NIH Public Access
Author Manuscript
Dev Sci. Author manuscript; available in PMC 2011 July 1.

Published in final edited form as:
Dev Sci. 2010 July ; 13(4): 622–635. doi:10.1111/j.1467-7687.2009.00921.x.

Selective Attention and Attention Switching: Toward a Unified Developmental Approach

Rima Hanania and **Linda B. Smith**
Indiana University

Abstract

We review and relate two literatures on the development of attention in children: one concerning flexible attention switching and the other concerning selective attention. The first is a growing literature on preschool children's performances in an attention switching task indicating that children become more flexible in their attentional control during the preschool years. The second literature encompasses a large and robust set of phenomena for the same developmental period that indicate a protracted course of development for selective attention in children. We ask whether developmental changes in processes of selective attention may contribute to more flexible attention switching. We consider the two sets of phenomena with respect to this question and propose an empirical agenda for their joint study that may lead ultimately to a unified account of the development of selective attention and attention switching.

Intelligent behavior requires selecting task-relevant information and minimizing interference from irrelevant information. Intelligent behavior also requires switching sources of information as task-relevance changes. Selective attention and attention switching are fundamental to almost all cognitive tasks, and deficits in these attentional processes have wide-ranging and cascading consequences for development (e.g., Mundy et al., 2007; Posner & Rothbart, 2007; Rueda, Posner, & Rothbart, 2005; Ruff & Rothbart, 1996). It is possible that different mechanisms and different neural processes underlie selection and switching (e.g., Fan, McCandliss, Fossella, Flombaum, & Posner, 2005). At the same time, the processes underlying these two aspects of attention may overlap and connect in important ways, leading to related developmental trends. So far, the development of selective attention and the development of flexible attention switching have been studied in separate literatures. In this paper, we present reasons for studying the development of the two together, and suggest that these two developing abilities may be fundamentally related and intertwined.

This paper is specifically concerned with tasks that ask perceivers to make judgments on one stimulus dimension (say, color) while ignoring another (say shape). Selective attention to stimulus dimensions is often conceptualized as the differential weighting of dimensions in perceptual decisions, as in similarity judgments (Shepard, 1964), in classification (Nosofsky, 1986), in attentional learning (Kruschke 2001), and in neurocomputational models of cognitive control (Rougier, Noelle, Braver, Cohen & O'Reilly, 2005; Reynolds, Braver, Brown & Van der Stigchel, 2006). Selective attention to stimulus dimensions is also critical in two well-documented developmental trends. The first comes from the current literature concerning the development of flexible attention switching in preschool children and children's performance in the Dimension Change Card Sort (DCCS) task. This literature largely ignores the issue of the development of selective attention and assumes that the problem is specifically in switching attention to a new dimension. The second literature is an

---

Correspondence: Rima Hanania (rhanania@indiana.edu), Department of Psychological and Brain Sciences, IUB, 1101 East 10th Street, Bloomington, Indiana 47405, Office: 812-855-8256.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

older one on the development of selective attention is preschool children. The well-documented trends in this old literature raise the possibility that immature selective attention skills may play a role in children's difficulty in switching attention.

The plan of the paper is as follows: First we review the current evidence on attention switching in the DCCS task, highlighting the evidence that suggests that selective attention per se may be a factor. Second, we review the older literature on the development of selective attention, a literature that strongly suggests that this is still a developing skill in the late preschool period. Third, we consider models of selective attention and of attention switching that suggest similar or related underlying processes that may be critical to the development of both selective attention and attention switching.

Flexible Attention Switching

The development of executive control and children's ability to adjust their behavior in accordance with changing task goals is a priority in the current developmental literature. The Dimensional Change Card Sort (DCCS) task (e.g., Frye, Zelazo, & Palfai, 1995), is one common experimental approach. The task, illustrated in Figure 1, works as follows: Two target cards are displayed, for example, drawings of a red truck and a blue star. Children are told to sort a stack of cards that match each target on one dimension only. In this example, the sorting cards would be blue trucks and red stars. The child might first be told to sort the cards by shape, putting the blue trucks with the red truck and the red stars with the blue star. After this so-called preswitch phase, the sorting rule is switched and the child is told to sort the cards by the other dimension, for example color, requiring now that the blue trucks be sorted with the blue star and the red stars be sorted with the red truck.

The main finding is that although younger children are able to follow instructions and sort without error on preswitch trials (no matter which rule is first), they have trouble switching sorting rules on postswitch trials and perseveratively sort by the preswitch dimension rather than switching to sort by the new dimension. Older children are more likely to succeed in sorting correctly both pre and postswitch, demonstrating an ability to follow the new instructions and to flexibly switch their responses appropriately (e.g., Brooks, Hanauer, Padowska, & Rosman, 2003; Frye et al., 1995; Kirkham, Cruess, & Diamond, 2003; Perner & Lang, 2002; Towse, Redbond, Houston-Price, & Cook, 2000; Zelazo, Frye, & Rapus, 1996). This developmental trend is robust, with the same pattern of behavior emerging for different pairs of dimensions (Frye et al., 1995), and for drawings of either abstract geometric shapes (Frye et al., 1995) or familiar objects (e.g., Zelazo et al., 1996). This pattern also persists whether the children are sorting the cards themselves or whether they are observing the cards being sorted by a puppet (Jacques, Zelazo, Kirkham, & Semcesen, 1999). Moreover, the perseverative bias sets up rapidly, in as few as 1 or 2 preswtich trials (Zelazo et al., 1996). The key developmental transitions appear to occur between 3 and 5 years of age. Table 1 summarizes the main experimental results. The developmental trend in the standard DCCS experiments is dramatic--while performance on preswitch trials is close to perfect for all age groups, young preschool children do not switch as successfully to the new dimension as do older preschool children, performing at rates as low as 20% average correct postswitch sorts whereas older children perform at rates reaching 80% or higher (Frye et al., 1995).

