

Automatic and Voluntary Shifts of Attention in a Dynamic Neural Field Model of the Dimensional Change Card Sort Task

Aaron Buss (aaron-buss@uiowa.edu) and John P. Spencer (john-spencer@uiowa.edu)

Department of Psychology, University of Iowa, 11 Seashore Hall East

Iowa City, IA 52242 USA

Abstract

Recently, Buss & Spencer (2008) proposed a Dynamic Neural Field (DNF) model of the Dimensional Change Card Sort task (DCCS). This model is able to not only capture the details of 3- and 4-year-olds' performance in the standard version, but also generalizes to account for performance in two other canonical variations. The distributions of features in space plays a central role in capturing these effects. To show that the model can generalize beyond space-based effects, we present simulations of DCCS variations reported by Fisher (2008) that examine the role of automatic and voluntary shifts of attention and randomize the spatial location of the target cards. Results show that the DNF model captures performance in these conditions as well.

Keywords: neural network models; dynamic systems theory; DCCS; attention; executive function; perseveration

The DNF Model of the DCCS

The Dimensional Change Card Sort task (DCCS) and other rule-based card sorting tasks provide an index of executive functioning both in early development and disordered populations (Zelazo et al., 2003). In the DCCS task, children sort cards by one dimension, for instance, the dimension of shape and are then instructed to switch and sort by another dimension such as color. Target cards are displayed on the trays where cards are sorted as clues for what features go where. Importantly, the test cards that children sort typically match each target card along one dimension (see Figure 1). Three-year-olds have little trouble using shape or color rules during the initial sorting phase; however, once these children are instructed to switch and sort by the other dimension, they typically perseverate and continue to use the first set of rules they were given. This task, then, measures aspects of cognitive flexibility, and children's behavior has been described in terms of rule-representation (Zelazo et al., 2003), object re-description (Kloo & Perner, 2005), inhibitory or dis-inhibitory control (Zelazo et al., 2003), and attentional control (Kirkham, Cruess, & Diamond, 2003).

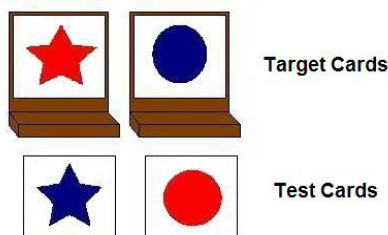


Figure 1: Example of cards used in the DCCS

More recently there has been a growing interest in using such tasks as an index of neural functioning. Specifically, these tasks have been linked to developmental changes in the prefrontal cortex as well as other brain regions (Bunge & Zelazo, 2006; Crone et al., 2006) and have been used to explore cognitive deficits in children with autism (Zelazo et al., 2002) and ADHD (Mulas et al., 2006). Although this move toward neural functioning adds new critical information about the nature of performance in this task, current theories are not well grounded in neural dynamics and neurally plausible concepts. A Dynamic Neural Field (DNF) model proposed by Buss & Spencer (2008) provides a critical bridge between the brain and behavior in this task. Using a neurally-grounded view of the processes that underlie the representations of shapes, colors, and space, the model provides an account of children's behavior and development. This is not explicitly a rule-use model; rather, it is couched in a general framework originally proposed to capture the integration of 'what' (i.e., ventral stream) and 'where' (i.e., dorsal stream) visual pathways in object representation. The DNF model is able to not only capture the details of 3- and 4-year-olds' performance in the standard task, but also generalizes to capture performance in other canonical versions of the DCCS.

