still satisfying the constraints for parameter selection described earlier. In this context, therefore, we contend that the model provided an excellent quantitative fit to children's performance.

Table 1

Comparison of Children's Mean Response Error on Trial B1 and the Mean Response Error of the Dynamic Field Model

Group	9-in. separation (Experiment 1)	6-in. separation (Experiment 2)	2-in. separation (Experiment 3)
	2-year-olds		
Children	−2.4″ (5.9″)	−3.0″ (5.7″)	
Model ($\sigma_w = 32$ units)	−2.0″ (4.9″)	−2.6″ (3.5″)	−1.2″ (3.2″)
	4-year-olds		
Children	−2.9″ (4.6″)	−1.7″ (3.4″)	−1.1″ (2.0″)
Model ($\sigma_w = 29$ units)	−2.5″ (3.4″)	−1.9″ (3.4″)	−1.4″ (2.6″)
	6-year-olds		
Children	0.1″ (2.3″)	−0.4″ (2.0″)	−0.9″ (1.5″)
Model ($\sigma_w = 11$ units)	−0.1″ (1.7″)	−0.6″ (1.3″)	−0.8″ (1.5″)

Note. Standard deviations are in parentheses.

Given this excellent quantitative fit, it is useful to come full circle and examine the landscape of the model's performance relative to the simulations presented in Figure 2. Recall that in our initial discussion of the spatial precision hypothesis, we presented an analysis of the dynamic field model showing an interaction between the spatial precision of interactions and the separation between A and B. Figure 7 shows a comparable landscape, but it is now constructed with a quantitatively precise parameter set (the three values of σ_w selected in Table 1 are marked with bold lines). Note that, as before, this landscape was constructed by simulating the model without noise. This provides a clear picture of the model's deterministic biases.

As in the previous analysis of the model, there is a clear interaction between age (i.e., spatial precision) and the separation between A and B (see Figure 7). In contrast to Figure 2, however, the magnitude of bias toward A, that is, the height of the landscape, also varies systematically with age. This captures the second aspect of the spatial precision hypothesis—that activation in working memory becomes more stable over development. This re-

sults in smaller biases toward A for the 6-year-olds than for the 2- or 4-year-olds. Given infinite time, however, the biases toward A shown in Figure 7, would eventually settle at the A location, as shown in Figure 2. Note that the landscape in Figure 7 shows a small amount of bias toward A for narrow interaction widths at large separations. As with the simulations shown in Figure 2, this reflects a slight numerical imprecision. In particular, at small interaction widths, the difference in activation between the site centered at B and the site one unit to the left of B was 0.0004. This produced a one-unit bias in the model, which resulted in the 0.25-in. bias toward A shown in Figure 7. Finally, note that we have simulated a broad range of separations between A and B. This highlights some of the predictive possibilities of the model. For instance, the model predicts that 2-year-olds will stop making A-not-B-type errors at a 15- to 16-in. separation.