The fact that both younger and older preschool children do well preswitch on this task suggests that the problem is switching attention not selective attention, and several theories of perseveration in the DCCS task reflect this (Kirkham et al., 2003; Towse et al., 2000; Kloo & Perner, 2003; 2005; Müller, Dick, Gela, Overton, & Zelazo, 2006; Zelazo, Müller, Frye, & Marcovitch, 2003). In these theories, the ability to selectively attend is not

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

considered a factor; switching is a separate ability which has to develop to overcome the stickiness of selective attention. There are, however, hints in the experimental evidence that the task requirement of selective attention is itself a factor in perseveration.

In one relevant study, Perner and Lang (2002) used what is sometimes called a “reversal shift” task (see also Kloo & Perner, 2003; Zelazo et al., 2003). The target cards, illustrated in panel 2 of Figure 2, were, for example, red cars and red suns and the sorting cards were identical to the targets, with the only varying dimension being shape. Children were first asked to sort sun cards with the sun target card and car cards with the car target card. The postswitch task asked them to switch by putting suns with the car target and cars with the sun target. This is not a selective attention task. The switch requires a switch in responses (which cards go where) but not a shift in attention (shape is always relevant). Young children were far more successful in this task than in a standard selective attention version of the DCCS task. Brooks et al. (2003) replicated and extended the result, showing that the need to selectively attend--even if switching attention to a new dimension was not required--contributed to young children’s perseverative responding (cf., Kloo et al, 2008). In one task, as illustrated in panel 1 of Figure 2, the stimuli were line-drawings of common objects, so only shape varied and only shape was task relevant. Selective attention was not required. Again, the postswitch task simply switched the stimulus-response rules and children sorted successfully. The second condition was the same but now color varied irrelevantly on both pre-and postswitch trials, as illustrated in panel 3 of Figure 2. Shape was always relevant and color irrelevant, so the task required selective attention but not attention switching. Children perseverated in this task unable to switch the sorting rule. The need to attend to shape while ignoring color was enough to cause increased perseverative errors.

Other DCCS studies also suggest that the relative ease and difficulty of the selective attention task influences children’s success in switching. In one series of studies, Zelazo et al. (2003; experiments 8 and 9) found that young children perseverate less when the need to selectively attend--in the sense of ignoring some of the available information--is reduced on preswitch trials (either because only one dimension varies or because the target and sorting cards match on both dimensions), even when the postswitch rule requires selective attention to a new relevant dimension and the ignoring of the formerly relevant dimension. Other studies have shown that stimulus manipulations that generally make selective attention easier--such as spatially separating the relevant and irrelevant sources of information (Diamond, Carlson, & Beck, 2005; Kloo & Perner, 2005), or giving children help in matching the dimensions through feedback (Bohlmann & Fenson, 2005), through demonstration (Towse et al., 2000), or by asking the children to label the relevant dimensions as they sort postswitch (Kirkham et al., 2003; Towse et al., 2000)--also reduce perserveration. Certainly, there are plausible explanations of these results that do not involve selective attention as a cause of difficulty in switching (e.g., Diamond et al., 2005; Zelazo, 2003). But the results are consistent with the idea that selective attention is an added demand that makes perseveration on the old sorting rule in the DCCS task somehow more likely.

One last and particularly elegant result raises the key theoretical question. Brace, Morton and Munakata (2006) showed that they could increase children’s flexible responding by systematically leading children’s attention away from the preswitch relevant dimension to the postswitch relevant dimension. They did this by a clever structuring of the sorting cards. Figure 3 illustrates the manipulation. Children began by sorting cards that varied on 2 dimensions with one relevant (shape) and the other irrelevant (color); then they were presented with cards that varied on only one dimension the postswitch relevant dimension (color). These then slowly morphed back adding in shape variation so that the formerly relevant properties were now irrelevant. Children readily succeeded in this task.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

These results are intriguing in that at switch, when the cards varying on just one dimension are introduced, children have no option but to attend to this formerly irrelevant dimension (because there is no variation on the formerly relevant dimension). Being forced to attend to a new dimension in this way appears to promote successful later selective attention when the formerly relevant but now irrelevant variation is reintroduced. It is as if being forced to attend to color, for example, made ignoring shape possible. A similar result and conclusion was reported in a pretty much forgotten study by Medin (1973). These results, along with the newer DCCS version by Brace et al. (2006), could be interpreted as attention that is too good or too sticky, an interpretation that has been offered and that we question. The theoretical issue is this: is stickiness a property of “good” selective attention or is it perhaps instead a property of not quite complete and immature selective attention?