The effects modeled by Buss & Spencer (2008) revealed an important role for spatial information in rule-use in the DCCS. Specifically, difficulty switching in the model stems from an inability to resolve spatial conflict between LTM traces accumulated from past decision and the current task inputs present when sorting a card. To provide a more comprehensive account of performance in this task and to show that the model captures phenomena that are not grounded in feature-space conflict, we present simulations of new versions of the DCCS reported by Fisher (2008). These versions distinguish automatic from voluntary shifts of attention by manipulating feature saliency and attentional weights for sets of features. To isolate these types of attentional switching, the spatial locations of the target cards were randomly assigned on each trial. With space thus randomized, children's errors cannot be a result of specific spatial habits or memories, eliminating space as a source of children's perseverative errors. These, then, seem like ideal conditions to provide a thorough test of the model. Before describing the implementation of these versions in the DNF model, we first start with a brief summary of the dynamics of rule-use.

Model Architecture

DNF models use fields of neurons that are tuned to continuous feature dimensions (e.g., shape, space, or color).

The fundamental unit of cognition within this framework is a peak of activation for a particular feature or set of features within these fields. Thus, peaks of activation capture representations that underlie decisions or working-memory for different spatial locations or object features. We utilize a generic two-layered architecture initially explored by Amari (1977) to capture the dynamics of neural activity in visual cortex. To highlight the dynamics that lead to peak formation, Figure 2 shows a 1-dimensional spatial system (the same architecture and dynamics are present for the 2-dimensional fields to be described below, except the interactions are along both dimensions). For these 1-dimensional fields, neurons tuned to a particular continuous feature dimension are lined up along the x-axis (in Figure 2, this is a spatial dimension), while each neuron's activation is displayed along the y-axis.

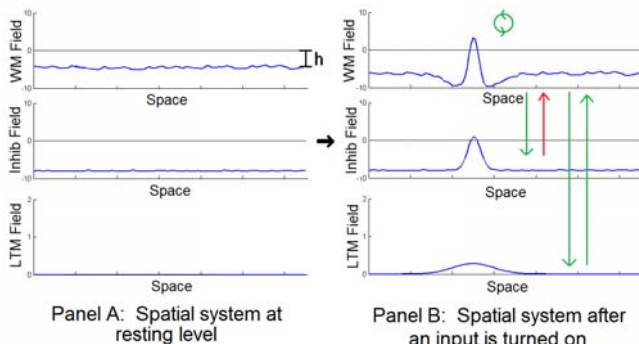


Figure 2: Excitatory and inhibitory architecture in a 1-dimensional system, along with an associated LTM layer.

The architecture in Figure 2 consists of a layer of self-excitatory working memory neurons (WM) arranged by functional topography (e.g., neurons that code for nearby spatial locations are neighbors in the network). These neurons project activation (see green arrows in panel B) to a layer of inhibitory interneurons (Inhib) which project inhibition broadly back into WM (see red arrow). Activation within these fields is passed through a sigmoid function so that only neurons that are sufficiently close to or above a 0 activation threshold contribute to interaction. This gives the WM field a critical type of non-linearity: a local population of neurons can go from a stable resting state to a stable active or ‘on’ state, creating a ‘peak’ of activation (see blue Gaussian profile in panel B). These peaks underlie behavioral choices in the model to, for instance, sort a card to a left location (as in Figure 2). An important parameter to mention here is ‘h’, or the resting-level of the field ($h < 0$). The closer a WM field is to 0 the quicker inputs will build to threshold, engage neural interactions, and build a peak.

In addition to the WM and Inhib layers, we add a long-term memory (LTM) layer (bottom layer in Figure 2) that is reciprocally coupled to the WM layer. The WM and LTM layers implement a form of Hebbian learning: peaks in WM leave traces of activity at associated sites in LTM which, in turn, influence the strength of activation in WM. As can be seen in panel B of Figure 2, the peak in WM is building a

LTM trace in the bottom LTM layer which is projecting activation back into WM. To summarize, then, the WM layer is influenced by inputs from the world, excitatory/inhibitory interactions, and LTM traces that accumulate over time in a given task.

