Cognition 192 (2019) 104003

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Original Articles

Dimensional attention as a mechanism of executive function: Integrating flexibility, selectivity, and stability

Aaron T. Buss[a,*], Anastasia Kerr-German[b]

[a] University of Tennessee – Knoxville, Department of Psychology, United States
[b] Boys Town National Research Hospital, Center for Neurobehavioral Research, United States

ARTICLE INFO

Keywords:
Dimensional attention
Executive function
Neural process model

ABSTRACT

In this report, we present a neural process model that explains visual dimensional attention and changes in visual dimensional attention over development. The model is composed of an object representation system that binds visual features such as shape and color to spatial locations and a label learning system that associates labels such as “color” or “shape” with visual features. We have previously demonstrated that this model explains the development of flexible dimensional attention in a task that requires children to switch between shape and color rules for sorting cards. In the model, the development of flexible dimensional attention is a product of strengthening associations between labels and features. In this report, we generalize this model to also explain development of stable and selective dimensional attention. Specifically, we use the model to explain a previously reported developmental association between flexible dimensional attention and stable dimensional attention. Moreover, we generate predictions regarding developmental associations between flexible and selective dimensional attention. Results from an experiment with 3- and 4-year-olds supported model predictions: children who demonstrated flexibility also demonstrated higher levels of selectivity. Thus, the model provides a framework that integrates various functions of dimensional attention, including implicit and explicit functions, over development. This model also provides new avenues of research aimed at uncovering how cognitive functions such as dimensional attention emerge from the interaction between neural dynamics and task structure, as well as understanding how learning dimensional labels creates changes in dimensional attention, brain activation, and neural connectivity.

1. Introduction

Executive function (EF) refers to the collection of mechanisms that give rise to goal-directed cognition and behavior. Measures of EF during early childhood are predictive of later quality of life outcomes (Moffitt et al., 2011) and academic achievement (Blair, Zelazo, & Greenberg, 2005; Bull, Espy, & Wiebe, 2008; Lee, Ng, Bull, Pe, & Ho, 2011; St Clair-Thompson & Gathercole, 2006; Swanson, Jerman, & Zheng, 2008), suggesting that EF is a foundational aspect of cognition. Although these studies suggest that strengthening EF could serve as a leverage point to improve developmental outcomes across multiple cognitive domains, efforts to improve EF through intervention have had only limited success (Diamond & Lee, 2011). Such efforts are complicated by a lack of consensus regarding the mechanisms that are involved in performance on specific tasks that are aimed to measure EF (Garon, Bryson, & Smith, 2008; Hanania & Smith, 2010; Happaney & Zelazo, 2003; Kirkham & Diamond, 2003; Munakata, Morton, & Yerys, 2003). Thus, it is not clear which mechanisms of EF should be targeted to expect gains in performance on EF tasks or in real-world behavior, suggesting that our current understanding of EF is incomplete.

Previous research has taken steps toward identifying mechanisms of EF using a latent variable approach (Miyake et al., 2000). This work has demonstrated that measures from batteries of EF tasks load onto factors that can be described as inhibition, memory updating, and set-shifting or switching. Developmentally, it has been demonstrated that EF skills differentiate over time. During early childhood a single factor typically explains performance on EF tasks (Fuhs & Day, 2011; Shing, Lindenberger, Diamond, Li, & Davidson, 2010; Wiebe et al., 2011; Wiebe, Espy, & Charak, 2008; Willoughby, Blair, Wirth, & Greenberg, 2010; Willoughby, Wirth, Blair, & Family Life Project Investigators, 2012). During later childhood two distinct factors are typically evident (Huizinga, Dolan, & van der Molen, 2006; Lee, Bull, & Ho, 2013; McAuley & White, 2011; Shing et al., 2010; van der Sluis, de Jong, & van der Leij, 2007; Van der Ven, Kroesbergen, Boom, & Leseman,

* Corresponding author.
E-mail address: abuss@utk.edu (A.T. Buss).

https://doi.org/10.1016/j.cognition.2019.06.015
Received 9 August 2018; Received in revised form 13 June 2019; Accepted 14 June 2019
0010-0277/ Published by Elsevier B.V.

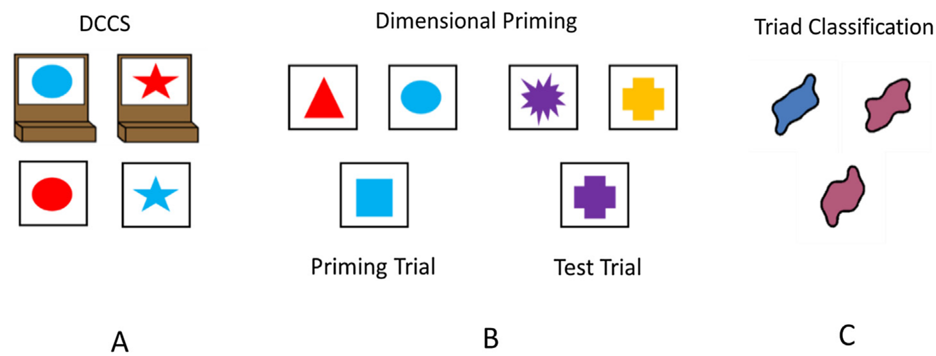

Fig. 1. Illustration of stimuli used in the DCCS (A), dimensional priming (B), and triad classification (C) tasks.

2012). Finally, during adolescence latent factors similar to those of adults are observable (Agostino, Johnson, & Pascual-Leone, 2010; Lee et al., 2013; Lehto, Juujarvi, Kooistra, & Pulkkinen, 2003; Rose, Feldman, & Jankowski, 2011; Wu et al., 2011). This literature is not without controversy, however. For instance, there is evidence for multiple components during early childhood, suggesting that the factor structure is impacted by the tasks, as well as which dependent variables are measured from those tasks (Miller, Giesbrecht, Müller, McInerney, & Kerns, 2012).

The latent variable approach is a data-driven method for identifying components of EF by determining how variance among measures from cognitive tasks cluster together. This approach is useful for describing the structure of EF, but it does not shed light on how EF works. To better understand the mechanisms and processes that give rise to EF, we instead take a theoretical approach to explore an alternative mechanism of EF: *dimensional attention to visual features*. To achieve these goals, we take a *process-based* approach to understand the dynamics that give rise to performance on tasks that require controlled processing of the visual features of objects. Previously, we have used neurocomputational simulations to develop a dimensional attention mechanism in the context of a canonical probe of EF during early childhood. In his report, our aim is to demonstrate how the processes that give rise to dimensional attention generalize to explain development across tasks that have distinct, and often competing, cognitive demands.

Attention is one of the most widely studied topics in cognitive development, so it is important to distinguish our definition of dimensional attention within this broader literature. Attention and EF are related to one another in many ways. For example, previous examinations of attention development have incorporated some aspect of endogenous control over attention as a central aspect of attentional development (Colombo & Cheatham, 2006; Sheese, Rothbart, Posner, White, & Fraundorf, 2008). Research has examined the foundations of attentional skills, including changes in the ability to orient, select, and control the allocation of attention from infancy through early childhood (e.g., Rueda, Posner, & Rothbart, 2005). Research has also examined the central role attention plays in the formation and maintenance of working memory for objects (Ross-Sheehy, Oakes, & Luck, 2011) and spatial locations (Schutte, Keiser, & Beattie, 2017), and the role of attention in inhibitory control (Reck & Hund, 2011). Much effort has also been given to examining the neural basis of attention development (Amso & Scerif, 2015; Richards, Reynolds, & Courage, 2010). Although these aspects of attention are likely foundational for the form of attention that we aim to study (Cuevas & Bell, 2014), in the current report we focus on dimensional attention during early childhood as a mechanism of EF. In particular, we explore the allocation of processing resources between different visual dimensions such as shape and color.

The dimensional attention mechanism that we have developed is grounded in a label system that is reciprocally connected to an object representation system. Dimensional labels, such as “color” or “shape”, modulate the activation of visual features within the object representation system. As associations are strengthened between neural representations for visual features and neural representations for labels of those features, the model comes to more strongly attend to visual features. One assumption of the model is that populations in frontal cortex are sensitive to both auditory and visual information which would allow neural representations to link labels with visual features and dimensions. In this way, labels act simply as another feature that is bound to visual features (Gliozzi, Mayor, Hu, & Plunkett, 2009). We have previously demonstrated that this mechanism unifies diverse behavioral findings and generates quantitative behavioral predictions (Buss & Spencer, 2014), as well as explains brain-behavior associations in early childhood (Buss & Spencer, 2018). Here, we explore whether this mechanism can generalize to explain performance on tasks that that require controlled processing of object features, but which are not typically characterized as EF tasks and do not use explicit verbal labels to instruct performance.

2. Theories of EF development

Various theories have been put forth to explain developmental changes in EF. These theories often focus on individual tasks as a benchmark measure of development. For example, the dimensional change card sort (DCCS) task has been the focus of many developmental theories and is often presented as a canonical probe of EF development. In this task, children are instructed to sort bi-dimensional test cards by either shape or color and then to switch and sort by the other dimension. Target cards are displayed at the trays where children sort to show which features go to which location for the shape or color rules (see Fig. 1A). The test cards children sort are constructed to contain conflict by matching both target cards along different dimensions.

The DCCS task has been a central focus in the literature for several reasons. First, it reveals a qualitative shift in children's abilities: 3-year-olds tend to perseverate and continue using the first set of rules when instructed to switch, but 4- to 5-year-olds have little difficultly switching rules. Thus, the task reveals the emergence of a new skill over a relatively short period of time. Second, subtle manipulations to the task can have amplified effects on children's performance (for a review, see Buss & Spencer, 2014; Zelazo, Muller, Frye, & Marcovitch, 2003). That is, children's abilities during this period are volatile and manipulations to different aspects of the task can illuminate the processes underlying performance. Finally, neuroimaging data have demonstrated that the development of flexible rule use is dependent upon activation in frontal cortex (Buss & Spencer, 2018; Moriguchi & Hiraki, 2009), which is widely regarded as the source of mature EF (e.g., Fuster, 2000). Thus, the DCCS task provides a wealth of behavioral and neural data to draw upon when explaining the development of dimensional attention.

The DCCS is an explicit dimensional attention task which requires children to understand and make use of verbal rules to sort cards according to the relevant dimension. Explaining performance in this task, however, is particularly challenging because it involves the coordination of multiple cognitive skills (Garon et al., 2008). First, the task requires an active representation of the currently relevant rules in working memory. This representation must then be used to selectively process the relevant feature of the test card and inhibit processing the irrelevant dimension. After the rules change, these processes must be updated—children must now form an active representation of the newly relevant dimension, selectively process the other dimension on the test card, inhibit processing of the previously relevant dimension, and inhibit the habits accumulated from the pre-switch phase. Thus, dimensional attention in the DCCS task involves the different components of EF that have been identified in the factor analyses approach: active working memory representations, inhibition, and switching. As we detail below, all of these functions are achieved by the processes that give rise to dimensional attention.

Explanations of performance in the DCCS task have centered on the mechanisms mentioned above. For example, it has been suggested that children fail during the post-switch phase due to a failure to inhibit attention to the pre-switch dimension which leads to attention being stuck on that pre-switch dimension (attentional inertia; Kirkham, Cruess, & Diamond, 2003; Rennie, Bull, & Diamond, 2004). Further, a connectionist model has been used to demonstrate that flexibility may develop from strengthening the active representations of the relevant rules (Chatham, Yerys, & Munakata, 2012; Morton & Munakata, 2002; Yerys & Munakata, 2006). Lastly, cognitive complexity and control (CCC) theory suggests that children have a limited ability to shift between representations of abstract rules created by a failure to reflect on the rules of the task (Zelazo et al., 2003). Although these accounts focus on specific mechanisms of EF, such as inhibition, representation strength, or shifting, they do not typically explain the processes that give rise to these mechanisms or developmental changes in these mechanisms. Consequently, these accounts of development are limited in their ability to generalize to other versions of the DCCS task or other types of tasks.

One perspective which expands beyond mechanisms and additionally emphasizes the processes underlying EF is dynamic field theory. Buss and Spencer (2014) presented a dynamic neural field (DNF) model that simulates real-time neural dynamics to explain the quantitative details of performance of 3- to 5-year-olds across a wide array of manipulations to the DCCS task. In the model, flexible rule-use arises from object representation processes which bind visual features (e.g., shapes and colors) to spatial locations (e.g., sorting locations). These object representation processes are coupled to a system which forms representations of dimensional labels (e.g., “shape” and “color”). Dimensional labels are instantiated as associations between labels and visual dimensions such as shape and color. For example, by activating a representation of the label “color”, the processing of color features in the object representation system of the model becomes enhanced. By enhancing the task-relevant visual dimension, the object representation system can overcome the conflict inherent to the task (i.e., the test card matches both target cards along different dimensions). As the model sorts during the pre-switch phase, memory traces accumulate which create biases to sort by the pre-switch dimension. These biases create an additional source of conflict that interferes with sorting by the post-switch dimension. Development in this task is implemented by increasing the strength of associations between label and visual feature representations. With weak associations, the model can sort by the pre-switch dimension but will perseverate when instructed to switch to sort by the other dimension. As the associations between labels and features are strengthened, the model transitions from performing like 3-year-olds to performing like 5-year-olds during the post-switch phase. That is, with strong associations, the model overcomes the biases that accumulate from the initial sorting phase and displays flexible sorting behavior in the DCCS task.