Before turning to the literature on the development of selective attention that motivates this alternative perspective, we note the important role of words and instructions in the DCCS task. The task does not require merely basing decisions on one dimension and then the other, but requires appropriately switching attention when given a verbal instruction. Words can be viewed as cues that strengthen attention to one dimension over another. Consistent with this view, Yerys & Munakata (2006) showed labeling the relevant dimension during instruction on preswitch trials results in increased perseveration, but not labeling the preswitch relevant dimension enhances children’s ability to later switch attention to a new dimension. These results might be viewed as indicating (contrary to our argument) that too strong selective attention makes for difficult switching. Similarly, labeling both relevant and irrelevant dimensions when asking children “knowledge” questions about where a card should be placed leads to more incorrect responding than labeling only the relevant dimension (Munakata & Yerys, 2001; Zelazo et al., 1996). Furthermore, having children label the relevant dimension on postswitch trials helps them sort by the new rule (Kirkham et al., 2003; Towse et al., 2000). Finally, the cueing power of words depends on children’s familiarity with the labels (Yerys & Munakata, 2006). All these results suggest that words in this task may serve as cues that direct attention to specific dimensions. A complete accounting of the development of selective attention and attention switching will have to take this into account, along with the fact that the learning of dimension words also undergoes significant development in the period of 3 to 5 years (see Gasser & Smith, 1998, for a review), the same period in which attention switching and, as we will see in the next section, the development of selective attention to dimensions also increases.

In summary, the DCCS task is a selective attention task and the evidence suggests that that fact matters to children’s perseveration. What is not clear is how and why it matters. Children generally sort correctly on the preswitch phase and so it would seem that selective attention per se is not the problem. Instead, the evidence suggests that selective attention might in a sense be too good, too sticky to the first dimension. If this is so, then separate processes that inhibit this prepotent first attentional state seem to be the key to what is developing. This interpretation fits with most current discussions and it has considerable face validity. But as we review next, there is plenty of data to suggest that preschool children’s selective attention is not too strong, but is instead weak and partial. Could this not fully developed selective attention be relevant to a poor ability to switch that attention?

Developmental Trends in Selective Attention to Dimensions

A vast literature about a wide variety of tasks--including discrimination learning, same/different judgments, incidental learning, and card sorting tasks--indicate that, in general, preschool children have considerable difficulty in making decisions or responses based on only one stimulus dimension, unaffected by irrelevant stimulus information (e.g. Enns & Girgus, 1985; Hagen and Hale 1972; Lane & Pearson, 1982; Maccoby & Hagen, 1965;

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Shepp & Barrett, 1991; Shepp & Swartz, 1976; Smith, 1989; Strutt, Anderson & Well, 1975; Tighe, Tighe, & Schechter, 1975). These phenomena and the related accounts about the development of selective attention should be relevant to an understanding of the DCCS task which, in its usual form, asks preschool children to first selectively attend to one dimension and then to another. The idea that children's typically errorless performance on the preswitch trails indicate good (perhaps even too good) selective attention, does not fit with this larger, older, and robust literature on the very protracted developmental course of selective attention.

Discrimination Learning

Discrimination learning is one task implicating generally poor selective attention to one dimension of variation in preschool children (Kendler, 1963; Kendler & Kendler, 1962; Tighe, Glick, & Cole, 1971). In this task, the child is not told which dimension is relevant but is asked to discover the relevant dimension through trial and error. The structure of the task is illustrated in Figure 4A. For example, a child might be shown two items at a time and asked to indicate which wins. In the illustrated example, across trials, black things win, white things lose, and shape is irrelevant. Presented with this task, both younger children and older children can learn the contingencies and reach near errorless performance, although generally speed of learning increases with age. What appears to change dramatically from the preschool period to the school-age period is how children solve the task and what they learn about winning and losing items. There are three different ways of solving the task: 1) memorizing whole stimuli--rather than learning that black wins and white loses, children could learn that black-triangle and black-circle win, and white-triangle and white-circle lose; 2) learning only about relevant attributes, for example, children could learn that black wins and white loses--a possibility that implies selective attention; 3) learning not just about the relevant attributes but also about dimensions; that is, children could learn that color is relevant and shape is irrelevant. Most of the research and the evidence concern the first and the third possibilities.

Performance on transfer tasks indicates that younger children are more likely to learn about wholes while older children, like adults, are more likely to learn about dimensions (Kendler & Kendler, 1962; Tighe, 1973; Tighe et al., 1971). One transfer task that supports this conclusion is the so-called reversal shift task (illustrated in Figure 4B) in which responses reverse, so that the stimuli that won before now lose. Older children do well on the reversal shift transfer task; younger children do not (Kendler, Kendler & Wells, 1960; Tighe, 1973). This task should be easy given selective attention to dimensions in that the same dimension is relevant with only the winning and losing attributes reversed, but it is difficult given learning about wholes (i.e., conjunctions of attributes) in that the response assignments of all individual items change. A second transfer task that also indicates developmental changes in attending to and learning about dimensions is the extradimensional shift (illustrated in Figure 4B). This task shifts the relevant dimension at transfer. The extradimensional shift is easier for the younger children than for the older children (Kendler et al., 1960; Tighe, 1973). Analyses of individual trial performance suggest that it is easier for younger children because half of the individual response assignments remain unchanged and so young children continue to succeed on those items and have to merely learn the other response assignments (Tighe, 1973; Tighe, et al., 1971). The task is presumably more difficult for older children because they must switch attention to a new, and formerly irrelevant, dimension. Finally, a third transfer task that also provides support for increased learning about dimensions rather than whole stimulus items is the intradimensional shift (illustrated in Figure 4B). In this task, new attributes are introduced at transfer for the same dimensions (e.g. new shapes and new colors), with no change in dimension relevance (e.g. color is still relevant and shape is not). Older children perform better on this task than do younger

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

children (Adams & Shepp, 1975; Kendler, Kendler, & Ward, 1972; see also Roberts, Robbins, & Everitt, 1988, with adults; cf. Dickerson, 1966). Again, this task should be easy given selective attention to the relevant dimension since only the new winning attribute needs to be identified, but it is difficult given learning about wholes (or both dimensions) in that all items are new and new response assignments must be learned for each item.