Object Representation

The DNF model of the DCCS is an extension of an object representation model put forth by Johnson, Spencer, & Schöner (2008). The basic architecture and the mapping to cortical pathways are displayed in Figure 3. For simplicity, only the WM fields are shown; however, each WM field is coupled to its own inhibitory and LTM layer. Research has revealed a dorsal/ventral dissociation in visual processing (Ungerleider & Mishkin, 1982). Dorsal stream processes are primarily devoted to encoding spatial information and are captured in a 1-dimensional spatial working memory (SWM) field (see top layer in Figure 3A and 3B). Ventral stream processes are primarily devoted to encoding object properties with different feature dimensions encoded by different populations of neurons. Importantly, many of these ventral stream areas are precise in their representation of features, but also have broad spatial receptive fields (Desimone & Gross, 1979). These processes are captured by 2-dimensional feature-space working memory (FWM) fields which contain a continuous feature dimension (e.g., color or shape, depicted along the y-axis) and a spatial dimension (depicted along the x-axis; see middle and bottom fields in Figures 3A and 3B). Relative levels of activation within these 2-dimensional fields are depicted by the color on the inset scale (red is high activation, blue is low activation).

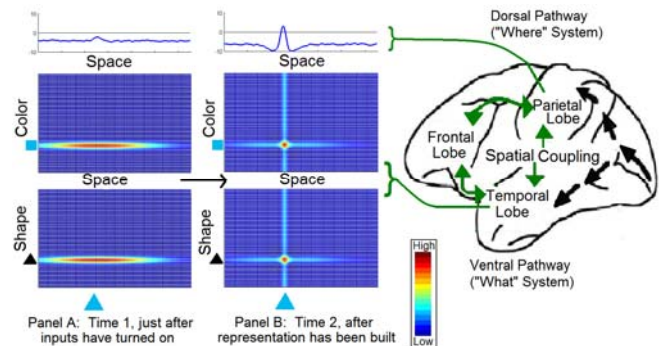


Figure 3: Space, shape, and color WM fields representing a blue triangle and the mapping to different visual pathways.

Given the broad spatial tuning of ventral stream neurons, how does our model solve the “binding” problem of vision (Treisman & Gelade, 1980)? That is, how would the model know that a “blue” peak in the color-space field should be matched with a “triangle” peak in a shape-space field in representing a blue triangle? The answer is through spatial coupling; that is, the SWM and FWM fields contain a common spatial dimension and thus share spatial activation. This achieves distributed but coupled peaks of activation across multiple cortical fields—our implementation of an “integrated” object (for related ideas, see Treisman &

Gelade, 1980).

Figure 3 shows the WM fields of the model representing a blue triangle at a location left of center in the task-space. The left panel shows the inputs just after they have been turned on. There is a small bump of activation building at the spatial location of the object in the SWM field. Meanwhile, the 2-dimensional inputs into the feature WM fields are broad in spatial information while being relatively precise in feature information (see Figure 3A). In Figure 3B, the inputs have built to threshold and excitatory/inhibitory interactions for each WM field have formed a peak of activation. The shared spatial activation is now visible in the feature WM fields as ridges of sub-threshold activation running vertically at the location of the object (there is also spatial coupling going into the spatial WM field; however, it is not readily visible). Thus, coupled spatial activation serves to anchor features together.

The final aspect of the model architecture highlighted in Figure 3 is the contribution from frontal lobe areas. As we discuss below, we hypothesize that neural processes in the frontal lobe can selectively modulate the resting level of populations of neurons in SWM and FWM (e.g., boosting the excitability of color neurons). This serves as a mechanism to bias representations of objects based on different visual dimensions and is a neurally plausible mechanism underlying ‘rule’-like behavior (Egner & Hirsch, 2005). This modulation of activity is putatively grounded in word-learning. As the labels ‘shape’ and ‘color’ are learned, stronger levels of resting-level boost can be applied to the populations of neurons coding for the respective type of information. Thus, developmental changes in rule-flexibility in the DCCS is hypothesized to be a product of a simple associative mechanism between the labels ‘shape’ and ‘color’ and the feature fields encoding the respective type of information.