General Discussion

This study tested the prediction of the DFT that A-not-B-type biases in the sandbox task depend on both target separation and age. Data from three experiments were
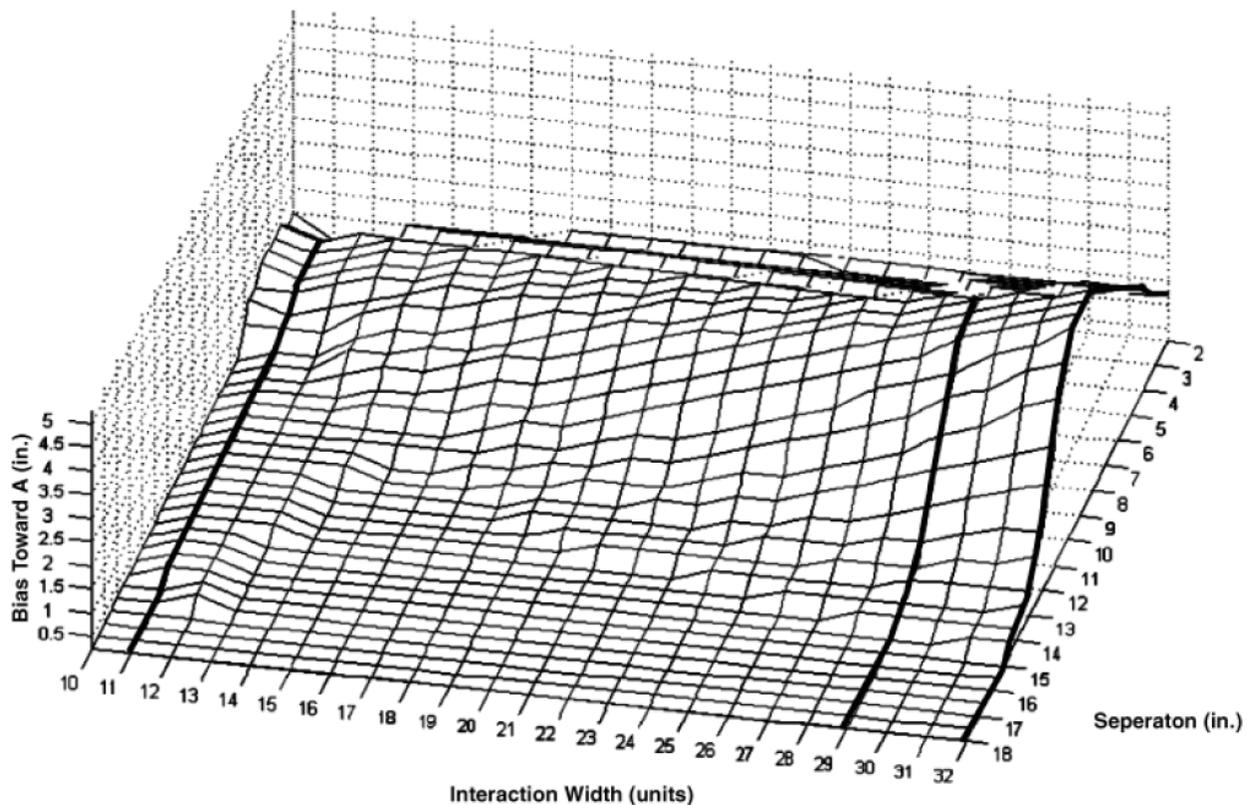

Figure 7. The amount of bias toward A (z axis) on the first B trial for different widths of neuronal interaction (x axis) and separations between A and B locations (y axis). Note that separation and the bias toward A are in inches, and the interaction width are in units. See text for further details.

consistent with this prediction. In Experiments 1 and 2 when A and B were separated by 9 in. and 6 in., 2- and 4-year-olds made A-not-B-type errors, but 6-year-olds did not. When we moved the A and B locations closer to 2 in., however, 4- and 6-year-olds made A-not-B-type errors, whereas 11-year-olds and adults were generally accurate. These results are consistent with the analysis of the DFT presented in Figures 2 and 7, and support the proposal by Spencer and colleagues (Spencer & Hund, in press; Spencer & Schöner, 2000) that neuronal interactions in SWM become more spatially precise over development. Furthermore, simulations of the DFT demonstrated that the model could quantitatively fit the data. The sections that follow discuss the implications of the present results for the DFT, as well as for other models of children's spatial memory abilities.

Implications for the DFT

The DFT predicted a priori that the separation between A and B and the age of the child would interact to create A-not-B-type effects in the sandbox task. It is important to note that this prediction is counter to most other accounts of the A-not-B error. The majority of other accounts suggest that A-not-B errors are either isolated to infancy or are caused by processes that differ qualitatively from the processes that underlie memory for locations in early childhood (e.g., Marcovitch & Zelazo, 1999; Wellman, Cross, & Bartsch, 1987; Zelazo et al., 1998). Our results show that this is not the case. Rather, there is developmental continuity in experience-dependent effects between 2 and 6 years. This extends previous findings from Spencer et al. (2001) and provides strong support for claims by these researchers that the general processes that underlie A-not-B effects in infancy are indeed general.

Although there is continuity across development in the processes that underlie experience-dependent effects, there are also important changes in how such effects arise in the model over development. As mentioned previously, the simulations analyzed in the present study used a dynamic field model operating in the self-sustaining mode, whereas Thelen et al. (2001) used a dynamic field model operating in an input-driven mode. These mechanistic differences are important because they capture three developmental effects. First, Thelen et al.'s account explains why 8- to 10-month-old infants make the A-not-B error in the canonical A-not-B task. Second, this account explains why 12-month-old infants stop making the A-not-B error in the canonical task—these infants are

able to actively sustain a memory of the B location. It is important to note that the task input at B (i.e., the lid) plays a central role in this account: The task input keeps the self-sustaining peak centered at the B location. In the absence of task input, self-sustaining peaks can drift during memory delays (Hund & Spencer, 2003; Schutte & Spencer, 2002; Spencer & Hund, 2002, in press). This is the third type of developmental effect captured by the DFT and was the focus of the present study.