This theoretical perspective is unique for many reasons. Most notably, the model explains flexibility through general perception/action systems rather than abstract cognitive processes or cognitive representations, such as those used by the previously mentioned accounts. Specifically, flexible rule-use is emergent from the interactions between a system that builds representations of objects and a system that learns associations between labels and visual features. Second, the flow of information in the model is not unidirectional as in typical theories of EF. Rather, activation between the object and label systems flows in both directions between visual features and labels. Thus, in contrast to previous theories which frame mechanisms of EF around top-down mechanisms, the DNF model implements a dimensional attention mechanism that depends on reciprocal interactions between different levels of representations (i.e., labels and features). Lastly, flexible rule-use is emergent from real-time neural dynamics. Therefore, this perspective allows us to explain not only behavioral developmental results but also associated neural activation across different conditions of the DCCS (Buss & Spencer, 2018).

To better understand the processes involved in the development of dimensional attention, we look beyond the DCCS task and examine other tasks which also involve attention to visual features of objects but with other distinct processing demands. The DCCS task is an explicit rule-use task that requires using verbal rules to attend to the task-relevant dimension. The primary developmental achievement revealed by the DCCS task is the ability to flexibly shift attention between dimensions. To test the generalizability of this model, we included two other tasks which are measures of implicit dimensional attention. In these implicit dimensional attention tasks, participants must infer the relevant dimension from the configuration of the stimuli. This demand places a strong emphasis on the bi-directionality of interactions between visual features and labels because the processing of object features must be used to recruit attention to the relevant dimension. Moreover, these tasks reveal developmental improvements in other distinct functions of dimensional attentional, namely selectivity and stability. In the next section, we discuss developmental improvements in other aspects of dimensional attention.

3. Beyond flexibility

3.1. Stability

Stability is another function of dimensional attention that shows developmental improvements during early childhood. This function of dimensional attention has been previously assessed with a dimensional priming (DP) task (Medin, 1973). In this task, children are first shown a reference object (e.g., a blue square) and are then shown two choice objects. Children are instructed to pick the choice object that goes best with the reference object. The first two trials children receive in this task are called priming trials (see Fig. 1B). Priming trials are configured such that there is only one matching option among the choice objects. If the child is being primed on the color dimension, for example, then the choice objects could be composed of a red triangle and a blue circle, as shown in Fig. 1B. In this configuration, only one object matches the reference object and it does so along the color dimension. After two priming trials with only a color matching choice object, children are then given a series of ten test trials in which the shape and color dimensions are pitted against one another (see Fig. 1B). On these test trials, the two choice objects match the reference object along different dimensions. The typical finding in this priming task is that 5-year-olds are more likely to continue selecting the primed dimension compared to 3-year-olds.

Recently, Benitez, Vales, Hanania, and Smith (2017) demonstrated that stability in the DP task is related to flexibility in the DCCS task. Children were given the DP and DCCS task, both with the shape and color dimensions counterbalanced. Children who flexibly switched rules in the DCCS task were more likely to be primed and select based on the primed dimension for more consecutive test trials (see Fig. 2). These results lead to challenging questions. How can the same

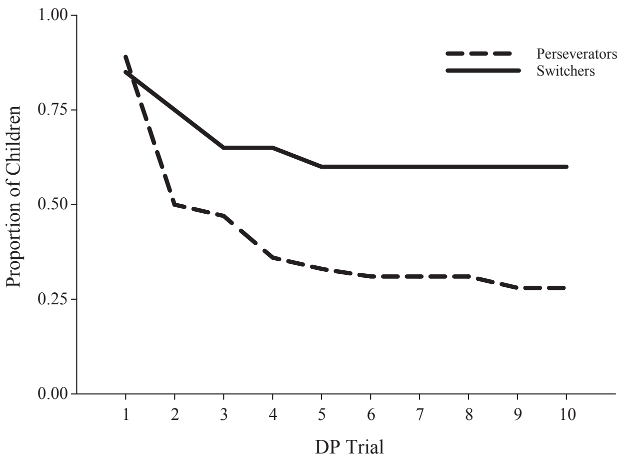

Fig. 2. Data from Benitez et al. (2017). Proportion of children who continued to select along the primed dimension in the dimensional priming (DP) task is higher among the children who switched rules in the DCCS task (Switchers) compared to children who perseverated in the DCCS task (Perseverators).

neurocognitive system give rise to flexibility in one context but stability in another? That is, stability in the DP task would seem to work in opposition to flexibility in the DCCS task—children who perseverate in the DCCS task would be more likely to continue selecting along the primed dimension in the DP task since children would be perseverating on a visual dimension in both cases. These data suggest that a more nuanced process drives the relationship in performance between these tasks. Additionally, if performance on these tasks tap into common aspects of dimensional attention over development, how is it that dimensional attention is driven explicitly in one task but implicitly in another? No previous theories have addressed these distinctions.

3.2. Selectivity

The development of dimensional attention has also been measured with a free classification task called the triad classification (TC) task (Smith & Kemler, 1977). The TC task is aimed at measuring implicit, selective dimensional attention. Similar to the DP task, children are given a series of trials in which they are shown a reference object and asked to pick which of two choice objects goes with the reference object. The stimuli are configured such that on every trial one object shares exactly the same feature of the reference object along one dimension while being maximally different along the other dimension. This object is called the identity (ID) choice object (object shown at the right in Fig. 1C). In contrast, the other choice object does not exactly match the reference object along either dimension but is overall more similar to the reference object when considering both dimensions (see object at the left in Fig. 1C). This object is called the holistic (H) choice object. Thus, the H object would be deemed a better match to the reference object if information is integrated across both dimensions. However, the ID object would be deemed a better match to the reference object if information is selectively considered along a single dimension. Previous studies have demonstrated that children's dimensional attention becomes more selective over development such that older children more frequently select the ID object over the H object compared to younger children (Smith & Kemler, 1977).

As with the DP task, the relevant dimension in the TC task must be determined implicitly based on the configuration of the stimuli since no explicit instructions are provided. In the TC task, the relevant dimension defined by the ID matching features is randomly assigned from trial to trial. Thus, there is no systematic dimension that is task-relevant as in the DP task. In this way, the TC task not only requires selectivity as described above, but also requires flexibly shifting attention between dimensions from trial to trial. No study has yet explored whether a relationship exists between performance on the DCCS and TC tasks (see Hanania & Smith, 2010). However, other research with the DCCS task suggests that selectivity and flexibility influence one another. For example, decreasing the demands on selective dimensional attention improves switching in 3-year-olds (Diamond, Carlson, & Beck, 2005; Kloo & Perner, 2005), and increasing the demands on selective dimensional attention makes switching more difficult for younger children (Fisher, 2011). These findings suggest that common processes might also underlie the development of flexibility and selectivity.

3.3. Summary

As this discussion reveals, various functions emerge over development in the context of attention to visual dimensions. Most theories of EF development are framed in such a way that they explain performance on individual tasks but do not offer insight as to how or why these functions of dimensional attention might be related from one task to another. For example, the attentional inertia account suggests that 'sticky' attention impairs DCCS performance, but it is not clear how the demands on stability in one task might be related to the demands on flexibility in another task. Further, these tasks involve either explicit or implicit dimensional attention which creates additional challenges for accounts of EF development that are centered on the DCCS task. For instance, the CCC account frames the development of cognitive flexibility in terms of reflection on the nature of the objects and the rules to be used when sorting. Thus, it is not clear how this account would explain performance on implicit measures of dimensional attention which do not specify any rules. Further, the connectionist model of the DCCS task implements a top-down mechanism but does not have a mechanism for bottom-up influences that would be required for these measures of implicit dimensional attention. Indeed, none of these accounts explain how bottom-up factors, such as the color and shape of the stimuli, can lead to the deployment of top-down dimensional attention.

In this report, our aim is to explore whether these functions of dimensional attention can be integrated within a common framework. The primary question we ask is whether the DNF model can explain how multiple cognitive functions arise from a common set of neural process. In the next section, we describe the DNF model that has been used to explain brain and behavior development in the DCCS task. We then describe how the dimensional attention mechanism adapts in the context of the DP and TC tasks. A key aspect of this mechanism is that it is driven through reciprocal connectivity between the object representation system and the dimensional label system. Specifically, recruitment of top-down modulation of information processing in the object representation system is dependent upon bottom-up signals from the object representation system. Across these different tasks, functions of dimensional attention emerge based on the demands imposed by a specific task and the developmental status of the neurocognitive system.

4. The dynamic neural field model

Dynamic Field Theory is a class of neural process models (called dynamic neural field, DNF, models) that explains how cognition arises from neural dynamics (for a review, see Schoner & Spencer, 2016). Dynamic Field Theory has a substantial history in being applied to the development of dimensional attention. Buss and Spencer (2014) first developed a DNF model to simulate the performance of 3- to 5-year-olds in an unprecedented 14 different variations of the DCCS task. Moreover, this model has also been used to predict: behavioral data in new variations of the DCCS task (Buss & Spencer, 2014), how prior exposure to visual dimension influences DCCS performance (Perone, Molitor, Buss, Spencer, & Samuelson, 2015; Perone, Plebanek, Lorenz, Spencer, & Samuelson, 2017), and neural activation across frontal, parietal, and temporal regions (Buss & Spencer, 2018). Thus, the DNF model offers

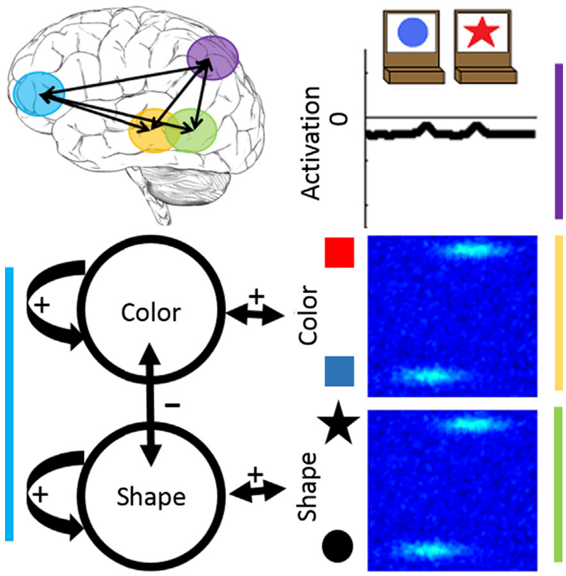

Fig. 3. Architecture of the DNF model and mapping to cortical regions. The spatial field is shown at top (purple). Below the spatial field is the color-space field (yellow) and shape-space field (green). On the left are the dimensional label units (blue) which represent labels for "color" and "shape". Inset brain image illustrates the mapping of model components to cortical regions. The double-headed arrows in the figure represent reciprocal relationship between the regions. Looped arrows in the dimensional label units represent self-excitation within each unit. The double headed arrow between the dimensional label units represents lateral inhibitions between each unit. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the most comprehensive explanation available of the behavioral and neural data from the DCCS task.

4.1. Model architecture

The DNF model is composed of populations of neural units that correspond to frontal, temporal, or parietal cortical processing. Neural units interact through local-excitation and lateral-inhibition to create 'peaks' of activation that correspond to cognitive representations or decisions about stimuli. Neural units have a learning mechanism which gradually increases the baseline level of activation for these units as they are activated over the course of a task.

Fig. 3 shows the architecture of the model and the mapping of model components to cortical regions. The parietal cortex component (purple) is composed of a population of neural units that is tuned to the spatial information of the task. Peaks within this field reflect representations of the spatial locations of objects or spatial decisions about objects in the context of the task-space. This component is coupled to a temporal cortex component (yellow and green) composed of a set of 2-dimensional fields. These fields are tuned to a combination of feature information (color or shape) and spatial information. The parietal and temporal fields share activation with one another along the spatial dimension. That is, when neural units are activated within the model, they pass activation to other units that are tuned to similar spatial information. Through this spatial coupling, these fields implement an object representation system that forms representations of objects by binding representations of shape and color features to spatial locations.