Rule-use and the Role of Space

Because space is a central component of the object representation model we are using to capture behavior in the DCCS, space is also an important component of rule-use. We highlight this in Figure 4 which shows how the model captures the sequence of events in the standard version of the DCCS task. Figure 4A shows the *target inputs* into the spatial and feature WM fields that capture the information on the target cards. Note that the hot spots in Figure 4A are not peaks, but are sub-threshold boosts of activity. This is apparent in the spatial WM field (top panel) which shows two sub-threshold bumps of activity at the location of the trays. Within the feature WM fields, these are hotspots of activity for particular features at the left and right locations. Here, the target cards are a red star on the left and a blue circle on the right. *Test inputs* are presented without spatial information as a ridge for a particular feature value across all spatial locations (see horizontal ridges in Figures 4B and 4D). The model, like children, must localize the features on the test cards to a spatial location. For the test inputs that could be presented (a blue-star or a red-circle), there is

spatial conflict between the relevant target inputs. For example, the blue feature input will overlap with a target input at the left location while the star feature input will overlap with a target input at the right location.

In order to resolve this conflict and get rule-like representations out of the model, we boost the resting-level for the relevant population of neurons. When ‘told’ to play the color game, for example, the resting level of the color field is given a slight boost. Figure 4 shows the sequence of events as the model sorts each test card according to the color dimension. Figure 4B shows the model sorting a blue-star. With the color field boosted, spatial activation at the location of the blue target input is shared with the spatial WM and shape WM field. This ultimately tips the balance of activation within the shape WM field and drives the binding of the star feature to the right sorting bin by building a peak at the rightward location. The model, then, binds the features together to represent the object at the location where it is to be placed. Figure 4C shows the accumulation of LTM as a result of making this decision. LTM is overlapping with the target input in the color WM field. However, LTM is at the opposite spatial location of the target input for that feature in the shape WM field (see white circle in Figure 4C); that is, the model is seeing a star at the left location in the task-space but has a memory for placing a star to the right.

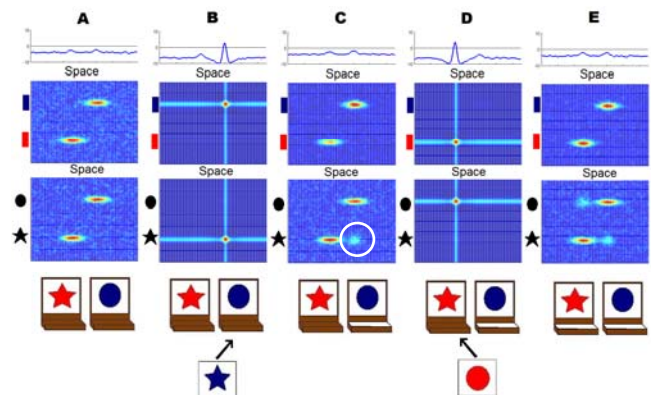


Figure 4: The sequence of events as the model sorts the two target cards.

The same sequence of events plays out as the model sorts a red-circle by color in Figure 4D. Figure 4E shows the accumulation of LTM after the pre-switch phase. Since the model consistently sorted by color, LTM is overlapping with the target inputs in the color field, while LTM is at conflicting locations with the inputs in the shape field. Thus, going into the post-switch phase, there is spatial conflict in the shape field.

What happens when we switch the ‘rule’ in the post-switch trials? To implement the rule switch in the model, we lower the resting level of the color field and boost the resting level of the shape field. With weak levels of resting-level boost, the model perseverates and uses the incorrect feature dimension to sort the test cards, that is, the model in

Figure 4 sorts by color. This occurs because activation in the color field builds quickly due to the strong overlap between the target inputs and the LTM traces accumulated during the pre-switch phase. By contrast, activation builds slowly in the shape field due to spatial conflict. If, however, we increase the resting-level boost for neurons within the post-switch field, the model sorts correctly and overcomes the spatial conflict created in the pre-switch phase. Thus, models with weak levels of resting level boost (i.e., weak representations of color and shape dimensions) will be successful in the pre-switch, but only models with strong levels of resting level boost will be able to switch rules. A model's ability to switch rules, then, is a function of the inputs to the model, the LTM that accumulates in the pre-switch phase, and the strength of resting-level boost that is supplied to the relevant feature WM field.