Differences between the Thelen et al. (2001) account and the model tested here are also informative in the context of the proposal that SWM becomes more spatially precise over development. To account for developmental changes in the A-not- B task, Thelen et al. manipulated a parameter in the model (*h*) that controls the resting level of neuronal interaction. By increasing the value of this parameter, these researchers changed the amount of inhibition in the model, making it easier to form a local peak of excitation. Thus, Thelen et al. changed the balance between local excitation and lateral inhibition over development—the same characteristic examined here. This provides a conceptual bridge between the present study and the dynamic field account of the A-not-B error in infancy. It is important that both manipulations of the model point toward an enhancement of interactions in SWM over development.

Although the data presented here are consistent with this developmental proposal, the DFT, at present, does not identify the factors that cause interactions in SWM to change over development. One possibility is that these changes are related to the development of prefrontal cortex. Recent neurophysiological evidence demonstrates that the sustained activation of populations of neurons in prefrontal cortex underlies some aspects of SWM performance. For instance, when nonhuman primates hold the location of a target stimulus in mind, neurons in prefrontal cortex sustain target-specific activation during the delay period, even when intervening stimuli are presented during the memory delay (e.g., Wilson et al., 1993). Moreover, performance on location memory tasks such as the A-not-B task has been linked to the development of dorsolateral prefrontal cortex (Diamond, 1990; Diamond & Doar, 1989). The prefrontal cortex is also one of the last regions of the cortex to develop and is not fully mature until adolescence (e.g., Huttenlocher, 1979, 1990). Therefore, it is possible that both early (i.e., infancy) and later (6 – 11 years) changes in experience-dependent effects are linked to the properties of prefrontal cortical areas.

Although we have focused primarily on the benefits gained by enhanced spatial precision in later development, it is worth noting the reverse—that broad interactions in SWM early in development might be adaptive as well. For example, less stability early in development can keep children from locking on to particular ways of behaving when they are still learning what is relevant in the environment. Broad interactions in working memory would also produce fewer categories of information, potentially simplifying the processing demands of a task. In this way, the DFT takes on a less-is-more characteristic similar to theoretical ideas generated in other domains of development (Munakata & McClelland, 2003; Newport, 1990).

A final question that we have not directly addressed is whether the dimension of reachable locations in the model reported here is grounded in an egocentric or allocentric frame of reference. Thelen et al. (2001) described their dynamic field model as a motor planning field, thereby linking responses of the model to an egocentric representation of space. In the sandbox task, however, children appear to represent space allocentrically. For example, Newcombe, Huttenlocher, Drummey, and Wiley (1998) found that children as young as 22 months could use external landmarks to help them remember a location in the sandbox. In addition, Huttenlocher et al. (1994) reported that children's errors did not depend on where children stood during the trial. Although these results point toward allocentric encoding, it is important to note that this cannot be the whole story. Rather, an account is needed that explains how egocentrically perceived spatial information (e.g., in retinal coordinates) is transformed into an allocentric representation of locations and then back into an egocentric response (see Spencer & Schutte, in press). We are currently extending the DFT to account for the egocentric-to-allocentric transformation. This more complete model will allow us to specify more precisely the ties between the egocentric responses of infants in the A-not-B situation and the more allocentric responses of toddlers.

Implications for Other Spatial Memory Models

Several models have been proposed to account for A-not-B errors in the canonical A-not-B task; however, these models have not been applied to tasks with homogeneous spaces. Despite this, it is useful to consider whether other models can account for the effects reported here. With some models, it is not currently possible to test the effect of varying the separation between hiding locations. For instance, in its present form, Munakata's parallel distributed processing (PDP) model (Munakata, 1998; Munakata et al., 1997) does not represent space using continuous, metric dimensions. Rather, this model has three location input nodes. Consequently, it is not possible to vary the distance between hiding locations using the current architecture. In addition, this model only has three output nodes; therefore, it cannot produce the response errors between hiding locations reported here.