The similarity of the transfer tasks in discrimination learning to the postswitch trials in the DCCS task have been noted by various researchers (Brooks et al., 2003; Perner & Lang, 2002; Zelazo et al., 2003). Further, adult performances in these tasks--and persevatory attentional processes--have been related to processes carried out by the PFC (Owen et al., 1993; Rogers, Andrews, Grasby, Brooks, & Robbins, 2000). Thus, the value of jointly studying development in discrimination learning and the DCCS task has precedence in the literature. However, the perhaps two most important lessons from the older developmental literature on discrimination learning may have been missed. First, there may be different ways of solving the preswitch task--the typically errorless performance of younger and older children on the preswitch DCCS task need not mean that they are doing this task in the same way nor that younger and older children are equally able to attend selectively to one dimension in the preswitch phase. Second, the findings in the discrimination learning literature clearly implicate significant developmental changes in attention to dimensions during the preschool period, and a shift in perception (and learning) from being organized by individual whole items to being organized by dimensions.

Integral and Separable Dimensions

In the mid 1970’s, research on the development of selective attention to dimensions incorporated the then new ideas of Garner (1974) about integral and separable dimensions. Garner famously warned (and showed) that many conclusions about dimensional processing in adults were wrong because experimenters inappropriately assumed that their experimenter-defined dimensions were perceptibly distinguishable sources of stimulus variation for their participants, which is not always the case. He made this point by contrasting separable and integral dimensions. Dimensions such as color and shape are perceptually separable, and, as such, afford selective attention to one dimension. However, dimensions such as saturation and brightness are not perceptually isolated forms of stimulus variation, they are integral, processed holistically, and selective attention is not possible. This distinction is developmentally interesting because adult performance with integral dimensions provides a diagnostic profile for nonselective or holistic attention.

Garner’s research with adults used two kinds of classification tasks to diagnose integral and separable dimensions: the speeded classification task and the free classification task. Further, and in contrast to the characterization of stimuli and dimensions in the DCCS and discrimination learning tasks, these tasks tended to use ordinal dimensions, such that stimuli were not just same or different on a dimension but could differ by degree. In the end, this proves theoretically significant because performance by adults with separable dimensions becomes more integral-like given small stimulus differences (Melara & Mounts, 1993; Nosofsky, 1985).

Speeded Sorting

Speeded classification tasks, like the preswitch phase in the DCCS task, explicitly ask participants to sort stimuli on one relevant dimension while ignoring an irrelevant dimension. Reaction time is the principle dependent measure. Figure 5 illustrates four speeded classification tasks used both in the study of integral and separable dimensions with adults and also in the study of selective attention in children. In all these tasks, participants are presented with multiple repetitions of unique instances and asked to sort them--as rapidly

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

as possible--into two groups by an explicitly stated single dimensional rule (e.g., by size: big versus little). In what is typically called the correlated task, participants sort items that vary redundantly on the relevant and irrelevant dimensions (e.g., big and black versus small and white). In the control task, the items vary only on the relevant dimension. In filtering tasks, the items vary on both the relevant and irrelevant dimensions, and across tasks the amount of irrelevant variation is manipulated.

Figure 5 also lists the main pattern of results for three critical stimulus-by-participant combinations. Given stimuli that vary on adult-integral dimensions such as saturation and brightness, adults sort more rapidly in the correlated task than in the control task, and more rapidly in the control task than in the filtering tasks (Garner, 1974). Moreover, performance in the filtering task deteriorates with increased variation on the irrelevant dimension (Smith & Kemler, 1978). This pattern exemplifies poor selective attention since performance is strongly determined by the nature of the irrelevant variation. Given stimuli that vary on adult-separable dimensions such as size and color, adults sort equally rapidly in all conditions, unaffected by the nature of irrelevant variation (Garner, 1974;Smith & Kemler, 1978). This pattern implies nearly perfect selective attention to the experimenter-defined relevant dimension.

The third pattern is for preschool children (4- to 6-year-olds) with adult-separable dimensions (Smith & Kemler, 1978). This pattern falls between the adult integral and separable patterns. Children sort more rapidly in the correlated than control condition (as do adults with integral dimensions), suggesting a benefit from attending to the redundant irrelevant dimension. Further, the addition of orthogonal irrelevant variation dramatically impairs performance and sorting time, again suggesting that selective attention is not easy. However, increasing irrelevant variation in the filtering task does not decrease performance further, a sharp difference from the adult-integral pattern of results. Smith and Kemler (1978) interpreted this pattern as follows: preschool children do not selectively attend (regardless of instructions) if they do not have to and therefore benefit from the redundant variation in the correlated condition relative to the control condition. However, they can selectively attend if forced by the task (the filtering conditions) although this either adds an extra step or is imperfect, and therefore slows response times.