This model not only captures the performance of 3- and 4-year-olds in the standard task, but also generalizes to two other canonical versions (a Negative Priming and an Absolute Negative Priming version; see Buss & Spencer, 2008). The spatial conflict shown in Figure 4 plays a central role in capturing these effects. To show that this account is not only about space, it is necessary to demonstrate that the model can capture effects that eliminate a role for spatial information. We now turn to new variations of the DCCS which provide such a test.

Automatic and Voluntary Shifts of Attention

The nature of attentional shifting in the DCCS was explored by Fisher (2008) because of the implications the results would have for early word-learning. Theories of early word learning either posit that developmental changes occur through *automatic attention* driven by bottom-up processing (Sloutsky & Fisher, 2005) or *voluntary attention* that is applied in a conscious and deliberate manner (Gelman & Medin, 1993). Proponents of the former perspective argue that developmental changes in word-learning occur prior to the development of voluntary control of attention. If voluntary attention is not yet developed in 3-year-olds, then these children should only have trouble switching rules in the DCCS when a voluntary shift of attention is required.

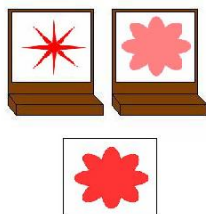


Figure 5: Example of cards used in Fisher (2008) with low saliency along the color dimension but high saliency along the shape dimension.

In one comparison, Fisher (2008) examined the role of feature saliency. Salient features have been shown to automatically capture attention (Treisman & Gelade, 1980); thus, when combined with a less salient set of features,

high-saliency should facilitate automatic switching of attention while low-saliency should require a voluntary switch. Fisher (2008) constructed the color features in the DCCS with low-saliency and the shape features with high-saliency. Using a more complex task design in order to isolate automatic and voluntary attentional processes, any influence of spatial information was eliminated by randomizing the location (left/right) of the target cards trial by trial. Thus, children's difficulty could not stem from spatial habits or memories associated with the images on the test cards. Further, half of the 8 trials during the pre- and post-switch phases were no-conflict (i.e., matched a target card along *both* dimensions).

Under these conditions, 3-year-olds had significantly more difficulty switching to a less-salient set of features than to a more-salient set of features (i.e., more difficulty switching to color than switching to shape). Thus, high-saliency facilitates an automatic switch of attention, while low-saliency is more attentionally demanding and requires a voluntary switch.

In another set of conditions, Fisher (2008) compared a high-attentional weights condition to a low-attentional weights condition (see Figure 6). Features are considered to have high attentional weights when there are fewer overall features (thus, each feature has a greater allotment of attention; Nosofsky, 1986). Again, randomizing the spatial location of the target cards trial-by-trial and using half no-conflict cards during each phase, 3-year-olds had significantly more difficulty switching rules when 2 features per dimension were used compared to 4 features per dimension. With lower attentional weights, attention for each feature essentially weighs less and is moved more easily to new features during the post-switch phase. Lower weights thus facilitate an automatic switch of attention, while higher weights are more attentionally demanding and require a voluntary switch.

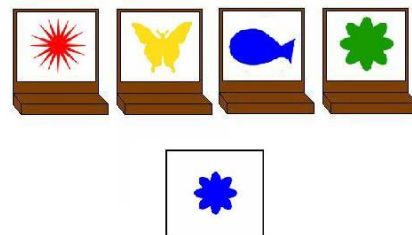


Figure 6: Example of cards used in a low-attentional weights version in Fisher (2008). Instead of 2, now there are 4 features within the color and shape dimensions.