The parietal and temporal components are reciprocally connected to a frontal cortex component (blue). The frontal component implements representations of labels for features, dimensions, or objects. In the model used for the following simulations, the frontal component consists of units corresponding to "shape" and "color" labels. These label units have reciprocal connectivity to their corresponding feature population in the temporal component. Additionally, these label units are homogenously connected to the parietal component. This connectivity

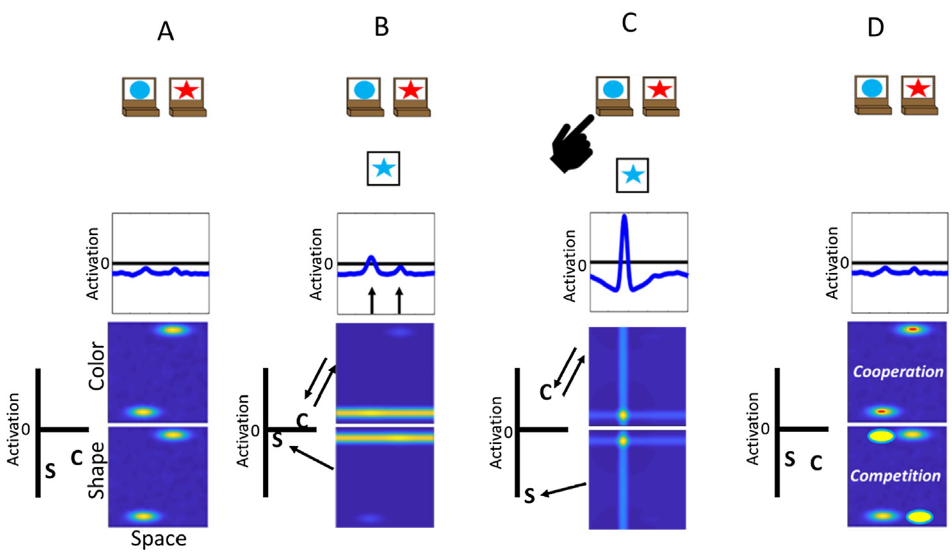

Fig. 4. Model performing a trial during the pre-switch of the DCCS task. Top: A representation of the stimuli shown to the model. Middle row: Activation profiles in the spatial field with the horizontal line representing the activation threshold (at 0). Bottom panels: Color-space (top) and shape-space (bottom) fields, along with their dimensional label units (C and S). When the label units are above the activation threshold (horizontal line at 0), activation is sent to the associated feature-space fields. In all figures, the thickness of the arrows is used to indicate the strength of activation with bolder arrows representing stronger activation than thinner arrows. (A) Model is shown the target cards and sorting trays. The 'color' label unit on the activation map (C) is at a higher level of activation due to a direct input reflecting that the model has been instructed to sort by color. (B) The model is presented with a blue star test card. Activation in the spatial field is equally strong at the left and right locations because the test card overlaps with both target cards along different dimensions (note the arrows below the spatial field). Further, activation is also equally strong within both the color and shape systems because the test card overlaps with the target cards within each dimension (note the arrows from the feature-space fields to their dimensional label units). The "color" label unit (C) has a competitive advantage due to the direct input and is sending activation back to the color-space field (note the arrow from the C label unit to the color-space field). (C) The model sorts the test card to the leftward location. (D) The model is shown at the beginning of the post-switch phase. The target card inputs and memory traces (illustrated as yellow ovals, the darker the color the stronger the memory trace) overlap with one another in the pre-switch field (noted with *cooperation*) whereas the memory traces are at the opposite location of the target card inputs within the post-switch field (noted with *competition*). The input to the C label unit is now removed, and a direct input is now given to the S label unit. The advantage provided to the S label unit, however, is not as strong due to the memory trace that has accumulated on the C label unit during the pre-switch phase. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

serves to build responses within the parietal component when a stimulus is presented.

4.2. Implementing the DCCS task

Fig. 4 shows the model over a sequence of events as it performs the DCCS task. Note that the arrows in this and subsequent figures of the model illustrate the strength of interactions along the spatial dimension and between labels and features. These are meant to facilitate comprehension of the model dynamics that are important in each task. The arrows illustrate the direction of interaction that is present at each time point and the boldness of the arrow is meant to illustrate the relative strength of interactions. In Fig. 4A, the model is displayed with inputs corresponding to the target cards (displayed at top). The model has small bumps of input at the left and right locations, corresponding to the locations of the sorting trays. Across the shape and color fields, the model has bumps of input at the leftward location for the blue and circle features in the color and shape fields, respectively. At the rightward location, the model has bumps of input corresponding to the red and star features in the color and shape fields, respectively. These inputs are sub-threshold and do not induce activation. Rather, these inputs served to pre-shape the fields based on the structure of the task. For this trial, the model is instructed to sort by color; thus, the "color" label unit, plotted in the bottom left of Fig. 4A, has a boosted level of activation from a direct input. As with the target card inputs, this rule input pre-shapes activation of the dimensional labels but does not induce activation.

In Fig. 4B, the model is given a blue star test card to sort by color. The test card is presented as a ridge of activation at the relevant feature values within the shape and color fields (note the horizontal stripe of activation). In this way, the test card contains no spatial information regarding where it is to be sorted. Rather, the model must use the overlap of the target card inputs and test card inputs along with the instructed dimension to make a decision. At this point in time, the "color" label unit and "shape" label unit are both being boosted in a bottom-up fashion due to the build-up of activation within the shape and color fields at the presentation of a test card (note the arrows drawn between the feature fields and the dimensional label units). However, the "color" label unit has an advantage from the input provided by the instructions to sort by color. The arrows drawn below the spatial field in this panel indicate the ambiguity in the stimuli—that is, the contributions of input across all fields at the left and right spatial locations are equivalent. Through activating the label for "color", however, the model enhances processing of color features and makes a decision based on color information. Specifically, the model uses the overlap of the inputs from the target card and test card to bind the features of the test card to the leftward location as illustrated in Fig. 4C. This decision is reflected by the peak of activation at the leftward location in the parietal field and the combination of peaks for the blue and star features at the left location in the feature space fields.

Fig. 4D shows the consequences of making decisions during the pre-switch phase. As mentioned above, the model has a learning mechanism which results the build-up of memory traces associated with the activated feature-space conjunctions. Within the color-space field, the memory traces overlap with the target input for the blue feature at the leftward location and the red feature at the rightward location. However, within the shape-space field, the memory traces are at the opposite location of the target card input. These memory traces create *co-operation* within the pre-switch color field, meaning that activation will build more quickly at these locations on subsequent test card presentations. These memory traces, however, create *competition* within the post-switch shape field. In this case, activation will build more slowly in this field at the target card location due to the location of the memory trace generating lateral inhibition when a test card is presented. Specifically, because the model has built a memory of sorting the start to the leftward location during the pre-switch phase, it will be more difficult for the model to build activation at the rightward location for this feature.

The original implementation of the model did not use active units for the "shape" and "color" labels of the frontal component. Buss and Spencer (2014) directly modulated the activation within the color and shape dimensions to reflect changes in the strength of dimensional attention. More recent instantiations of the model have used active units for the frontal component to simulate hemodynamic responses in frontal cortex (Buss & Spencer, 2018) as well as the influence of prior exposure to visual dimension on DCCS performance (Perone et al., 2015, 2017). Here, we continue using the model architecture with active frontal units to enable simulating both explicit and implicit dimensional attention tasks. These more recent instantiations of the model used a 'young' and an 'old' model to simulate the performance of 3.5- and 4.5-year-olds on the DCCS, respectively. These models were defined based on parameter differences for two different types of coupling within the model. First, relative to the 'old' model, the 'young' model had weaker local-excitation/lateral-inhibition within the frontal component of the model. Second, relative to the 'old' model, the 'young' model had weaker reciprocal connectivity between the frontal and posterior systems. Taken together, the 'old' model activated representations in the frontal component more quickly and strongly and was able to boost activation within the posterior system more strongly. As a consequence, the 'young' model produced low rates of switching similar to 3.5-year-olds, while the 'old' model showed higher rates of switching similar to 4.5-year-olds.

This model is the first computational framework to simulate the trial-to-trial response of children in a way that captures both pre- and post-switch performance. Aside from the explanation offered for the existing literature, our theory has also been used to generate data that calls into question assumptions of previous accounts, motivating a reinterpretation of the cognitive processes giving rise to performance on the DCCS task over development. First, the emergence of flexibility in the DCCS task is not only associated with increased involvement of frontal cortex as suggested by previous theories (Bunge & Zelazo, 2006; Moriguchi & Hiraki, 2009; Morton & Munakata, 2002), but also increases in activation of parietal and temporal regions (Buss & Spencer, 2018). Moreover, frontal cortex activity can be influenced by manipulations to the task. For example, children who perseverate in the standard condition, and show weak frontal activity when doing so, produce strong activation of frontal cortex in an easier version of the DCCS task in which they can correctly switch rules (Buss & Spencer, 2018). This pattern of data is explained by stronger activation sent to the frontal label system from the posterior object system in the easier version of the task.

Behaviorally, it has also been demonstrated that giving children pre-exposure to the post-switch dimension can facilitate performance of 3-year-olds in the DCCS task. Perone et al. (2015, 2017) administered a memory game where children flipped over cards to find matching features before performing the standard DCCS task. If the memory game involved the same dimension used during the post-switch phase of the DCCS task, children performed significantly better in the DCCS task compared to when the matching task did not involve the post-switch dimension. Again, this change in performance occurs through boosting activation for the relevant dimension in the posterior object system which increases the strength of interaction between the frontal and posterior systems. Lastly, (Buss & Spencer, 2014) directly manipulated the spatial configuration of the DCCS task to either increase activation within the posterior system which facilitated switching of 3-year-olds or increase inhibitory competition in the posterior system which impaired switching of 3-year-olds. These studies were all based on predictions generated by the same model regarding the behavior or hemodynamic activity which were supported by data from 3.5- and 4.5-year-olds. In summary, the model performs the DCCS task and its variants in a similar fashion as children and replicates patterns of hemodynamic activity across frontal and posterior regions.

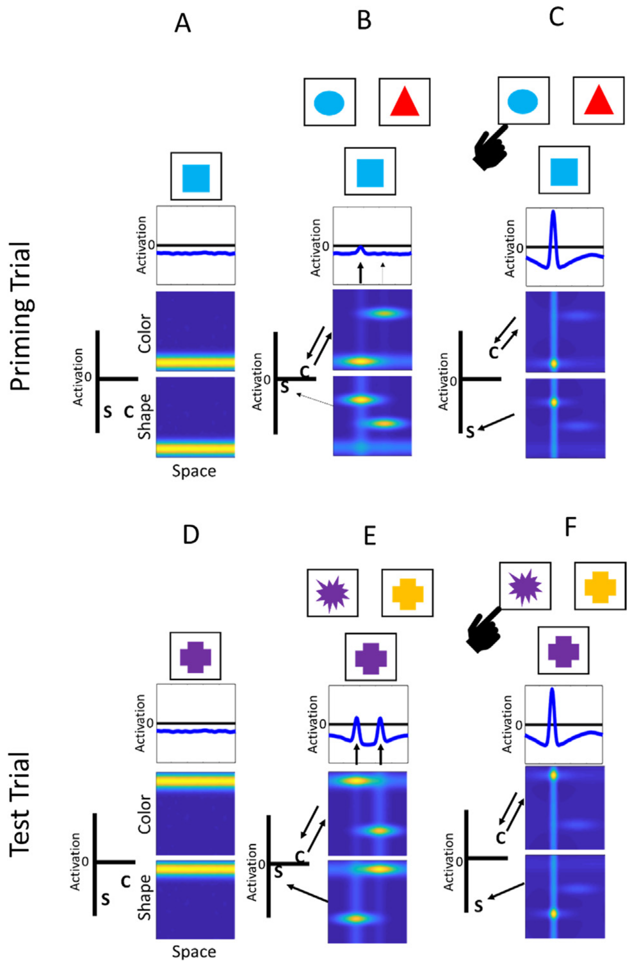

Fig. 5. Model performing the dimensional priming task. A–C illustrate a priming trial and D–F illustrate a test trial. (A) Model is shown the reference object. In contrast to the DCCS task, the model is not provided with a direct input to the dimensional label units. Therefore, the activation levels of the C label unit and S label unit are the same. (B) Model is shown the two choice objects. Spatial activation is stronger at the leftward location where the choice object matches the reference object (note the bold arrow at the leftward location under the spatial field). Activation is stronger within the color system due to the overlap of color features between the reference object and the choice object (note the bold arrow from the color-space field to the C label unit). (C) Model selects the choice object on the left (note the reciprocal interactions within the color system). (D) Model is shown the reference object after two priming trials. Note that C label unit has stronger resting-state activation due to memory traces that accumulated during the priming trial. (E) Model is presented with the choice objects. Spatial activation is equally strong at both spatial locations because the choice objects both match the reference object along one dimension. Activation is also equally strong from the shape-space and color-space fields to their label units because there is overlap between the reference object and choice objects within both dimensions. Due to its higher resting-state activation, the C label unit passes the activation threshold and sending activation to the color-space field. (F) The model selects the choice object on the left that matches along the color dimension.

An important assumption of the DNF model is that EF skills are built from the perceptual and motor processes involved in specific tasks. If this is the case, then the same model should be able to simulate behavioral performance from other tasks that are not explicitly EF tasks, but still require attention to visual dimensions. Next, we test the generalizability of this model in the context of other dimensional attention tasks. First, we describe the inputs provided to simulate the DP and TC tasks and how the dynamics of the model generate responses in these tasks. We then run batches of the 'young' and 'old' model to determine whether these models explain quantitative associations in performance across tasks.

4.3. Implementing the DP task

Fig. 5 shows the sequence of events as the DNF model performs the DP task. Fig. 5A shows the input for the reference object. Similar to the test card input in the DCCS task, the reference object input is given as a ridge of activation (note the horizontal stripes of activation) to the shape and color fields without specifying location. This input does not induce above threshold activation, but instead serves to pre-shape activation in the field. Fig. 5B–C shows the presentation of a priming trial. The choice objects are administered as a set of spatially localized inputs at a leftward and rightward location. These inputs are strong and induce above-threshold activation. In the color field, the leftward object overlaps with the ridge for the reference object while the rightward object is at a different color value. In the shape field, both objects are at a different feature value than the shape feature on the reference object. The arrows drawn below the spatial field indicate that the inputs aggregated across all fields are stronger at the leftward location compared to the rightward location. This is due to the overlap of the reference object feature with the choice object feature at the leftward location—that is, both objects are blue. In this way, the imbalance of spatial activation associated with the stimuli indicate which object is the correct choice.