In sum, preschool children’s performance in speeded selective attention tasks suggests a partially integral-like pattern and again implicates both more holistic (or distributed) attention across relevant and irrelevant dimensions and also extra cognitive effort in tasks that demand selective attention.

Free Classification

The second widely used task in studying integrality-separability in adults and related developmental trends was free classification. This task is of particular interest in this review for four reasons: 1) The task itself is the most different from the DCCS task, discrimination learning and speeded classification, and thus broadens the window onto developing abilities; 2) Performance in this task is strongly predicted by the performance patterns in speeded classification--for adults with integral dimensions, adults with separable dimensions, and for young children (Garner, 1974; Shepp & Swartz, 1976; Smith & Kemler, 1977); 3) the task demands and instructions are minimal and thus the task has been successfully used with children as young as 2 years of age as well as with adults (Smith, 1989); and 4) Formal mathematical models of performance in this task implicate two potentially distinct but related developments that may be critical to the development of attention switching: from more graded to more categorical discriminations of stimulus difference, and from graded to more all-or-none selective attention (Smith, 1989).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

The free classification task works as follows: participants are given a small set of stimuli (typically 3 items) and are asked to partition them into classes of like items. These are not speeded tasks and there is no right or wrong answer. Instead, this task measures the perceived similarities of the items and subjects' *preferences* for forming groups. Figure 6 shows an example of two sets of triads represented in a two-dimensional stimulus space--for example, the stimuli could all be squares that vary on size and shade. As depicted in the Figure, stimuli A and B match in value on one of the two dimensions (e.g., size) and differ markedly on the other dimension; stimuli B and C do not match on either dimension but are similar on both. Given this set of stimuli, if participants group A with B and apart from C, the partitioning would seem to be based on identity on one dimension, implying selective attention to that dimension. In contrast, if participants group B with C and apart from A, the partition is considered by Garner (1974) to be based on overall similarity and to imply nonselective attention to both varying dimensions. Participants are presented (in random order) with multiple versions of these triads with sometimes one dimension and sometimes the other dimension offering the dimensional identity match, a key factor for our analysis.

Figure 6 summarizes patterns of performance for adults with integral dimensions, adults with separable dimensions and children at different developmental levels with adult-separable dimensions. Table 2 provides some specific developmental results and the relevant citations. Given integral dimensions such as saturation and brightness, adults form groups (BC) that maximize within group similarity on both varying dimensions, consistent with nonselective attention to both dimensions (Garner, 1974;Handel & Imai, 1972). Given separable dimensions such as color and size, adults form groups (AB) well organized by a single matching dimensions, consistent with selective attention (Garner, 1974;Handel & Imai, 1972), and, because trial-to-trial they must form such groups sometimes by one dimension and sometimes by the other, their performance also implies the facile switching of attention.

The evidence from children with adult-separable dimensions suggests a developmental trend from more integral-like to more separable-like performance--from attention to both varying dimensions, to selective attention to one dimension, to, ultimately, the facile switching of attention trial-to-trial as the dimension offering the identity match changes. The middle pattern is intriguing: these children exhibit a kind of perseveration--forming AB groups when there is a match on one dimension but not on the other. These children were often called preferrers--children who seem to systematically favor one dimension in their classifications (Aschkenasy & Odom, 1982; Cook & Odom, 1992). So-called preferrers have also been noted in discrimination learning and speed sorting (Caron 1969; Mumbauer & Odom, 1967.).

These results again suggest a protracted course in the development of selective attention--from nonselective attention, to attention that, like preschool children in the DCCS task, appears "sticky", to selective attention that is flexible. Here then is the question, and one that would seem deeply relevant to understanding the current literature on the DCCS task: is this one developmental trend in which the processes that underlie selective attention culminate in switching, or is it perhaps two trends one about selective attention and one about switching? Pertinent to answering this question, Smith (1989) proposed a formal mathematical analysis of the developmental trend in the classification task which showed that they were not explainable without positing two distinct developmental trends, not selection and switching, but two underlying processes each moving from being more graded to being more all-or-none, that in the end may be the key to understanding how selection and switching are related.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

From Graded to All-or-None

Smith (1989) specifically extended the Generalized Context Model (GCM; Nosofsky, 1986) to account for adult performance and the developmental trend in the free classification task. The GCM itself is a well-established model of perceived similarity and selective attention that accounts for extensive sets of experimental findings in the adult literature (e.g., Nosofsky 1992; Zaki, Nosofsky, Stanton & Cohen, 2003). Smith found that within this formal account, two developmentally changing parameters were required to explain the developmental trends in the free classification task. First the pattern from younger to older children's performances could only be explained if one assumed a developmental trend from less perfect or partial selective attention to more perfect or all-or none attention to a dimension. That is, whereas adult performances with separable dimensions could be fit by assuming weights to the two varying dimensions that approximated values of .99 and .01 (almost all attention to one dimension, almost no attention to the other), children's performances--even when they seemed to be attending selectively--were best fit by assuming less perfect selective attention, say weights of .7 to the more attended dimension and .3 to the less attended dimension. Second, the developmental pattern could only be explained by also assuming sharper, more categorical, stimulus discrimination functions for older participants and more graded discrimination functions for younger participants. Sharper, more all-or-none, discrimination is needed to get the mature preference for selective attention to the dimension that offers an exact match in attributes (rather than a near match on the other dimension). Conceptually, this aspect of the developmental trend may be understood as progressing from broad and noisy activation (or generalization) given some stimulus value to sharper and more tightly tuned activation (see also Simmering, Schutte, & Spencer, 2008).