Other models of the A-not-B error are difficult to apply to the sandbox task because they have not been specified in enough detail. Diamond et al.'s (1994) memory+inhibition account does not provide a formalized model. Consequently, it is difficult to determine a priori what this account predicts when applied to different tasks with different ages. In contrast to the account by Diamond et al., Marcovitch and Zelazo's (1999) response versus representation account has been mathematically formalized. Moreover, this model has been used to explain how target separation affects infants' performance in the A-not-B task. As such, the Marcovitch and Zelazo account is particularly relevant to the present study. This model predicts that A-not-B-type errors should dissipate in early childhood as children shift from relying on a response-based memory for locations to a representation-based memory for locations. Our data are not consistent with this proposal (see also Spencer & Schutte, in press).

The final model considered here is the category adjustment (CA) model proposed by Huttenlocher et al. (1991). According to the CA model, children encode two types of location information in spatial recall tasks—categorical information and fine-grained information. For instance, Huttenlocher et al. (1994) proposed that young children treat large, homogeneous spaces such as the sandbox as one category with a spatial prototype at the center. Children encode both the fine-grained, or metric, location of the target and the location of the prototype. During recall, these two types of information are combined, which results in a bias toward the prototype. This provides an account for why 2- and 4-year-olds were biased toward midline on the A trials. Huttenlocher et al. (1994) also proposed that older children and adults divide the sandbox into two categories with a prototype at the center of each half. This produces biases away from midline, consistent with 11-year-olds' responses on the A trials.

Although the CA model provides an account for midline-related biases, this model does not provide an explicit account of experience-dependent biases in the spatial

domain. Nevertheless, Huttenlocher, Hedges, and Vevea (2000) recently used the central ideas of the CA model to capture induced category effects in object categorization. Specifically, Huttenlocher et al. proposed that adults use two types of information when recalling object properties— fine-grained information about the details of the object and a representation of the category in which the object belongs constructed via task-specific experience. Although these ideas have not been applied within the spatial domain, they suggest that a modified version of the CA model might be able to account for A-not-B-type effects.

Conclusion

The current study tested a novel prediction of the DFT—that A-not-B-type errors in the sandbox task depend on both target separation and age. Data from three experiments were consistent with this prediction. This successful test of the model combined with quantitative fits demonstrates the generativity of the DFT. Moreover, our results provide support for the proposal that interactions in SWM become more spatially precise over development (Spencer & Hund, in press). A central goal of future research will be to determine the factors that influence this developmental change.

References

Amari, S. (1977). Dynamics of pattern formation in lateral-inhibition type neural fields. *Biological Cybernetics*, 27, 77 – 87.

Bastian, A., Riehle, A., Erlhagen, W., & Schöner, G. (1998). Prior information preshapes the population representation of movement direction in motor cortex. *NeuroReport*, 9, 315 – 319.

Compte, A., Brunel, N., Goldman-Rakic, P. S., & Wang, X. J. (2000). Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cerebral Cortex, 10*, 910 – 923.

Diamond, A. (1990). The development and neural bases of memory functions as indexed by the A-not-B and delayed response tasks in human infants and infant monkeys. *Annals of the New York Academy of Sciences, 608*, 267 – 317.

Diamond, A., Cruttenden, L., & Neiderman, D. (1994). AB with multiple wells: 1. Why are multiple wells sometimes easier than two wells? 2. Memory or memory+inhibition? *Developmental Psychology, 30*, 192 – 205.

Diamond, A., & Doar, B. (1989). The performance of human infants on a measure of frontal cortex function, the delayed response task. *Developmental Psychobiology, 22*, 271 – 294.

di Pellegrino, G., & Wise, S. P. (1993). Visuospatial versus visuomotor activity in the premotor and prefrontal cortex of a primate. *Journal of Neuroscience, 13*, 1227 – 1243.

Erlhagen, W., Bastian, A., Jancke, D., Riehle, A., & Schöner, G. (1999). The distribution of neuronal population activation (DPA) as a tool to study interaction and integration in cortical representations. *Journal of Neuroscience Methods, 94*, 53 – 66.

Erlhagen, W., & Schöner, G. (2002). Dynamic field theory of movement preparation. *Psychological Review, 109*, 545 – 572.

Georgopoulos, A. P., Kettner, R. E., & Schwartz, A. B. (1988). Primate motor cortex and free arm movements to visual targets in three-dimensional space. II. Coding of the direction of movement by a neuronal population. *Journal of Neuroscience, 8*, 2928 – 2937.