The activation of the "shape" and "color" label units is plotted at the bottom left of each panel. When the choice objects are presented, as shown in Fig. 5B, the "color" label unit gains stronger activation than the "shape" label unit. As indicated by the arrows drawn between the feature fields and the dimensional label units, the overlap of the reference and target input creates stronger synaptic output from the color field to its associated dimensional label unit. Note that this is an implicit dimensional attention task; therefore, the model is not given a direct input to the dimensional labels as in the DCCS task. The configuration of stimuli creates a signal regarding which dimension is relevant. Fig. 5C shows the model selecting the color-matching object on the leftward location which is indicated by the peak in the spatial field. At this point in time, the model has strongly activated the "color" label unit and is attending to the color dimension.

Recall that in the DP task, children are first given two priming trials followed by ten test trials. Fig. 5D–F illustrate the sequence of events on a test trial in the DP task. First, Fig. 5D shows that the activation of the "color" label unit is stronger than that of the "shape" label unit. This difference is due to the accumulation of a memory trace for this dimensional unit during the priming trials. In Fig. 5E, the model is given stimuli corresponding to the test items. Now, the reference object overlaps with a choice object input within each field (note the arrows below the spatial field indicating equal input strength at each spatial location across fields). Specifically, the reference object overlaps with the choice object on the left within the color field and overlaps with the choice object on the right within the shape field. In Fig. 5E, the "color" label unit is at an advantage relative to the "shape" label unit due to the memory trace. In Fig. 5F, the model has selected the choice object at the leftward location which matches along color. Thus, in this example, the model has become primed on the relevant dimension and has selected the choice object that matches along color. It is important to point out one difference in the implementation of this task relative to the DCCS task. Although memory traces accumulate within the label units, the accumulation of memory traces *within the feature fields* is disabled for this task. This change was implemented for practical purposes. The DP task uses different features from trial-to-trial and there is no feature-space conjunction built into the task as there is with the DCCS. Rather than implementing a larger field that would allow for sampling a wider range of features but would require more time to perform the field computations, we opted to eliminate the accumulation of feature-space memories. Thus, from the model's perspective, new features are

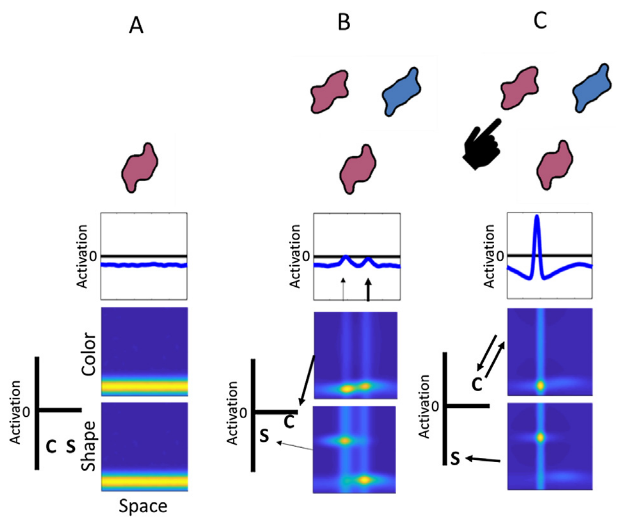

Fig. 6. Model performing the triad classification task. (A) Model is shown the reference object. As in the dimensional priming task, the model is not provided with a direct input, so the activation levels of the C label unit and S label unit are the same. (B) Model is presented with the choice objects (H item on the right, ID item on the left). Activation is stronger at the location of the holistic object (note bold arrow at rightward location under the spatial field) due to the partial overlap of features with the reference object within both dimensions. Activation is also stronger within the color system (note the bolder arrow from the color-space field to the C label unit) because of the direct overlap of the color feature on the ID object (left) with the reference color feature. (C) Model selecting the ID object (note reciprocal interactions within the color system).

presented on every trial since it does not have a mechanism to remember which specific features were presented on previous trials. As illustrated by this example, however, the key property of the model that creates dimensional priming is the strength of the memory traces accumulated on the dimensional label units during the priming trials.

4.4. Implementing the TC task

Now we turn to the TC task. This task is similar to the DP task in terms of the demands on implicit dimensional attention. In both tasks, the relevant dimension must be deduced from the configuration of stimuli. However, there are new demands placed on flexibility in the TC task. That is, in the TC task, the relevant dimension is randomly selected from trial to trial; thus, children must be able to not only determine which dimension is relevant based on the configuration of stimuli but must also be able to selectively attend to shapes or colors in a flexible manner. Further, the source of competition induced by the non-target object is different than in the DP task. Specifically, the H object is similar to the reference object along both the relevant and irrelevant dimension.

Fig. 6 illustrates the sequence of events and configuration of stimuli in this task. In Fig. 6A the model is given the reference object similar to the priming task (note the horizontal stripes of activation). Again, this is a pre-shaping input that does not induce above-threshold activation. In Fig. 6B, the model is shown the ID choice object on the left and the H choice object on the right. Note that the ID choice object directly overlaps with the ridge for the reference object in the color field but is maximally far away from the reference object input within the shape field. The H choice object, on the other hand, is just adjacent to and partially overlapping with the reference object input within both the shape and the color fields. The source of competition is illustrated with the arrows at the bottom of the spatial field. Specifically, there is stronger excitation from the stimuli at the rightward location due to the partial overlap of the features of the H choice and the reference objects at the rightward location. Selective dimensional attention in the model is illustrated with the arrows between the feature fields and the dimensional label units. Specifically, there is stronger activation within the field containing the ID choice object because of the direct overlap of the ridge input for the reference object and the feature-space input of the identity choice object. In contrast, the other field generates weaker excitation because only one of the inputs for the two choice objects overlaps with the reference object input. Thus, at the point in time illustrated in Fig. 6B, the “color” label unit has a competitive advantage due to the stronger input being received from the color feature field. In Fig. 6C, the model is shown selecting the ID choice object. This is reflected by the presence of an activation peak within the spatial field at the location of the ID choice object. At this point, the model has also activated the “color” label unit and is attending to the color dimension to make this decision.

As with the DP task, the formation of memories within the feature fields is disabled for this task. Again, this is a simplification is for practical purposes, reflects the random variation in features from trial to trial, and is based on the assumption that children do not form systematic feature-space biases in this task. However, memory formation on the dimensional label units is enabled since the same dimensions are retained across trials.

4.5. Summary

As illustrated above, the model that explains development in the DCCS task is also able to perform the DP and TC tasks. Modeling these new tasks required no modification to the model architecture, but simply a reconfiguration of the stimuli to reflect the implementation of each task. Although these example illustrations are useful to think about the common and distinct processing demands across tasks, they do not tell us the nature of associations in performance between tasks. That is, at the individual level, is performance on these tasks associated with one another? In the next section, we present the results of batches of the ‘young’ and ‘old’ models which have distributions of parameters corresponding to the strength of local-excitation/lateral-inhibition within the frontal component and the strength of reciprocal connectivity between the frontal and posterior systems. These distributions reflect individual differences in the developmental status of children between the ages of 3 and 5. We use these ‘young’ and ‘old’ models to determine whether the model is able to explain the existing association between performance in the DCCS and DP tasks (Benitez et al., 2017). Further, we use these models to make formal predictions regarding the relationship between DCCS and TC tasks.

5. Explaining and predicting dimensional attention development

In this section, we ask whether the model can provide a formal account of dimensional attention development that goes beyond the DCCS task. First, we determine whether the model replicates the data presented by Benitez et al. (2017). Specifically, children who showed stable dimensional attention in the DP task were more likely to switch rules in the DCCS task. Do models that switch in the DCCS task also become primed in the DP task? Considering the processing demands imposed by the DP task, it is not trivial that a model should be able to simulate these effects. In particular, the DP task taps into implicit dimensional attention whereas the DCCS task measures explicitly cued dimensional attention. Additionally, these tasks require distinct dynamics of dimensional attention: stability is required in the DP task, but flexibility is required for the DCCS task. Can the model show stability in the DP task while showing flexibility in the DCCS task? Second, we examine the predictions the model makes regarding associations in performance between the DCCS and TC tasks. The TC task is also a measure of implicit dimensional attention that requires selectively processing a single visual dimension. Can the model that shows flexibility in an explicit rule-use task also show selectivity and flexibility in the TC task?

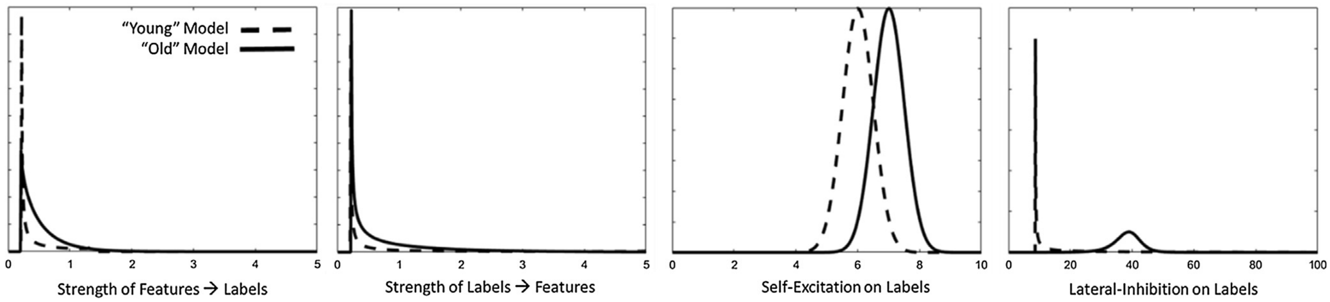

Fig. 7. Probability density functions for the parameters manipulated to explain performance of 3.5-year-olds ("Young" model) and 4.5-year-olds ("Old" model). See Table 1 for lists of parameters used to define these distributions.

5.1. Simulation methods

Simulations were conducted in Matlab 8.4 (Mathworks, Inc.) on a PC with an Intel® i7™ 3.5 GHz quad-core processor. We defined 'young' and 'old' models based on distributions of parameters for four interactions that are critical for the model's performance: from feature fields to associated dimensional units, from dimensional units to associated feature fields, the strength of self-excitation on the dimensional units, and the strength of mutual inhibition on the dimensional units. Fig. 7 shows the distributions used for these parameters (see Table 1 for the parameters defining these distributions). As illustrated in these plots, these parameters become stronger from the 'young' to the 'old' model. The full set of parameters for the model is shown in Tables S1–S3. The equations were the same as reported by Buss and Spencer (2014). The 'old' and 'young' models were iterated for 200 runs each (corresponding to 200 participants). For a given run, a parameter was generated from each of the four distributions that defined the 'young' and 'old' models. Each run of the model was given the DCCS, the DP, and the TC tasks.

5.1.1. The DCCS task

The DCCS task was simulated in the same manner as in Buss and Spencer (2014; 2017). The model was given five trials during each of the pre- and post-switch phases. To 'instruct' the model about which dimension is relevant for each sorting phase, the model was given an input of strength 0.5 to one of the dimensional units during the pre-switch phase. During the post-switch phase, the input was removed from this dimensional unit and given to the other dimensional unit. Throughout each simulation, target inputs were presented at specific feature and spatial values to capture the relevant details of the targets cards for the pre-switch and post-switch phases. After 500 time-steps, the model was presented with ridges of input for the features displayed on the test cards. Each trial was simulated for 1500 time-steps, with the test card stimulus presented for 1000 time-steps. The models always generated an active response by the end of the trial.

5.1.2. The DP task

The sequence of inputs for DP task was administered to mimic key properties of the task. It should be noted that no inputs were given to the dimensional units in this task since it is a measure of implicit dimensional attention and no direct instructions are given to children. First, inputs were presented that represented the reference object. This was administered as a 'ridge' at the target feature value across all spatial locations (see Fig. 5). After 500 time-steps, inputs were presented that corresponded to the choice objects. During the first two trials (the priming trials), the choice objects only overlapped with the reference ridge in one field. During the subsequent ten test trials, the model was given inputs for choice objects that each overlapped with the ridge in one field. The model always generated a response by the end of the 1500 time-steps trial.

Table 1
Parameters defining the probability density functions for developmental parameters manipulated in the model.

		Mean	St-Dev	Skew	Kurtosis
Label → Feature	"Old"	0.9	1	2.5	11
	"Young"	0.6	1	2.95	11
Feature → Label	"Old"	0.52	0.35	2.15	10
	"Young"	0.3	0.3	2.75	10
Label Self-Exc	"Old"	7	0.05	0.8	18
	"Young"	6	0.05	1	18
Label-Inhib	"Old"	38	1	−1	12
	"Young"	15	15	3	12

5.1.3. The TC task

The sequence of inputs for the TC task was administered to mimic key properties of the task described in the experiment reported below. First, inputs were presented that represented the reference object. This was administered as a 'ridge' at the target feature value across all spatial locations (see Fig. 6). After 500 time-steps, the choice objects were presented. These inputs are spatially-localized and the features were configured to reflect either the ID choice object or the H choice object. The inputs for the H choice object were centered 5 units away from the center of the reference ridge. The input for the ID match was centered on the center of the reference ridge in one field but was centered maximally far away (20 units) from the center of the reference field in the other field. The relevant dimension in which the ID choice object matched the reference ridge was randomly selected on each trial with the constraint that the trials were evenly divided between the shape and color dimensions. To match the properties of the other two tasks, the TC was administered with 12 trials.