Figure 7 illustrates this idea showing possible distributions of internal activation to two different stimuli, X1 and X2, on one dimension. At the top is a system with broadly tuned and overlapping internal responses; at the bottom is one with sharply tuned, distinct, nearly all-or-none responses. The top is consistent with Smith's model of the perception of small differences as they relate to younger children's responses in the free classification task and the bottom is consistent with the model of older children's responses.

Smith's (1989) theoretical analysis--and the two trends from more graded to more categorical processes--were supported by a number of experimental studies including independent studies of the development of selective attention (Shepp & Swartz, 1976; Smith and Kemler 1977) and the function that relates perceived difference to stimulus differences (Smith & Evans, 1989) as well as tests of new predictions from the model that examined children's classifications of different kinds of sets (Smith, 1989). Smith's analysis of performances in the free classification task did not address the dependence of mature performance on switching the attended dimension trial-to-trial (to the dimension offering the identity match), nor did it offer a coherent account of perseverators (or preferrers) who attend to only one dimension. This last gap in Smith's account was important to several critics of the model (Cook & Odom, 1992; Thompson, 1994).

In summary, the developmental pattern in the free classification task suggests concurrent developmental changes: 1) from more graded to more categorical perceptual discrimination; 2) from more imperfect to more all-or-none selective attention; 3) from "sticky" attention, to the flexible switching of attention from one dimension to another. How are these trends related? And, is graded activation--with respect to discrimination and selective attention--related to attention switching?

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Neurocomputational model

More recently, similar ideas about a transition from graded to all-or-none processes has been suggested in theories of attentional control (such as in the DCCS task). One of these ideas relates to Munakata and colleagues' explanation of the DCCS task in terms of latent and active memory representations. By this account, preswitch sorting forms a long-lasting, latent memory which is strengthened by repetition, while the current, postswitch rule is an active memory which has to be maintained (Morton & Munakata, 2002; see also Unsworth & Engle, 2007). The maintenance of the more recent active memory is proposed to be due to recurrent connections (related to sustained firing in PFC neurons; Fuster & Alexander, 1971; Kojima & Goldman-Rakic, 1982)--i.e., to an internal pattern of activation that feeds back on itself and thus sustains itself even in the face of competing latent memories (see also Clearfield, Diedrich, Smith, & Thelen, in press; Schöner & Dineva, 2007). These recurrent connections, by feeding back on themselves, can also create more sharply tuned (more all-or-none) patterns of activation.

Building on this last idea, O'Reilly (2006) further proposed that these recurrent and thus self-sustaining connections instantiate an intrinsic *bi*stability, yielding a *non-graded* rule-like, go/no go, or all-or-none response that enables the system to rapidly jump from one pattern of self-sustaining activation to a completely different one. O'Reilly has suggested that this pattern of bistability is characteristic of the PFC *and* differentiates it fundamentally from the more graded, incremental and probabilistic processes of other cortical areas.

The potential of these ideas about more graded versus more all-or-none (or rule-like) processes to understanding the development of both selective attention and attention switching is made clear in a related paper by Rougier et al. (2005). That paper, attempted to model in an artificial system the development of selective attention to a single dimension (while ignoring variation on other dimensions) by instantiating three core properties thought to characterize the PFC: (1) recurrent and thus stabilizing patterns of activation, (2) an adaptive gating mechanism which maintains a rule so long as the outcome is predicted but that destabilizes the rule when the predictability of the outcome changes, and (3) modulation of processing in other brain areas, thus enabling the actively maintained rule to orchestrate and capture supporting processes (the so-called executive function of the PFC). A PFC layer exhibiting these processes was embedded within a more standard connectionist network.

The network was taught three types of tasks: (1) labeling attributes on single dimensions (e.g., red, blue), (2) making judgments of sameness and difference on a single dimension (e.g., same color or same shape), and (3) making comparisons on a single dimension (bigger, smaller). Networks with PFC units as described above that were taught all three types of these tasks (but not just any one of them) developed representations, not of individual attributes (red, blue) but of whole dimensions (color, shape) that were highly abstract and orthogonal to each other. This result replicates a previous recurrent connectionist model by Smith, Gasser, & Sandhofer (1997) which also showed that highly abstract (i.e. all-or-none) dimensional representations required training on both labeling (or selective responding to) attributes *and* judgments on the dimension as a whole (e.g., same or different color). The difference in internal activations that underlie early learning about attributes versus the late learning about sameness and difference on dimensions are again about a trend from broadly to more tightly tuned perceptual generalization and are consistent with the ideas in the Smith (1989) model.

This distinction between more graded attribute categories and more all-or-none representations of dimensional sameness, as in both Rougier et al.'s model and Smith et al.'s model, also fit proposals in the adult task-switching literature (e.g., Waszak, Hommel & Allport, 2003) that the stimulus-response bindings that contribute to slowed response at task

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

switch involve matching attributes, not just the same relevant dimensions. This last fact also fits findings that changing the attributes of the objects postswitch in the DCCS task (while keeping the same dimensions) reduces perseveration (Zelazo et al., 2003; Müller et al., 2006; Hanania & Smith, in preparation).