Georgopoulos, A. P., Taira, M., & Lukashin, A. (1993). Cognitive neurophysiology of the motor cortex. *Science, 260*, 47 – 52.

Hund, A. M., & Spencer, J. P. (2003). Developmental changes in the relative weighting of geometric and experience-dependent location cues. *Journal of Cognition and Development, 4*, 3 – 38.

Huttenlocher, J., Hedges, L. V., & Duncan, S. (1991). Categories and particulars: Prototype effects in estimating spatial location. *Psychological Review, 98,* 352 – 376.

Huttenlocher, J., Hedges, L. V., & Vevea, J. L. (2000). Why do categories affect stimulus judgement? *Journal of Experimental Psychology: General, 129,* 220 – 241.

Huttenlocher, J., Newcombe, N., & Sandberg, E. H. (1994). The coding of spatial location in young children. *Cognitive Psychology, 27,* 115 – 147.

Huttenlocher, P. R. (1979). Synaptic density in human frontal cortex—Developmental changes and effects of aging. *Brain Research, 163,* 195 – 205.

Huttenlocher, P. R. (1990). Morphometric study of human cerebral cortex development. *Neuropsychologia, 28,* 517 – 527.

Marcovitch, S., & Zelazo, P. D. (1999). The A-not-B error: Results from a logistic meta-analysis. *Child Development, 70,* 1297 – 1313.

McDowell, K., Jeka, J. J., Schöner, G., & Hatfield, B. D. (2002). Behavioral and electrocortical evidence of an interaction between probability and task metrics in movement preparation. *Experimental Brain Research, 144,* 303 – 313.

Munakata, Y. (1998). Infant perseveration and implications for object permanence theories: A PDP model of the AB task. *Developmental Science, 1,* 161 – 184.

Munakata, Y., & McClelland, J. L. (2003). Connectionist models of development. *Developmental Science, 6,* 413 – 429.

Munakata, Y., McClelland, J. L., Johnson, M. H., & Siegler, R. S. (1997). Rethinking infant knowledge: Toward an adaptive process account of successes and failures in object permanence tasks. *Psychological Review, 104,* 686 – 719.

Newcombe, N., Huttenlocher, J., Drummey, A. B., & Wiley, J. G. (1998). The development of spatial location coding: Place learning and dead reckoning in the second and third

years. *Cognitive Development, 13,* 185 – 200.

Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive Science, 14,* 11 – 28.

Piaget, J. (1954). *The Construction of Reality in the Child.* New York: Basic.

Schöner, G., Kopecz, K., & Erlhagen, W. (1997). The dynamic neural field theory of motor programming: Arm and eye movements. In P. G. Morasso & V. Sanguineti (Eds.), *Self-Organization, Computational Maps and Motor Control* (Vol. 119, pp. 271 – 310). Amsterdam: Elsevier-North Holland.

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.

Smith, L. B., Thelen, E., Titzer, R., & McLin, D. (1999). Knowing in the context of acting: the task dynamics of the A-not-B error. *Psychological Review, 106*, 235 – 260.

Spencer, J. P., & Hund, A. M. (2002). Prototypes and particulars: Geometric and experience-dependent spatial categories. *Journal of Experimental Psychology: General, 131,* 16 – 37.

Spencer, J. P., & Hund, A. M. (in press) Developmental continuity in the processes that underlie spatial recall. *Cognitive Psychology.*

Spencer, J. P., & Schöner, G. (2000, August). *A dynamic field model of location memory.* Paper presented at the 22nd Annual Conference of the Cognitive Science Society, Philadelphia.

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

Spencer, J. P., & Schöner, G. (2003b). Bridging the representational gap in the dynamic systems approach to development. *Developmental Science.*

Spencer, J. P., & Schutte, A. R. (in press). Do perseverative biases in early development solely reflect prepotent responses? *Psychological Science.*

Spencer, J. P., Smith, L. B., & Thelen, E. (2001). Tests of a dynamic systems account of the A-not-B error: The influence of prior experience on the spatial memory abilities of 2-year-olds. *Child Development, 72,* 1327 – 1346.

Suzuki, W. A., Miller, E. K., & Desimone, R. (1997). Object and place memory in the macaque entorhinal cortex. *Journal of Neurophysiology, 78,* 1062 – 1081.

Tanaka, S. (2000). Roles of intracortical inhibition in the formation of spatially tuned delay-period activity of prefrontal cortical neurons: Computational study. *Progress in Neuropsychopharmacology & Biological Psychiatry, 24,* 483 – 504.