It is important to note that different strengths were used for different inputs. However, these strengths were matched across tasks to constrain these simulations. The target card inputs in the DCCS and the reference object input for the DP task and TC task were a strength of 2 and width of 4 units. Importantly, these inputs were sub-threshold and would not induce a peak of activation by themselves. Rather, they serve to pre-shape activation within the neural fields based on the structure of the task. However, the test card input in the DCCS and the choice object inputs for the DP and TC tasks had a strength of 7.75 and a width of 6. Based on this configuration, the test card input and choice object inputs would induce a peak of activation based on their interaction with the pre-shaping inputs.

5.2. Results

Simulation results for the DCCS task replicated previous simulation and behavioral results (e.g., Buss & Spencer, 2014, 2018). All models sorted correct during all pre-switch trials. Models were scored as passing the DCCS task if they sorted at least 4 out of 5 trials correctly during the post-switch phase and as perseverating if it sorted 1 or fewer correctly during the post-switch phase. The majority of runs of the

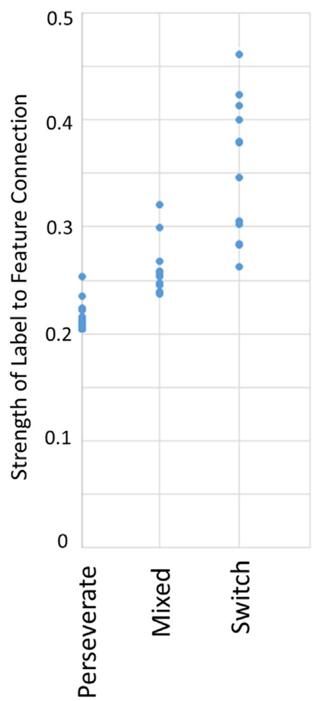

Fig. 8. Illustration of the relationship between model parameters and performance on the DCCS task. As the strength of coupling from the labels to the feature fields is strengthened, the model transitions from a tendency to perseverate to a tendency to switch with mixed performance occurring within a critical range of values.

'young' model failed to switch rules (67.5%), relatively few correctly switch rules (27.5%), and even fewer showed mixed responding (5%). The majority of runs of the 'old' model correctly switched rules (64%), relatively few failed to switch rules (27%), and even fewer showed mixed responding (9%). Fig. 8 illustrates the relationship between model parameters and performance on the DCCS task using the parameter for the strength of connection from the dimensional label units to the feature-space fields as an example. At low parameter values, the model tends to perseverate, but at high values the model tends to switch rules. Mixed performance occurs within a range of parameters values where neither switching nor perseverating are stable.

Compared to the findings of Benitez et al. (2017), the model produced a similar relationship between performance on the DCCS and the DP tasks. Fig. 9 represents proportion of models that selected along the primed dimension over the course of the full experiment. That is, models that switched rules in the DCCS task continued to select along the primed dimension in the DP task for more consecutive trials ($M = 8.07$), compared to models that perseverated ($M = 3.95$) (t (3 6 9) = 11.209, $p < .001$). Further, models that switched rules also selected the primed-match option on a higher number of trials ($M = 8.90$) than models that perseverated ($M = 6.46$) (t (3 6 9) = 11.068, $p < .001$). Although the model performed the DP task better than children in the Benitez et al. (2017) study, the average number of trials the models selected from the primed dimension are similar to those reported by Benitez et al. (2017) for both consecutive trials ($M = 6.50$ for switchers and $M = 4.03$ for perseverators) and across the length of the experiment ($M = 8.00$ for switchers and $M = 6.61$ for perseverators).

We next compared performance in the DCCS and TC tasks (see Fig. 10A). Here, we also observed an association in performance between tasks. Models that switched rules in the DCCS task selected the identity object in the TC task at a significantly higher rate ($M = 0.94$) than models that perseverated ($M = 0.66$) (t (3 6 9) = 22.466, $p < .001$). To explore dimension-switching effects on performance, we

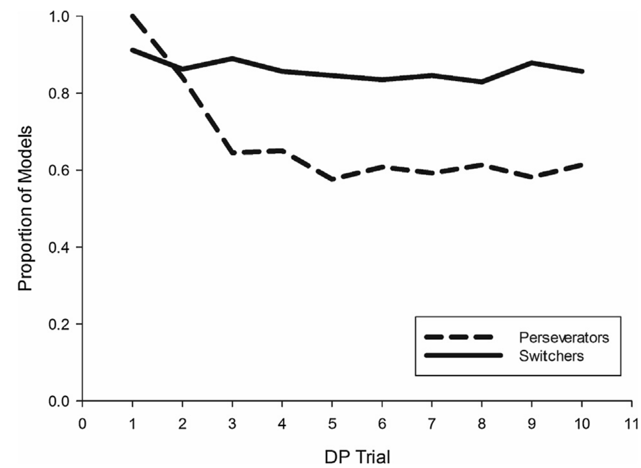

Fig. 9. Proportion of models that selected based on the primed dimension over the course of dimensional priming (DP) trials in the DP task. Solid line plots data for models that switched in the DCCS task, and the dashed line plots data for models that perseverated in the DCCS task.

conducted a supplementary batch of 50 simulations each of the 'young' and 'old' models that only ran the DCCS and TC tasks. We increased the number of trials simulated for the TC task to 36 to more robustly estimate trial-to-trial performance. Specifically, we examined the model's performance on trials which the dimension of the identity-match switched from the previous trial as a function of the number of trials which the other dimension repeated. As shown in Fig. 10B, models that switched rules in the DCCS did not differ in performance based on whether there were one or two repetitions of the other dimension prior to the dimension switch ($t(37) = 0.408$, $p = 0.69$). However, for models that failed to switch rules in the DCCS, performance was significantly poorer when the previous dimension repeated twice compared to when the previous dimension repeated once on switch trials ($t(50) = 2.443$, $p = .018$).

5.3. Discussion

The model results revealed a significant relationship between the DCCS task and the DP task, replicating Benitez et al. (2017), as well as a significant association between the DCCS task and the TC task. Regarding the association between the DCCS and the DP tasks, the model offers insight about the nature of the processes underlying performance across these two tasks. Performance in both tasks depended on the strength of connectivity between the frontal and posterior systems and connectivity within the frontal system. The 'young' model with weak coupling between dimensional units and feature fields, and weak local-excitation/lateral-inhibition among the dimensional units did not robustly engage the dimensional units during the priming trials, leading to a weaker accumulation of memory traces on these units relative to the 'old' model. The 'old' model robustly engaged the dimensional units, built up stronger memory traces on the relevant dimensional unit, and was able to consistently engage this primed dimensional representation over the course of the test trials. In relation to DCCS performance, across both the 'young' and the 'old' models, runs of the model that were able to correctly switch rules in the DCCS task showed significantly higher levels of priming over the course of 10 test trials in the DP task. Moreover, the model also produced an association between performance in the TC task and the DCCS task. Runs of the model that correctly switched rules in the DCCS task showed a higher rate of ID choice object selection. Next, we examine these predictions in an experiment with 3- and 4-year-olds.

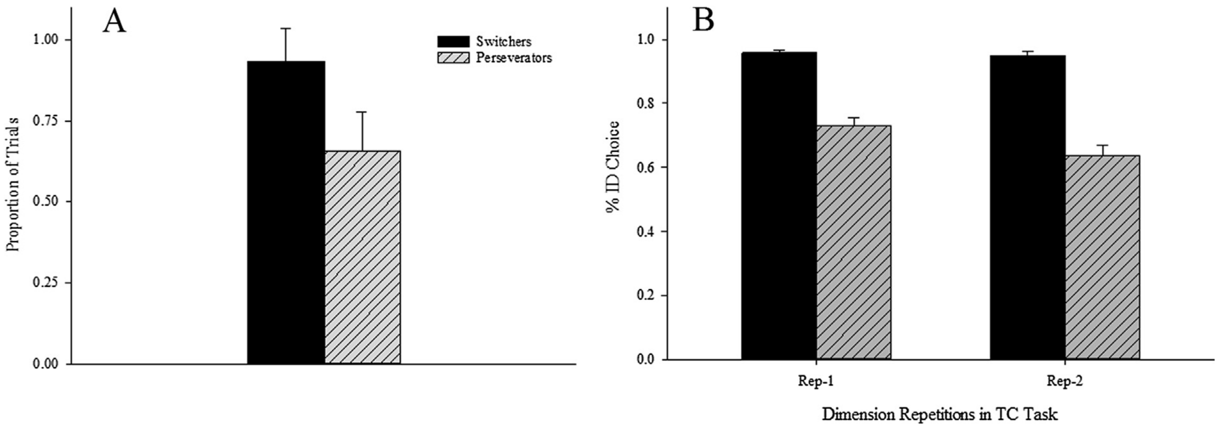

Fig. 10. (A) Proportion of trials the model selected the identity choice object in the triad classification task based on DCCS performance. (B) Performance of model during TC task on dimension-switch trials as a function of number of repetitions of other dimensions before the dimension switched.

6. Testing model predictions

We tested the predictions generated by the model with a group of 3- and 4-year-olds who were given both the TC and DCCS tasks. The TC task provides a quantitative metric of performance based on the proportion of trials on which the ID object is picked. To obtain a quantitative metric in the context of flexible dimensional attention, we administered the NIH Toolbox version of the DCCS (Zelazo & Bauer, 2013). This version is similar to the standard DCCS task, except that a mixed block is included after the post-switch phase. During this mixed block phase, all of the features in the task are changed from the previous sorting phases. Thirty trials are administered, 10 of which children are instructed to sort by the pre-switch dimension and 20 of which children are instructed to sort by the post-switch dimension. In the NIH Toolbox version, the mixed block is only administered to children who pass the post-switch phase. However, we administered the mixed block to all children with the aim of detecting more variation in children's dimensional attention switching abilities. Thus, we can measure whether children pass or fail based on their post-switch performance as is traditionally done with this task. We can also generate a quantitative measure of flexibility based on the number of trials sorted correctly during the mixed block. Functional near-infrared spectroscopy data were also collected from all children, but those data will not be reported here since we are focusing on the behavioral predictions of the model.

6.1. Methods

6.1.1. Participants

We recruited 23 3.5-year-olds (M age = 42.7 mo; 9 males and 14 females) and 18 4.5-year-olds (M age = 54.3 mo; 13 males and 5 females) from Knoxville, TN and surrounding communities. Research protocols were approved by the University of Tennessee, Knoxville, IRB. Informed consent was obtained from parents or legal guardians. Children were given the TC and the DCCS tasks in counterbalanced order. Half of the children had color as the pre-switch dimension in the DCCS task while the other half had shape as the pre-switch dimension. From this sample, one 3-year-old was dropped for not completing both tasks, two 3-year-olds were dropped for failing to sort correctly during the pre-switch phase, and two 3-year-olds were dropped for excessively long reaction-times during the DCCS (> 40% trials with RTs longer than 8 s). The final sample included in the analyses included 18 3-year-olds and 18 4-year-olds.

6.1.2. DCCS procedure

Children were first oriented to the DCCS task with a set of trials using physical cards and sorting trays. Children were instructed to sort by the dimension that would be relevant during the pre-switch phase. Target cards were composed of a blue circle and a red star. Test cards were composed of a red circle and a blue star. During this orienting phase, the experimenter explained the rules of the game by saying, "This is a sorting game. It is called the (color/shape) game. In this game, we are going to sort by (color/shape). That means all of the (red ones/circles) go here and (blue ones/stars) go there." Two cards were sorted by the experimenter to show children how to sort the cards. Children were then given 5 trials to perform on their own. These practice trials were not included in the analyses.

The formal DCCS task was administered on a 40″ LCD monitor that was connected to a PC running E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA). Children were given 5 pre-switch trials and 5 post-switch trials. The number of trials that are used in the DCCS can vary widely, from as many as 10 to as few as 3 during each sorting phase. Given the robust all-or-none aspect of the majority of children's behavior, in order to have as expeditions of a session as possible, and to be consistent with the NIH Toolbox version of the DCCS (Zelazo et al., 2013), we opted to use 5 trials during the pre- and post-switch phases (Buss & Spencer, 2014; Zelazo et al., 2003). During these sorting phases, test card images were composed of a yellow house and a purple fish. Target card images were composed of a yellow fish and a purple house. The mixed block used red bunny and green chair images as test cards. The target cards were composed of a green bunny and a red chair. The mixed block included 10 trials for the pre-switch dimension and 20 trials of the post-switch dimension which were administered in random order.

Fig. 11A shows the sequence of events over the course of a trial in the computerized task. The task was initiated with the presentation of images showing sorting trays and target cards (Fig. 11A, top panel). These images were 200 × 200 pixels and appeared on the screen approximately 5 × 5 cm. During this screen, the experimenter explained the rules of the game that was to be played as described above. When the child was ready, the experimenter pressed the space bar to trigger and auditory prompt saying, "Let's play the (color/shape) game!" and the presentation of a test card that was centered above the sorting tray images (Fig. 11A, middle panel). The child indicated their response by pointing to one of the sorting trays. The experimenter entered the response with a keyboard. Each session was also video recorded to later validate the keyboard responses entered by the experimenter. After the pre-switch trials were completed, the experimenter then instructed the child to sort by the other dimension by saying, "We are all done with that game. Now we are going to play a new game. This new game is called the (shape/color) game. In this game we are going to sort by (shape/color). That means all of the (fish/yellow ones) go here and (houses/purple ones) go there." Post-switch trials proceeded in a similar fashion as the pre-switch trials.