Across these two variations in the DCCS, voluntary shifts of attention prove to be more difficult for 3-year-olds than automatic shifts of attention. This account leaves unanswered, however, what exactly makes a switch of attention automatic or voluntary and what makes voluntary shifts of attention difficult. As we will show, the DNF model provides a way to ground these concepts.

Simulation of Feature Saliency Conditions

To implement different degrees of feature saliency in the model, the relative separation between the inputs along the feature dimension was manipulated. As can be seen in Figure 7, the inputs to the color field are now closer along the feature dimension. Instead of a 31 neuron separation as in the shape field, the color inputs are now only 11 neurons apart. Inputs to the model were implemented exactly as in Fisher (2008). The spatial location of the target cards was randomized for each trial. Eight cards were presented during each the pre- and post-switch phase, half of which were no-conflict.

Figure 7 shows the post-switch results from 100 runs of the model with a weak resting-level boost for each post-switch dimension compared to the results from 3-year-olds in Fisher (2008). The model, like children, performed well when sorting by both shape and color during the pre-switch phase, sorting 98% and 96% correct, respectively. As can be seen, there is a tight fit between the model and the performance of 3-year-olds. Specifically, the model exhibited high rates of correct sorting for no-conflict cards for both shape and color rules. Further, the model was able to successfully switch to the shape rules but had a very poor level of performance switching to the color rules.

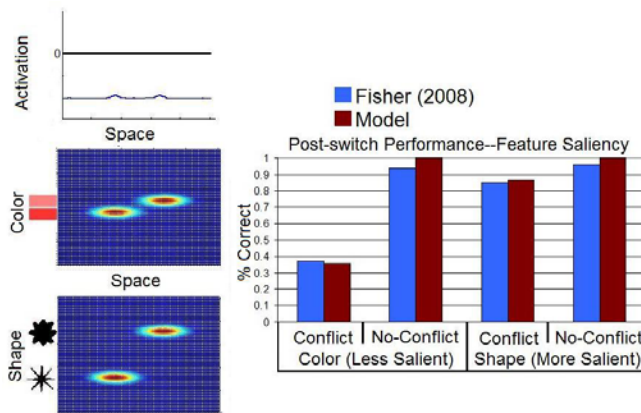


Figure 7: Inputs to WM fields for the feature saliency condition (left panel) and fit of performance between the model and 3-year-olds.

To explore the source of difficulty with similar feature metrics, we examined the latency in peak-build time on correct trials within the shape and color fields when they were boosted and relevant for the pre-switch. Since the model performed with a high level of success during the pre-switch, the boosted field is reliably driving the decision making process. Thus, this gives a measure of how long it takes to build a representation of a relevant feature in order to make a decision based on either color or shape. Average peak latency within the color field with less-salient inputs was 703 time-steps, while this average latency was 677 time-steps in the shape field with more-salient inputs. Thus, activation builds more slowly with close feature metrics in the color field. This is due to the partially overlapping representation of the color features. The color field is more

active overall when presented with a test card which requires more inhibition during peak formation. Although this delay in peak build-time does not impact performance during the pre-switch trials, once LTM has been established and the model is switching to color rules, there is activation at each spatial location for each feature. This induces additional conflict, requiring even more inhibition to resolve the conflict and select a choice along the color dimension. Thus, it is difficult to sort by color and shape wins the competition during peak formation.

The difference between automatic and voluntary shifting in this situation is the amount of resting-level boost required to sort by a given dimension's inputs. Switching to more salient features is 'automatic' in that distinct features win-out by default given the overall slowing down of activation in the color field. Thus, switching to distinct features is supported in a bottom-up manner by the inputs to the model. Conversely, switching to less-salient features is voluntary in that it requires extra resting-level boost (i.e., a stronger representation of shape and color dimensions within frontal areas) to correctly switch to these features. This condition, then, highlights how these two types of attention can be conceptualized within the dynamic field framework and provides neural grounding for these concepts.