Thelen, E., Schöner, G., Scheier, C., & Smith, L. B. (2001). A dynamic field theory of infant perseverative reaching errors. *Behavioral and Brain Sciences, 24,* 1 – 86.

Tipper, S. P. (1992). Selection for action: The role of inhibitory mechanisms. *Current Directions in Psychological Science, 1,* 105 – 109.

Wellman, H. M., Cross, D., & Bartsch, K. (1987). Infant search and object permanence: A meta-analysis of the A-not-B error. *Monographs of the Society for Research in Child Development, 51,* (3, Serial No. 214).

Wilson, F. A. W., Scalaidhe, S. P., & Goldman-Rakic, P. S. (1993). Dissociation of object and spatial processing domains in primate prefrontal cortex. *Science, 260,* 1955 – 1958.

Zelazo, P. D., Reznick, J. S., & Spinnazzola, J. (1998). Representational flexibility and response control in a multistep multilocation search task. *Developmental Psychology, 34,* 203 – 214.

Appendix

In the simulations reported in the text, activation in the working memory field was governed by the following equation:

$$\tau \dot{u}(x,t) = -u(x,t) + h + \int dx' w(x-x') f(u(x'))$$
$$+ u_{tar}(x,t) + u_{mem}(x,t) + q \int dx' g_{noise}(x$$
$$- x') \xi(x',t)$$

where $u(x, t)$ is the current activation in the field, h is the resting level, $\int dx' w(x–x') f(u(x'))$ is the interaction function, $u_{tar}(x, t)$ is the target input, $u_{mem}(x, t)$ is the longer term memory input, and $q \int dx' g_{noise}(x - x') \xi(x', t)$ is spatially correlated noise. The τ parameter on the left side of the equation sets the time scale of the dynamical system. The target input took the form of a Gaussian represented by the following equation:

$$u_{tar}(x,t) = S_{tar} \exp\left[- \frac{(x - x_{tar})^2}{2\sigma_{tar}^2} \right] \chi^{(t)[2100,4100]},$$

where S_{tar} is the strength of the input and star determines the spatial spread (i.e., spatial precision) of the input. The target input was centered at the target location given by x_{tar}. Because the target was only visible for a few seconds in our studies, this input was turned on at 2,100 ms and turned off at 4,100 ms (the first 100 ms of the simulation allowed the dynamic field to relax to a stable resting state). This step function is specified by $\chi^{(t)}$.

In the DFT model reported by Thelen et al. (2001), the longer term memory input was constructed dynamically over trials. In the present set of simulations, however, we only simulated performance on the first B trial. There-

fore, the memory input was represented by a Gaussian as well:

$$u_{mem}(x, t) = S_{mem} \exp\left[-\frac{(x - x_{mem})^2}{2\sigma_{mem}^2}\right],$$

where S_{mem} is the strength of the input and σ_{mem} is the spatial spread. The memory input was centered at the A location given by x_{mem}. The characteristics of interaction in the field were specified by the convolution of two functions. A threshold function determined which sites participated in the locally excitatory–laterally inhibitory interaction: Only sites that were sufficiently active had an effect on other sites. The level of activation required to enter into the interaction was determined by a sigmoid or S-shaped function of the form:

$$f(u) = \frac{1}{1 + \exp[-\beta u]},$$

where β is the slope of the sigmoid. In addition to the threshold function, an interaction function, $w(x–x')$, spec-

ified the spatial structure of the interaction (see main text for additional details). Noise was added to the simulations by convolving a noise field composed of independent noise sources with a Gaussian kernel specified by:

$$g_{noise}(x - x') = \frac{1}{\sqrt{2\sigma_{noise}}} \exp\left[-\frac{(x - x')^2}{2\sigma_{noise}^2}\right],$$

where σ_{noise} is the spatial spread of the noise kernel. Although this type of noise differs from the Gaussian white noise used by Thelen et al. (2001), more recent work in our laboratory has shown that spatially correlated noise can capture variations in peak position in the field model more effectively than white noise (Spencer & Schöner, 2003a). In contrast to white noise, spatially correlated noise introduces slightly wavy perturbations to the field—random fluctuations can take on a spatially distributed characteristic because of the noise kernel. A large enough wave can shift the spatial position of a peak in working memory by giving an extra boost of activation to sites to the left or right of the maximally activated site.