Mixed block trials were administered similarly to the pre- and post-

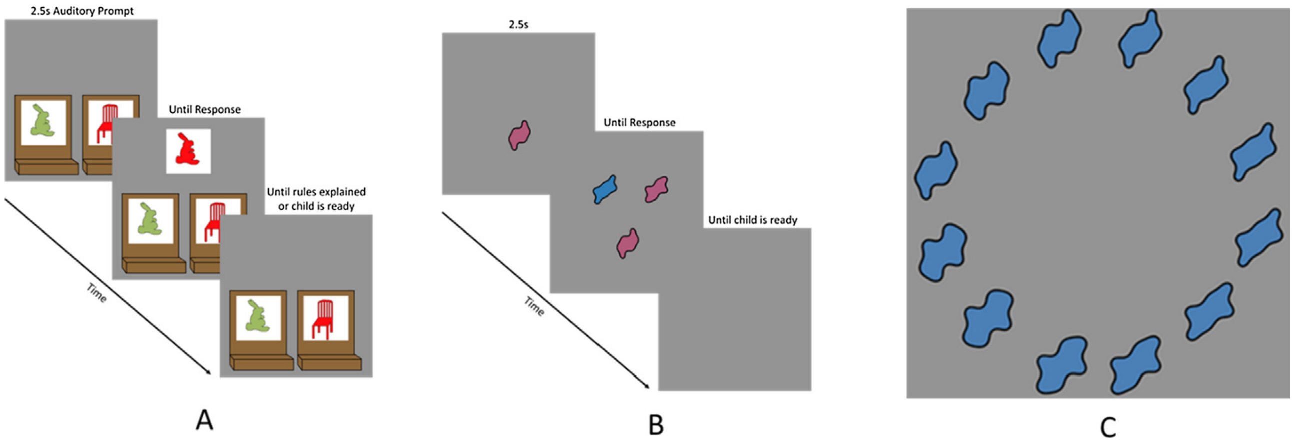

Fig. 11. (A) Sequence of events in the DCCS task. During presentation of the first panel, an auditory track played which said, "Let's play the [shape/color] game!". The trays stayed on the screen throughout the experiment, but the target cards changed depending on the condition. (B) Sequence of events in the triad classification task. (C) Depiction of shape-space used in the triad classification task.

switch trials. At the start of the mixed block, the experimenter instructed the child that they were going to sometimes play the color game and sometimes play the shape game. The experimenter initially instructed the child for both rules by saying, "If we are playing the color game, then you should sort by color. That means that the red ones go here but the green ones go there. If we are playing the shape game, though, then you should sort by shape. That means that the bunnies go here but the chairs go there." Trials were administered in the same fashion as during the previous phases.

6.1.3. Triad classification procedure

In the TC task, children were told that they were going to determine which objects "go together" and are "most similar" to one another. Children were first shown the reference object at the bottom of the screen and were directed to attend to the object by the experimenter (Fig. 11B, top panel). After 2.5 s, the H and ID choice objects were presented above the reference object to the left and right (Fig. 11B, middle panel; left/right placement of the objects was randomly selected). Children were instructed to "Pick which object is most similar to this object [pointing to the reference object]." Children responded by pointing to one of the choice objects. The experimenter then entered the response with a keyboard, and these responses were later validated using the video recordings of each session. Children were not provided any feedback and each trial was initiated by the experimenter once the child was ready and attending to the screen. Half of the trials had a shape identity match and half of the trials had a color identity match. The dimension of the identity match was randomly selected from trial to trial.

The stimuli used in the TC task were constructed with metrically controlled shapes and colors. These images were also 200 × 200 pixels and took up approximately 5 cm × 5 cm on the screen. Colors were sampled from CIE Lab (1976) color space and shapes were constructed using Fourier components as described by Drucker and Aguirre (2009). Fig. 11C shows an example of shapes used in the task. Shapes and colors were selected from a list of 60 items that were separated by 6 degrees in shape or color space. The ID object was chosen to have the exact color or shape as the reference object while the other dimension was chosen to be 180 degrees different from the reference object. The H object's features were chosen based on pilot data from adults (see Supplemental Materials). Specifically, the H choice object could vary between 90 degrees and 114 degrees in color and shape space. These values were found to be the closest separation of features at which adults achieved above 85% ID choices.

6.1.4. Results

First, we categorized children based on their post-switch performance in the DCCS task. Children were scored as switchers if they sorted at least 4 out of 5 correctly in the post-switch phase and as perseverators if they sorted 2 or fewer trials correctly (see Fig. 12, left). A Pearson χ^2 test revealed that significantly more 4-year-olds were able to switch compared to 3-year-olds ($\chi^2(1) = 5.46p = .019$). Performance on the DCCS task was also scored based on the number of correctly sorted trials throughout all sorting blocks of the task (see Fig. 12, right). An independent-samples t-test showed that 4-year-olds sorted significantly more cards correctly than 3-year-olds ($t(34) = 3.04$, $p = .004$).

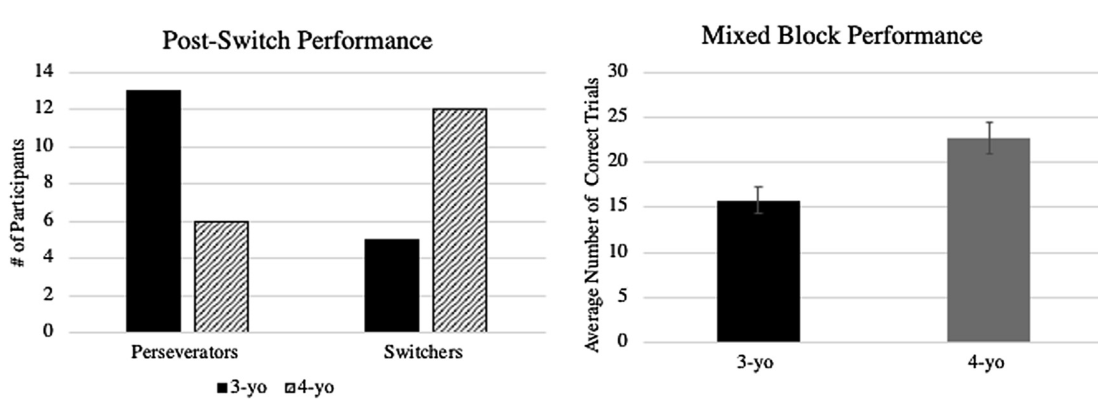

Fig. 12. Performance on the DCCS task. Left: Number of children who successfully switched and who perseverated during the post-switch phase of the DCCS by age. Right: Average number of correct trials during mixed block by age.

Performance in the TC task was scored based on the number of trials during which participants selected the ID choice object. In addition to testing for the effect of age on TC performance, we included a dimension factor based on which dimension the ID object matched the reference object since this was a within-subjects factor for this task. We first conducted a 2 (age: 3yo, 4yo) × 2 (dimension: color, shape) mixed ANOVA on the TC scores, where age was a between-subjects factor and dimension was a within-subjects factor. This analysis yielded a significant main effect of age (F (1,34) = 4.30, p = .046, η_p^2 = 0.112), no significant main effect of dimension (F (1,35) = 2.187, p = .148), nor an interaction between age and dimension (F (1,34) = 0.551, p = .463). Thus, age influenced children's performance in the TC task as it did in the DCCS task, but performance was not influenced by the dimension along which the identity object matched the reference object.

We next examined whether performance on the DCCS task is related to performance on the TC task. TC task performance was analyzed using a 2 (DCCS: switch, perseverate) × 2 (dimension: color, shape) mixed ANOVA to determine whether children performed differently in the TC task based on their performance in the post-switch phase of the DCCS task. This analysis again revealed no significant main effect of dimension (F (1,34) = 2.17, p = .150) nor an interaction between dimension and DCCS performance (F (1,34) = 0.016, p = .901). However, we found a significant main effect of DCCS performance (F (1,34) = 5.158, p = 0.030, η_p^2 = 0.132). As shown in Fig. 13A, children who successfully switched during the post-switch phase of the DCCS selected the ID choice object more frequently in the TC task, supporting the main prediction of the DNF model.

To examine switch-related performance in the TC task, we divided trials based on the number of trial repetitions before the dimension switched. We examined trials which were preceded by a single correctly-performed trial of the other dimension (e.g., color – shape) and trials that were preceded by two correctly performed trials of the other dimension (e.g., color – color – shape) before a dimension change (See Fig. 13B). We conducted a 2 (repetitions before switch: 1, 2) × 2 (DCCS: switch, perseverate) mixed ANOVA which revealed no main effect of repetition (F (1,32) = 1.219, p = .278). Again, we found a significant main effect of DCCS performance (F (1,32) = 6.780, p = .014, η_p^2 = 0.175) with children who successfully switched during the post-switch phase of the DCCS task selecting the ID choice object at a significantly higher rate. Additionally, a significant interaction between repetition and DCCS performance was detected (F (1,32) = 4.34, p = .045, η_p^2 = 0.120). Pairwise comparisons revealed that switchers and perseverators performed equally well when only one repetition preceded a dimension switch (t (34) = 1.30, p = 0.200). However, switchers performed significantly better that perseverators when two repetitions preceded a dimension switch (t (34) = 2.88, p = 0.007).

Lastly, to investigate whether overall performance in the DCCS task significantly predicted performance in the TC task, we conducted regressions using linear, quadratic, and cubic functions with total score in the DCCS (including the mixed block) as the independent variable and the number of trials in which children selected the ID object in the TC task as the dependent variable. In line with the model's prediction, a linear relationship was significant (F(1,34) = 10.981, p = .002). Further, a linear relationship fit the data as well as the quadratic (F (2,33) = 7.940, p = .002) and cubic (F(2,33) = 7.849, p = .002) functions (see Fig. 13C).

7. General discussion

In this report, we demonstrated how a neural process model previously used to explain the development of flexible dimensional attention in the DCCS task (Buss & Spencer, 2014) also explains patterns of behavior of 3- to 5-year-olds across two other tasks that require selective and stable dimensional attention. Flexibility, selectivity, and stability were revealed as properties of neurocomputational dynamics

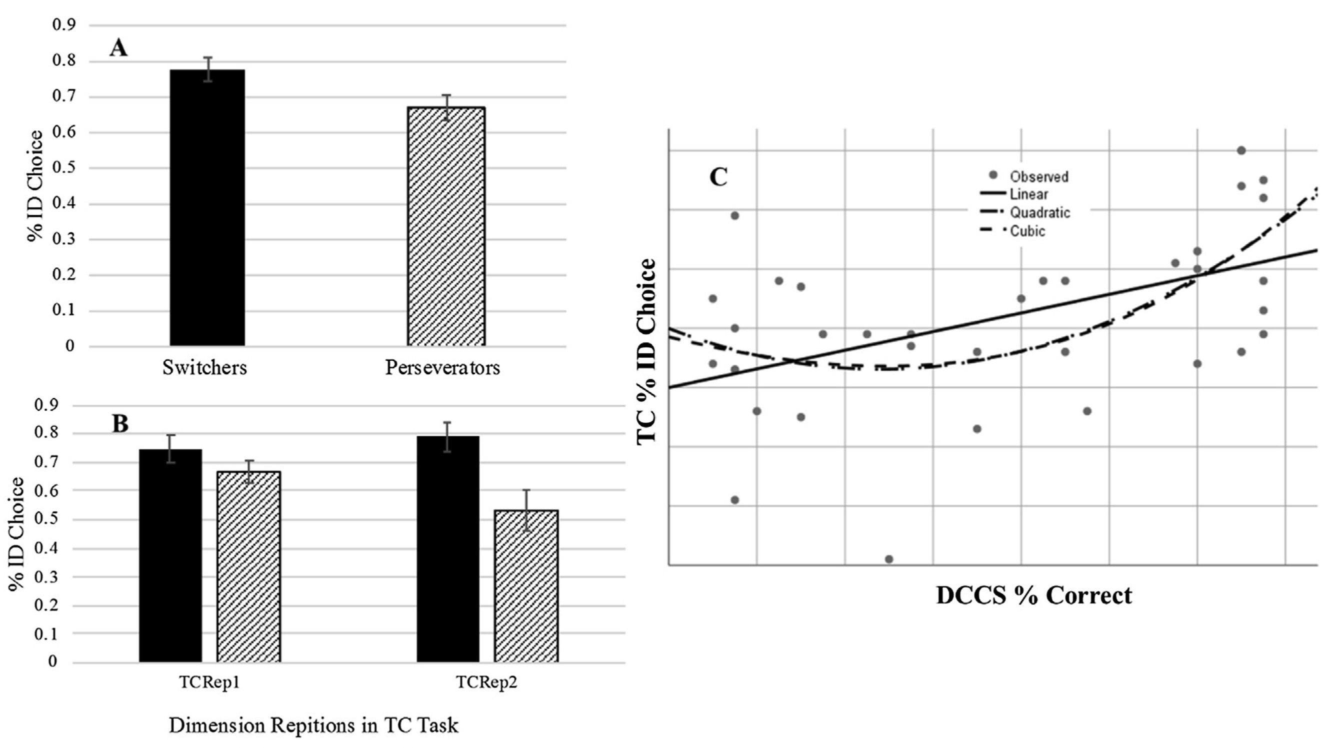

Fig. 13. Performance on the triad classification (TC) task as a function of DCCS performance. (A) Average percent of trials children selected the identity (ID) choice object, separated by their DCCS performance. (B) Average percent of trials children selected the ID choice object as a function of preceding trials before a dimension switch for children who switched and perseverated on the DCCS task (Rep-1: one preceding trial of the other dimension; Rep-2: two preceding trials of the other dimension) (C) Percent of ID choices plotted against the average number of trials sorted correctly in all phases of the DCCS task, along with the regression line.

in the context of different tasks. First, we demonstrated that the model can explain previously reported associations in performance on the DCCS task and the DP task. Specifically, models that showed flexible dimensional attention in the DCCS task (switchers) also had better stable dimensional attention in the DP task. Further, the model generated novel predictions regarding an association between DCCS and TC task performance. Data from 3- and 4-year-olds supported these predictions: children who switched rules in the DCCS task showed better selective dimensional attention in the TC task compared to children who perseverated in the DCCS task.