Simulation of Attentional Weights Conditions

In the second condition from Fisher (2008), the high-attentional weights inputs have 2 different feature values for each dimension and 2 sorting locations, while the low-attentional weights inputs have 4 different feature values and 4 sorting locations (see Figure 8). Once again, we simulated this condition with no special modifications to the model; rather, we simply implemented Fisher's procedure. The location of the target cards was randomized so that each target card occupied each location an equal number of times. Eight trials were administered for each of the pre- and post-switch phases with half of the cards being no-conflict. The comparison of 100 runs of the model with each set of features and the 3-year-olds from Fisher (2008) is displayed in Figure 9 (the model showed a high level of performance during pre-switch for both the low- and high-attentional weight versions, sorting 100% and X% correct, respectively). Again, there is a tight fit between the model and children's performance. Specifically, both the model and children displayed a high rate of correct sorting with the no-conflict cards. Further, the model had a high rate of correct switching in the low-attentional weights condition but a low rate of switching in the high-attentional weights condition.

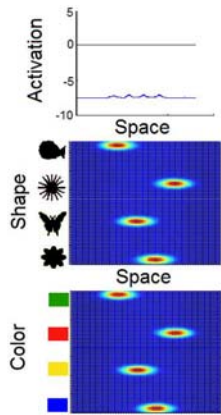


Figure 8: Inputs to WM fields for the low-attentional weights condition (left panel) and fit of performance between the model and 3-year-olds.

What underlies this ‘attention’ related performance? In the model, LTM for visual features is specific to spatial locations within the 2-dimensional feature WM fields. With randomized space and 4 cards, the LTM for the features involved in the task becomes distributed across the 4 spatial locations. Further, with 4 features per dimension, the individual features are presented less often. With 2 features, by contrast, LTM accumulates at more consistent feature and spatial locations. As a result, LTM with 2 features becomes stronger and leads to more interference during the post-switch phase. This, in turn, requires a stronger resting-level boost to sort correctly. With 4 features, however, LTM is weaker and, thus, does not create strong interference in the post-switch field. Indeed, the average strength of LTM with 2 features was X units, while this average was only X units with 4 features. The distinction between automatic and voluntary is, again, the increase in the strength of resting-level boost required to sort in the post-switch relative to the pre-switch. Distributing LTM across 4 features builds up weak LTM which allows the inputs to the feature fields to drive correct switching in an automatic and bottom-up fashion. However, more consistent LTM accumulation for only 2 features builds up strong LTM and requires a stronger resting-level boost, that is, a voluntary shift of attention.

Discussion

Our object representation system coupled with a simple mechanism of resting-level modulation has proven to be a versatile framework for capturing children’s rule-use and the development of flexible rule-use. This shows the power and utility of formalizing and grounding abstract concepts such as rule-use and rule-representation in a neural framework. Within the same model, concepts such as negative-priming, inhibition, and various attentional processes can all be united and understood.

The previous simulations presented by Buss & Spencer (2008) highlighted the role of spatial information and showed that boosting the resting level of relevant feature fields is enough to capture developmental changes in

flexible rule-use. Here we showed that more subtle effects can emerge from the interactions of LTM, inputs, and resting-level modulation such that correct switching can be supported in a bottom-up *and* top-down fashion based on the nature of activation within the feature fields.

Further, these effects provide a formal tie between word-learning and the hypothesized mechanism of resting-level modulation in the model. Fisher (2008) constructed the versions simulated here to reconcile the literature on early word-learning and rule-flexibility. Under debate was whether attentional processes in word learning were automatic or under voluntary control. Fisher (2008) found evidence that voluntary attention is not developed at 3-years of age, and children were only able to switch rules when automatic shifts of attentions were supported by the task. In the model, different inputs (i.e., salient features and low-attentional weights) support automatic switching even with weak levels of resting-level modulation, that is, weak representations of ‘shape’ and ‘color’ categories; however, stronger levels of resting-level boost is required in order to switch when voluntary attention is required (i.e., non-salient features and high-attentional weights). In the model, then, automatic attention is an early development, while voluntary attention only emerges later in development with stronger category representations. Further work will probe how thoroughly this rule-use mechanism is grounded in early word-learning and how the developmental changes needed to support rule-use might emerge from word or category learning.