Critically, within Rougier et al.'s (2005) model, the recurrent connections of the PFC *as well as the three training tasks*, enabled highly selective attention--that is finely tuned activation patterns to a single dimension, rather than more graded or distributed representations. These kinds of representations were also critical to the rapid switching of attention among dimensions. According to the model simulations, if the PFC and gating mechanism are not present, then there is little abstraction, and learning is about attributes (not whole dimensions) with more graded, similarity-based decisions. In this model, graded processes shift only incrementally from one state to another while non-graded (or "bistable" or "go/no-go") processes shift rapidly from one state to another. *If this is so, then graded imperfect selective attention and discrimination may be a cause of perseveration*. Thus, perseveration in the DCCS task may derive not from selective attention which is too "good" but from selective attention which--even in the preswitch phase of the task--is not complete and is imperfect.

Thus far, we have juxtaposed two literatures. One is an older literature on the development of selective attention that indicates that during the period between 3 and 6 years of age, children's ability to attend selectively to one dimension undergoes considerable change, becoming less graded, and more all-or-none. Also during this period, children's discrimination of differences along stimulus dimensions becomes sharper and more all-or-none. The more recent literature, on attention switching, tells us that during the same period, children become remarkably better at first sorting cards by one dimension and then switching to sort the same cards by the formally irrelevant dimension. Clearly, these are related phenomena and require a unified account. One possibility is that they are manifestations of the very same underlying processes.

What now?

In many ways, this paper presents a conjecture--not a knock down case--that attention switching in the DCCS task is fundamentally about the immaturity of children's selective attention. The evidence is circumstantial: The DCCS task in its usual form asks children to selectively attend and so the large literature suggesting that selective attention is still undergoing considerable development seems relevant. Aspects of the trend in the two literatures are similar and include what would seem to be overly restricted attention to one dimension prior to switching. Finally, formal--and independently developed--models of the underlying process share some marked similarities, in proposing that the key development is a transition from more graded to more all-or-none processing of dimensions. Given this circumstantial case, the next steps should be empirical studies that specifically examine the relations between children's performances in the two tasks. Table 3 lists a set of "known" and "unknown" aspects of the phenomena as a guide to future studies.

The set of "knowns" and "unknowns" in the first block in the table concerns children's performance in the preswitch phase of the DCCS task. Although performance looks the same at the surface for children who will perseverate postswitch and for children who will flexibly switch, is that performance achieved in the same way? Are children who perseverate, not selectively attending as effectively in the preswitch phase, but instead, perhaps treating the cards as holistic items or only partially, not completely, focusing attention on the relevant dimension? Similarly, if one tests children in an assortment of selective attention tasks that are perhaps more sensitive than the preswitch task of the DCCS

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

task, can one predict from these selective attention performances which children will perseverate in the DCCS task?

The second set of “knowns” and “unknowns” concern the role of words, the instructions that cue children as to the relevant and irrelevant dimensions in the DCCS task. Here there are mostly “unknowns.” Children’s learning of dimension words grows during the same period that selective attention to dimensions grows and during which success in the DCCS task grows. Further, Rougier et al.’s (2005) model and the Smith, Gasser and Sandhofer (1997) both suggest that learning dimension words and making dimensional comparisons builds the representations necessary to support all-or-none selective attention and switching. This is an empirical domain ripe to be explored.

The third block concerns the developmental trend from more graded to more finely tuned discriminations. We know this developmental trend is associated with the development of selective attention. Is it also associated with developmental trends in the DCCS task? If one trains children to more finely discriminate small differences, does one enhance selective attention and switching? Or is this trend not about discrimination per se, but about the stability of internal representations?

Clearly, there is more “unknown” than “known.” Executive function, selective attention and attention switching are fundamentally important developmental achievements. A complete theory of their development should include an account of their developmental relation to each other.

Acknowledgments

This research was supported by grants R01MH060200 and R01HD028675 from the National Institutes of Health.

References

Adams MJ, Shepp BE. Selective attention and the breadth of learning: A developmental study. Journal of Experimental Child Psychology. 1975; 20(1):168–180.

Aschkenasy JR, Odom RD. Classification and perceptual development: Exploring issues about integrality and differential sensitivity. Journal of Experimental Child Psychology. 1982; 34(3):435–448. [PubMed: 7153703]

Bohlmann NL, Fenson L. The Effects of Feedback on Perseverative Errors in Preschool Aged Children. Journal of Cognition and Development. 2005; 6(1):119–131.

Brace JJ, Morton JB, Munakata Y. When Actions Speak Louder Than Words: Improving Children's Flexibility in a Card-Sorting Task. Psychological Science. 2006; 17(8):665–669. [PubMed: 16913947]

Brooks PJ, Hanauer JB, Padowska B, Rosman H. The role of selective attention in preschoolers' rule use in a novel dimensional card sort. Cognitive Development. 2003; 18(2):195–215.

Caron AJ. Discrimination shifts in three-year-olds as a function of dimensional salience. Developmental Psychology. 1969; 1(4):333–339.

Clearfield MW, Dineva E, Smith LB, Diedrich FJ, Thelen E. Cue salience and infant perseverative reaching: Tests of the Dynamic Field theory. Developmental Science. in press.