The DNF model performed these tasks like 3- and 4-year-olds without any modifications to the model architecture or parameters to accommodate performance on the other tasks. By simply presenting inputs to this model that reflect the structure of these tasks, the dynamics of the model gave rise to these cognitive functions. Performance on these tasks emerge based on (1) the processes of object representation, (2) the demands imposed by a particular task, and (3) the developmental status of the dimensional label system. Specifically, the properties of the object representation system account for how the configuration of features influences the neurocognitive system. The primary function of this system is to bind features to spatial locations. Thus, this system is sensitive to the unique aspects of each task, such as the relationship between the features of different objects and the history of feature-space bindings over a series of trials. The developmental status of the model is defined as the strength of associations between labels and features. In this way, the dimensional label system differentially modulates processing of visual features based on the strength of the links between labels and visual features over development.

Modulation of visual features can be achieved either explicitly by instructing the model to attend to a specific dimension as in the DCCS task, or implicitly based on the signals propagating from the object representation system as in the TC and DP tasks. In the DCCS task, the model is given explicit instructions to sort by a particular dimension so that there is no ambiguity regarding which dimensional label should be activated. With weak associations between labels and features, the model predominantly perseverates due to conflict between the rules given for the post-switch phase (e.g., sort red to the right and blue to left) and the pattern of memories that accumulate during the pre-switch phase (e.g., red was previously sorted to the left and blue was previously sorted to the right). With strong associations between labels and features, however, the model can resolve this conflict and switches rules.

In implicit tasks, such as in the DP task, there is greater ambiguity regarding which dimension is relevant. If color is the matching dimension during the priming trials, the overlap of the choice and reference objects' color values sends strong activation from the object representation system to the "color" label. With strong associations between labels and features (as in older children or adults), this activation engages the label system more strongly which helps the model to build memory traces on the relevant dimensional label. On subsequent test trials, this memory trace can provide a bias toward the primed color dimension and can maintain attention to that dimension. However, if these label-feature links are weak (as in younger children), the model engages the frontal label system weakly and builds weaker memory traces on the relevant dimensional label, leading to a weaker bias for the primed dimension. Thus, the functions of dimensional attention the model ultimately achieves is inherent to the demands imposed by the task. Mechanistically, the label system operates at the level of labels and visual features, whereas the object system operates at the level of spatial locations and visual features; however, by coupling these systems together and allowing to interact reciprocally instead of hierarchically, these different combinations of representation can give rise to different functions of dimensional attention in the context of specific task-demands. Stronger links between labels and features give rise to flexibility, selectivity, and stability by creating the ability to activate dimensional labels which can more strongly intervene in the object representation processes that bind features to spatial locations.

The simulation results from the DNF model point to an exciting possibility that a dimensional label learning mechanism drives the development of EF. Specifically, we propose that learning labels for visual features (e.g., red and blue) and dimensions (e.g., color and shape) can change how attention is allocated to features of objects and the EF skills that children can display. Previous accounts of EF development have mainly focused on biological or maturational changes in frontal cortex as a primary cause of EF development (Bunge & Zelazo, 2006; Moriguchi & Hiraki, 2009; Morton & Munakata, 2002). Although it has not yet been explored whether dimensional label learning is predictive of dimensional attention development, previous research has demonstrated a powerful influence of labels on dimensional attention. For example, the labels used during instruction influence DCCS performance (Yerys & Munakata, 2006) and categorization reversal tasks (Schonberg, Atagi, & Sandhofer, 2018), and using labels for objects has been shown to enhance children's selective dimensional attention in a free classification task (Perry & Samuelson, 2013). Moreover, Plunkett and colleagues have demonstrated that engaging labels increases selective attention to visual dimensions (Althaus & Plunkett, 2016; Mather & Plunkett, 2010), and labels can disrupt perceptual category learning, leading to broader perceptual categories if the same label is applied to different perceptual properties (Plunkett, Hu, & Cohen, 2008). Thus, future research will need to explore the role of dimensional label learning on different measures of dimensional attention and weather facilitating dimensional label learning also facilitates dimensional attention.

The model formalizes the mechanisms involved in object representation and specifies how strengthening feature-label associations can give rise to a diverse set of dimensional attention skills. Many researchers have suggested that language and cognitive control are related over development (Cragg & Nation, 2010; Deak, 2003; Jacques & Zelazo, 2005). In these proposals, the logic that is supported by language is typically the source of EF or cognitive flexibility. However, the proposal here is unique because we suggest that forming associations between labels and visual features is the causal force that drives the formation of cognitive control networks. By embodying the label learning process in this way, our account can explain associations between explicit and implicit dimensional attention. That is, the label learning process is not only an abstract learning process that exerts top-down control over cognitive functioning, but can be flexibly engaged in both a bottom-up and top-down fashion based on the visual properties of the task.

Current research in our lab is also exploring the changes in neural dynamics associated with comprehension and production of dimensional labels and associated changes in dimensional attention skills. Motivated by the top-down view of cognitive control, previous theories and research have focused on local changes to prefrontal cortex activity as the primary driver of developmental changes in EF (Bunge & Zelazo, 2006; Moriguchi & Hiraki, 2009; Morton & Munakata, 2002). However, more recent data implicates changes in activation across parietal and temporal cortex in addition to frontal cortex (Buss & Spencer, 2018; Morton, Bosma, & Ansari, 2009). The DNF model explains these results through stronger coupling between the label system and the object representation system (Buss & Spencer, 2018). Thus, it is possible that changes in EF during early childhood are driven by the formation of neural networks that integrate frontal regions involved in label learning and posterior regions involved in object representation.

There are many important limitations to this study that can be addressed by future work. For example, previous research has revealed important dynamics regarding how behavior is linked across the timescales of in-the-moment performance, learning across trials, and developmental changes. van Bers, Visser, van Schijndel, Mandell, and Raijmakers (2011) used a hidden Markov model approach to examine the latent states underlying development as a function of the trial-to-trial response patterns of children. These analyses demonstrated that children's behavior is best explained by two qualitatively different latent states, one corresponding to switching and another corresponding to perseverating rather than a latent state that changes in the likelihood of switching. Moreover, children can switch between developmental states of perseverating and switching in-the-moment during a series of post-switch trials. Notably, switching between states typically occurs from a perseveration state to a switching state, but not *vice versa*. The DNF model has not yet been probed regarding whether such latent states also underlie performance of the model, whether transitions between states can also arise as a function of the neural dynamics in the model, or how underlying latent state dynamics are related to the parameter distributions of the model. Moreover, even though both the 'young' and the 'old' model produced bi-modal distributions of performance in the DCCS task, it is also unclear how more continuous shifts in the distribution of parameters would influence the rates of performance and the bimodal nature of flexibility.

Relatedly, initial research using the TC task proposed a holistic to analytic shift such that children initially process the features of object holistically and later analytically attend to only a single dimension (Smith & Kemler, 1977). However, Raijmakers, Jansen, and van der Maas (2004) used a latent class analysis method to demonstrate that children's performance on the TC task is best explained by underlying cognitive states that correspond to attending to shape or color in a rule-like fashion rather than states that correspond to holistic or analytic processing. In many ways, the model's dynamics are consistent with either of these explanations. It is true that the model shifts in how much processing priority is placed on a particular dimension as a function of dimensional label learning. With weaker priority given to any particular dimension the model can be said to be integrating information across dimensions. However, it is also true that the model does 'select' a dimension to which it selectively attends on each trial which is the basis for rule-like behavior when the model performs the DCCS task. This is based on the competitive dynamics between the 'shape' and 'color' dimensional label units. However, it is currently unclear whether the data from our TC task that used shape and color as relevant dimensions replicated the latent class analyses from Raijmakers et al. (2004) and how the developmental states of the model shift within sets of parameters and as a function of changes in parameters.

Another important limitation of this work is that only a simple triad of stimuli were involved across all the tasks simulated here. Thus, it is unclear how the model would operate in task conditions that required more complex consideration of spatial locations or features when more items are relevant (e.g., Enns & Cameron, 1987; Hommel, Li, & Li, 2004; Jennings, Dagenbach, Engle, & Funke, 2007; Rueda et al., 2005; Trick & Enns, 1998). In this regard, we are also currently developing the model to simulate comprehension, production, and matchings tasks presented by Sandhofer and Smith (1999) as a means of assessing how the label associations formed by the model impact the ability of the model to perform these tasks. Importantly, the comprehension and matching tasks from this literature provide more complex configurations of stimuli than those presented here and will test of the ability of the model to generalize to tasks that involve processing objects in more complex contexts and with stronger spatial competition.

In conclusion, we demonstrated that neural systems approach can provide a framework to explain cognitive function and changes in cognitive function over development. The work presented here addresses various debates regarding the mechanisms underlying changes in EF and attention during early childhood. We propose a unified theory of development which is grounded in a dimensional attention mechanism. This dimensional attention mechanism arises from dimensional label and object representation systems. Further, this mechanism can achieve diverse attentional functions either explicitly or implicitly depending on the task. Importantly, we emphasize the strengths and advantages of a process-based approach to study development. This perspective goes beyond focusing on individual mechanisms of EF. Instead, it explains how these mechanisms arise from real-time processes and provides a way of linking performance on tasks that tap into distinct aspects of cognition.

Acknowledgements

Research was supported by R01HD092484 awarded to ATB. Behavioral and simulation data can be found at https://osf.io/j6knx/.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2019.06.015.

References

Agostino, A., Johnson, J., & Pascual-Leone, J. (2010). Executive functions underlying multiplicative reasoning: Problem type matters. *Journal of Experimental Child Psychology, 105*(4), 286–305. https://doi.org/10.1016/J.JECP.2009.09.006.

Althaus, N., & Plunkett, K. (2016). Categorization in infancy: Labeling induces a persisting focus on commonalities. *Developmental Science, 19*(5), 770–780. https://doi.org/10.1111/desc.12358.

Amso, D., & Scerif, G. (2015). The attentive brain: Insights from developmental cognitive neuroscience. *Nature Reviews Neuroscience, 16*(10), 606–619. https://doi.org/10.1038/nrn4025.

Benitez, V. L., Vales, C., Hanania, R., & Smith, L. B. (2017). Sustained selective attention predicts flexible switching in preschoolers. *Journal of Experimental Child Psychology, 156*, 29–42. https://doi.org/10.1016/J.JECP.2016.11.004.

Blair, C., Zelazo, P. D., & Greenberg, M. T. (2005). The measurement of executive function in early childhood. *Developmental Neuropsychology, 28*(2), 561–571.

Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology, 33*(3), 205–228. https://doi.org/10.1080/87565640801982312.

Bunge, S. A., & Zelazo, P. D. (2006). A brain based account of the development of rule use in childhood. *Psychological Science, 15*(3), 118–121.

Buss, A. T., & Spencer, J. P. (2014). The emergent executive: A dynamic field theory of the development of executive function. *Monographs of the Society for Research in Child Development, 79*(2), https://doi.org/10.1002/mono.12096.

Buss, A. T., & Spencer, J. P. (2018). Changes in frontal and posterior cortical activity underlie the early emergence of executive function. *Developmental Science*. https://doi.org/10.1111/desc.12602.

Chatham, C. H., Yerys, B. E., & Munakata, Y. (2012). Why won't you do what I want? The informative failures of children and models. *Cognitive Development, 27*(4), 349–366. https://doi.org/10.1016/j.cogdev.2012.07.003.

Colombo, J., & Cheatham, C. L. (2006). The emergence and basis of endogenous attention in infancy and early childhood. *Advances in Child Development and Behavior, 34*, 283–322. https://doi.org/10.1016/S0065-2407(06)80010-8.

Cragg, L., & Nation, K. (2010). Language and the development of cognitive control. *Topics in Cognitive Science, 2*(4), 631–642. https://doi.org/10.1111/j.1756-8765.2009.01080.x.

Cuevas, K., & Bell, M. A. (2014). Infant attention and early childhood executive function. *Child Development, 85*(2), 397–404. https://doi.org/10.1111/cdev.12126.

Deak, G. O. (2003). The development of cognitive flexibility and language abilities. *Advances in Child Development and Behavior, 31*, 271–327.

Diamond, A., Carlson, S. M., & Beck, D. M. (2005). Preschool children's performance in task switching on the Dimensional Change Card Sort task: Separating the dimensions aids the ability to switch. *Developmental Neuropsychology, 28*(2), 689–729.

Diamond, A., & Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science, 333*(6045), 959–964.

Drucker, D. M., & Aguirre, G. K. (2009). Different spatial scales of shape similarity representation in lateral and ventral LOC. *Cerebral Cortex, 19*(10), 2269–2280.