Acknowledgements

Research funded by NSF HSD0527698 awarded to JPS.

References

- Amari, S. (1977). Dynamics of pattern formation in lateral-inhibition type neural fields. *Biological Cybernetics*, 27, 77-87.
- Bunge, S. A., & Zelazo, P. D. (2006). A brain-based account of the development of rule use in childhood. *Current Directions in Psychological Science*, 15(3), 118-121.
- Buss, A. T. & Spencer, J. P. (2008). The role of space in children’s rule-use. In A. T. Buss (Chair), *Cognitive Development*. Paper symposium presented at the 30th Annual Meeting of the Cognitive Science Society, Washington, DC.
- Crone, E. A., Donohue, S. E., Honomichl, R., Wendelken, C., and Bunge, S. A. (2006). Brain regions mediating flexible rule use during development. *The Journal of Neuroscience*, 26(43), 11239-11247.
- Desimone, R & Gross, C. G. (1979). Visual areas in the temporal cortex of the macaque. *Brain Research*, 178(2-3), 363-380.
- Egner, T., & Hirsch, J. (2005). Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. *Nature Neuroscience*, 8(12), 1784-1790.

- Fisher, A. V. (2008). Automatic and voluntary shifts of attention of the Dimensional Change Card Sorting task. In A. T. Buss (Chair), *Cognitive Development*. Paper symposium presented at the 30th Annual Meeting of the Cognitive Science Society. Washington, DC.
- Gelman, S. A., & Medin, D. L. (1993). What's so essential about essentialism? A different perspective on the interaction of perception, language, and concrete knowledge. *Cognitive Development*, 8, p.157-167.
- Johnson, J.S., Spencer J.P., and Schöner, G. (in press). Moving to higher ground: the Dynamic Field Theory and the dynamics of visual cognition. In F. Garzòn, A. Laakso, T. Gomila (Eds.). *Dynamics and Psychology* [special issue]. *New Ideas in Psychology*.
- Kirkham, N. Z., Cruess, L., & Diamond, A. (2003). Helping children apply their knowledge to their behavior on a dimension-switching task. *Developmental Science*, 6(5), 449-476.
- Kloo, D., & Perner, J. (2005). Disentangling dimensions in the dimensional change card sort. *Developmental Science*, 8(1), 44-56.
- Mulas, F., Capilla, A., Fernández, S., Etchepareborda, M. C., Campo, P., Maestú, F., Fernández, A., Castellanos, F. X., & Ortiz, T. (2006). Shifting-related brain magnetic activity in attention-deficit/hyperactivity disorder. *Biological Psychiatry*, 59, 373-379.
- Nosofsky, R. M. (1984). Choice, similarity, and the context theory of classification. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 10, 104-114.
- Sloutsky, V. M. & Fisher, A. V. (2005). Similarity, Induction, Naming, and Categorization (SINC): Generalization or Inductive Reasoning? Response to Heit and Hayes. *Journal of Experimental Psychology: General*, 134, 606-611.
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12, 97-136.
- Ungerleider, L. G. & Mishkin, M (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale, and R. J. W. Mansfield (Eds.), *Analysis of Visual Behavior*. Cambridge, MA: MIT Press.
- Zelazo, P. D., Jacques, S., Burack, J. A., & Frye, D. (2002). The relation between theory of mind and rule use: evidence from persons with autism-spectrum disorders. *Infant and Child Development*, 11, 171-195.
- Zelazo, P. D., Müller, U., Frye, D., & Marcovitch, S. (2003). The development of executive function. *Monographs of the Society for Research in Child Development*, 68(3).