Enns, J. T., & Cameron, S. (1987). Selective attention in young children: The relations between visual search, filtering, and priming. Retrieved from *Journal of Experimental Child Psychology, 44*(1), 38–63. http://www.ncbi.nlm.nih.gov/pubmed/3612023.

Fisher, A. V. (2011). Automatic shifts of attention in the Dimensional Change Card Sort task: Subtle changes in task materials lead to flexible switching. *Journal of Experimental Child Psychology, 108*(1), 211–219. https://doi.org/10.1016/j.jecp.2010.07.001.

Fuhs, M. W., & Day, J. D. (2011). Verbal ability and executive functioning development in preschoolers at head start. *Developmental Psychology, 47*(2), 404–416. https://doi.org/10.1037/a0021065.

Fuster, J. M. (2000). Executive frontal functions. *Experimental Brain Research, 133*(1), 66–70. https://doi.org/10.1007/s002210000401.

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin, 134*(1), 31–60.

Gliozzi, V., Mayor, J., Hu, J.-F., & Plunkett, K. (2009). Labels as features (Not Names) for infant categorization: A neurocomputational approach. *Cognitive Science, 33*(4), 709–738. https://doi.org/10.1111/j.1551-6709.2009.01026.x.

Hanania, R., & Smith, L. B. (2010). Selective attention and attention switching: Towards a unified developmental approach. *Developmental Science, 13*(4), 622–635. https://doi.org/10.1111/j.1467-7687.2009.00921.x.

Happaney, K., & Zelazo, P. D. (2003). Inhibition as a problem in the psychology of behavior. *Current,* 468–470.

Hommel, B., Li, K. Z. H., & Li, S. (2004). Visual search across the life span. Retrieved from *Developmental Psychology, 40*(4), 545–558. https://insights.ovid.com/developmental-psychology/depsy/2004/07/000/visual-search-across-life-span/7/00063061.

Huizinga, M., Dolan, C. V., & van der Molen, M. W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia, 44*(11), 2017–2036.

Jacques, S., & Zelazo, P. D. (2005). Language and the development of cognitive flexibility: Implications for theory of mind. In J. Astington, & J. Baird (Eds.). *Why language matters for theory of mind* (pp. 144–162). Oxford, UK: Oxford University Press.

Jennings, J. M., Dagenbach, D., Engle, C. M., & Funke, L. J. (2007). Age-related changes and the attention network task: An examination of alerting, orienting, and executive function. *Aging, Neuropsychology, and Cognition, 14*(4), 353–369. https://doi.org/10.1080/13825580600788837.

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), 449–467. https://doi.org/10.1111/1467-7687.00300.

Kirkham, N. Z., & Diamond, A. (2003). Sorting between theories of perseveration: performance in conflict tasks requires memory, attention and inhibition, pp. 474–476.

Kloo, D., & Perner, J. (2005). Disentangling dimensions in the dimensional change card-sorting task. *Developmental Science, 8*(1), 44–56.

Lee, K., Bull, R., & Ho, R. M. H. (2013). Developmental changes in executive functioning. *Child Development, 84*(6), 1933–1953. https://doi.org/10.1111/cdev.12096.

Lee, K., Ng, S. F., Bull, R., Pe, M. L., & Ho, R. H. M. (2011). Are patterns important? An investigation of the relationships between proficiencies in patterns, computation, executive functioning, and algebraic word problems. *Journal of Educational Psychology, 103*(2), 269–281. https://doi.org/10.1037/a0023068.

Lehto, J. E., Juujarvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology, 21*, 59–80.

Mather, E., & Plunkett, K. (2010). Novel labels support 10-month-olds' attention to novel objects. *Journal of Experimental Child Psychology, 105*(3), 232–242. https://doi.org/10.1016/J.JECP.2009.11.004.

McAuley, T., & White, D. A. (2011). A latent variables examination of processing speed, response inhibition, and working memory during typical development. *Journal of Experimental Child Psychology, 108*(3), 453–468. https://doi.org/10.1016/j.jecp.2010.08.009.

Medin, D. L. (1973). Measuring and training dimensional preferences. *Child Development, 44*(2), 359. https://doi.org/10.2307/1128060.

Miller, M. R., Giesbrecht, G. F., Müller, U., McInerney, R. J., & Kerns, K. A. (2012). A latent variable approach to determining the structure of executive function in preschool children. *Journal of Cognition and Development, 13*(3), 395–423. https://doi.org/10.1080/15248372.2011.585478.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis. *Cognitive Psychology, 41*(1), 49–100. https://doi.org/10.1006/COGP.1999.0734.

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., ... Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences of the United States of America, 108*(7), 2693–2698.

Moriguchi, Y., & Hiraki, K. (2009). Neural origin of cognitive shifting in young children. *Proceedings of the National Academy of Sciences of the United States of America, 106*(14), 6017–6021.

Morton, J. B., Bosma, R., & Ansari, D. (2009). Age-related changes in brain activation associated with dimensional shifts of attention: An fMRI study. *NeuroImage, 46*(1), 249–256.

Morton, J. B., & Munakata, Y. (2002). Active versus latent representations: A neural network model of perseveration, dissociation, and decalage. *Developmental Psychobiology, 40*(3), 255–265.

Munakata, Y., Morton, J. B., & Yerys, B. E. (2003). Children's perseveration: Attentional inertia and alternative accounts. *Developmental Science, 6*(5), 471–473. https://doi.org/10.1111/1467-7687.00302.

Perone, S., Molitor, S. J., Buss, A. T., Spencer, J. P., & Samuelson, L. K. (2015). Enhancing the executive functions of 3-year-olds in the dimensional change card sort task. *Child Development, 86*(3), https://doi.org/10.1111/cdev.12330.

Perone, S., Plebanek, D. J., Lorenz, M. G., Spencer, J. P., & Samuelson, L. K. (2017). Empirical tests of a brain-based model of executive function development. *Child Development*. https://doi.org/10.1111/cdev.12885.

Perry, L. K., & Samuelson, L. K. (2013). The role of verbal labels in attention to dimensional similarity. *Proceedings of the Cognitive Science Society, 35*(35).

Plunkett, K., Hu, J.-F., & Cohen, L. B. (2008). Labels can override perceptual categories in early infancy. *Cognition, 106*(2), 665–681. https://doi.org/10.1016/J.COGNITION.2007.04.003.

Raijmakers, M. E., Jansen, B. R., & van der Maas, H. L. J. (2004). Rules and development in triad classification task performance. *Developmental Review, 24*(3), 289–321. https://doi.org/10.1016/j.dr.2004.06.002.

Reck, S. G., & Hund, A. M. (2011). Sustained attention and age predict inhibitory control during early childhood. *Journal of Experimental Child Psychology, 108*(3), 504–512. https://doi.org/10.1016/j.jecp.2010.07.010.

Rennie, D. A. C., Bull, R., & Diamond, A. (2004). Executive functioning in preschoolers: Reducing the inhibitory demands of the dimensional change card sort task. *Developmental Neuropsychology, 26*(1), 423–443.

Richards, J. E., Reynolds, G. D., & Courage, M. L. (2010). The neural bases of infant attention. *Current Directions in Psychological Science, 19*(1), 41–46. https://doi.org/10.1177/0963721409360003.

Rose, S. A., Feldman, J. F., & Jankowski, J. J. (2011). Modeling a cascade of effects: The role of speed and executive functioning in preterm/full-term differences in academic achievement. *Developmental Science, 14*(5), 1161–1175. https://doi.org/10.1111/j.1467-7687.2011.01068.x.

Ross-Sheehy, S., Oakes, L. M., & Luck, S. J. (2011). Exogenous attention influences visual short-term memory in infants. *Developmental Science, 14*(3), 490–501.

Rueda, M. R., Posner, M. I., & Rothbart, M. K. (2005). The development of executive attention: Contributions to the emergence of self-regulation. *Developmental Neuropsychology, 28*(2), 573–594. https://doi.org/10.1207/s15326942dn2802_2.

Sandhofer, C. M., & Smith, L. B. (1999). Learning color words involves learning a system of mappings. Retrieved from *Developmental Psychology, 35*(3), 668–679. http://www.ncbi.nlm.nih.gov/pubmed/10380858.

Schonberg, C. C., Atagi, N., & Sandhofer, C. M. (2018). Two-year-olds' executive functioning: The influence of task-specific vocabulary knowledge. *Infant Behavior and Development, 53*, 33–42. https://doi.org/10.1016/J.INFBEH.2018.09.004.

Schoner, G., & Spencer, J. P. (2016). *How we think: An introduction to Dynamic Field Theory.* New York: Oxford University Press.

Schutte, A. R., Keiser, B. A., & Beattie, H. L. (2017). Developmental differences in the influence of distractors on maintenance in spatial working memory. *Journal of Cognition and Development, 18*(3), 338–357. https://doi.org/10.1080/15248372.2017.1300155.

Sheese, B. E., Rothbart, M. K., Posner, M. I., White, L. K., & Fraundorf, S. H. (2008). Executive attention and self-regulation in infancy. *Infant Behavior and Development, 31*(3), 501–510. https://doi.org/10.1016/J.INFBEH.2008.02.001.

Shing, Y. L., Lindenberger, U., Diamond, A., Li, S.-C., & Davidson, M. C. (2010). Memory maintenance and inhibitory control differentiate from early childhood to adolescence. *Developmental Neuropsychology, 35*(6), 679–697. https://doi.org/10.1080/87565641.2010.508546.

Smith, L. B., & Kemler, D. G. (1977). Developmental trends in free classification: Evidence for a new conceptualization of perceptual development. *Journal of Experimental Child Psychology, 24*(2), 279–298. https://doi.org/10.1016/0022-0965(77)90007-8.

St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *Quarterly Journal of Experimental Psychology, 59*(4), 745–759. https://doi.org/10.1080/17470210500162854.

Swanson, H. L., Jerman, O., & Zheng, X. (2008). Growth in working memory and mathematical problem solving in children at risk and not at risk for serious math difficulties. *Journal of Educational Psychology, 100*(2), 343–379. https://doi.org/10.1037/0022-0663.100.2.343.

Trick, L. M., & Enns, J. T. (1998). Lifespan changes in attention: The visual search task. *Cognitive Development, 13*(3), 369–386. https://doi.org/10.1016/S0885-2014(98)90016-8.

van Bers, B. M. C. W., Visser, I., van Schijndel, T. J. P., Mandell, D. J., & Raijmakers, M. E. J. (2011). The dynamics of development on the Dimensional Change Card Sorting task. *Developmental Science, 14*(5), 960–971. https://doi.org/10.1111/j.1467-7687.2011.01045.x.

van der Sluis, S., de Jong, P. F., & van der Leij, A. (2007). Executive functioning in children, and its relations with reasoning, reading, and arithmetic. *Intelligence, 35*(5), 427–449. https://doi.org/10.1016/J.INTELL.2006.09.001.

Van der Ven, S. H. G., Kroesbergen, E. H., Boom, J., & Leseman, P. P. M. (2012). The development of executive functions and early mathematics: A dynamic relationship. *British Journal of Educational Psychology, 82*(1), 100–119. https://doi.org/10.1111/j.2044-8279.2011.02035.x.

Wiebe, S. A., Espy, K. A., & Charak, D. (2008). Using confirmatory factor analysis to understand executive control in preschool children: I. Latent structure. *Developmental Psychology, 44*(2), 575–587. https://doi.org/10.1037/0012-1649.44.2.575.

Wiebe, S. A., Sheffield, T., Nelson, J. M., Clark, C. A. C., Chevalier, N., & Espy, K. A. (2011). The structure of executive function in 3-year-olds. *Journal of Experimental Child Psychology, 108*(3), 436–452.

Willoughby, M. T., Blair, C. B., Wirth, R. J., & Greenberg, M. (2010). The measurement of executive function at age 3 years: Psychometric properties and criterion validity of a new battery of tasks. *Psychological Assessment, 22*(2), 306–317. https://doi.org/10.1037/a0018708.

Willoughby, M. T., Wirth, R. J., Blair, C. B., & Family Life Project Investigators, F. L. P. (2012). Executive function in early childhood: longitudinal measurement invariance and developmental change. Psychological Assessment, 24(2), 418–31. http://doi.org/10.1037/a0025779.

Wu, K. K., Chan, S. K., Leung, P. W. L., Liu, W.-S., Leung, F. L. T., & Ng, R. (2011). Components and developmental differences of executive functioning for school-aged children. *Developmental Neuropsychology, 36*(3), 319–337. https://doi.org/10.1080/87565641.2010.549979.

Yerys, B. E., & Munakata, Y. (2006). When labels hurt but novelty helps: Children's perseveration and flexibility in a card-sorting task. *Child Development, 77*(6), 1589–1607.

Zelazo, P. D., Anderson, J. E., Richler, J., Wallner-Allen, K., Beaumont, J. L., & Weintraub, S. (2013). NIH Toolbox cognition battery (CB): Measuring executive function and attention. *Monographs of the Society for Research in Child Development, 78*(4), 16–33.

Zelazo, P. D., & Bauer, P. J. (2013). National Institutes of Health Toolbox— Cognition Battery (NIH Toolbox CB): Validation for Children between 3 and 15 Years. *SRCD Monograph, 78*(4), 1–146.

Zelazo, P. D., Muller, U., Frye, D., & Marcovitch, S. (2003). The development of executive function in early childhood. *Monographs of the Society for Research in Child Development, 2*(68), 1–137.