Cook GL, Odom RD. Perception of multidimensional stimuli: A differential-sensitivity account of cognitive processing and development. Journal of Experimental Child Psychology. 1992; 54(2):213–249. [PubMed: 1431734]

Diamond A, Carlson SM, Beck DM. Preschool Children's Performance in Task Switching on the Dimensional Change Card Sort Task: Separating the Dimensions Aids the Ability to Switch. Developmental Neuropsychology. 2005; 28(2):689–729. [PubMed: 16144433]

Dickerson DJ. Performance of preschool children on three discrimination shifts. Psychonomic Science. 1966; 4(12):417–418.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Enns JT, Girgus JS. Developmental changes in selective and integrative visual attention. Journal of Experimental Child Psychology. 1985; 40(2):319–337. [PubMed: 4045383]

Fan J, McCandliss BD, Fossella J, Flombaum JI, Posner MI. The activation of attentional networks. NeuroImage. 2005; 26(2):471–479. [PubMed: 15907304]

Frye D, Zelazo PD, Palfai T. Theory of mind and rule-based reasoning. Cognitive Development. 1995; 10(4):483–527.

Fuster JM, Alexander GE. Neuron activity related to short-term memory. Science. 1971; 173(3997): 652–654. [PubMed: 4998337]

Gasser M, Smith LB. Learning nouns and adjectives: A connectionist account. Language and Cognitive Processes.Special Issue: Language Acquisition and Connectionism. 1998; 13(2&3): 269–306.

Garner, WR. The processing of information and structure. Oxford, England: Lawrence Erlbaum; 1974. p. 203

Hagen, JW.; Hale, GH. The development of attention in children. In: Pick, A., editor. Minnesota symposia on child psychology. Oxford, England: U. Minnesota Press; 1972.

Hanania R, Smith LB. The development of selective attention and attention shifting: Are they related? (in preparation).

Handel S, Imai S. The free classification of analyzable and unanalyzable stimuli. Perception & Psychophysics. 1972; 12(1-B):108–116.

Jacques S, Zelazo PD, Kirkham NZ, Semcesen TK. Rule selection versus rule execution in preschoolers: An error-detection approach. Developmental Psychology. 1999; 35(3):770–780. [PubMed: 10380867]

Kemler DG. Classification in young and retarded children: The primacy of overall similarity relations. Child Development. 1982; 53(3):768–779. [PubMed: 7094679]

Kemler DG, Smith LB. Accessing similarity and dimensional relations: Effects of integrality and separability on the discovery of complex concepts. Journal of Experimental Psychology: General. 1979; 108(2):133–150. [PubMed: 528902]

Kendler HH, Kendler TS. Vertical and horizontal processes in problem solving. Psychological Review. 1962; 69(1):1–16. [PubMed: 14455127]

Kendler HH, Kendler TS, Ward JW. An ontogenetic analysis of optional intradimensional and extradimensional shifts. Journal of Experimental Psychology. 1972; 95(1):102–109. [PubMed: 5070275]

Kendler TS. Development of Mediating Responses in Children. Monographs of the Society for Research in Child Development. 1963; 28(2):33–48. [PubMed: 14031906]

Kendler TS, Kendler HH, Wells D. Reversal and nonreversal shifts in nursery school children. Journal of Comparative and Physiological Psychology. 1960; 53(1):83–88. [PubMed: 14408484]

Kirkham NZ, Cruess L, Diamond A. Helping children apply their knowledge to their behavior on a dimension-switching task. Developmental Science. 2003; 6(5):449–476.

Kloo D, Perner J. Training transfer between card sorting and false belief understanding: Helping children apply conflicting descriptions. Child Development. 2003; 74(6):1823–1839. [PubMed: 14669898]

Kloo D, Perner J. Disentangling dimensions in the dimensional change card-sorting task. Developmental Science. 2005; 8(1):44–56. [PubMed: 15647066]

Kojima S, Goldman-Rakic PS. Delay-related activity of prefrontal neurons in rhesus monkeys performing delayed response. Brain Research. 1982; 248(1):43–50. [PubMed: 7127141]

Kruschke JK. Toward a unified model of attention in associative learning. Journal of Mathematical Psychology. 2001; 45(6):812–863.

Lane DM, Pearson DA. The development of selective attention. Merrill-Palmer Quarterly. 1982; 28(3): 317–337.

Maccoby EE, Hagen JW. Effects of distraction upon central versus incidental recall: Developmental trends. Journal of Experimental Child Psychology. 1965; 2(3):280–289.

Medin DL. Measuring and training dimensional preferences. Child Development. 1973; 44(2):359–362. [PubMed: 4705562]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Melara RD, Mounts JR. Selective attention to Stroop dimensions: Effects of baseline discriminability, response mode, and practice. Memory & Cognition. 1993; 21(5):627–645.

Morton JB, Munakata Y. Active versus latent representations: A neural network model of perseveration, dissociation, and decalage. Developmental Psychobiology. Special Issue: Converging method approach to the study of developmental science. 2002; 40(3):255–265.

Müller U, Dick AS, Gela K, Overton WF, Zelazo PD. The Role of Negative Priming in Preschoolers' Flexible Rule Use on the Dimensional Change Card Sort Task. Child Development. 2006; 77(2): 395–412. [PubMed: 16611180]

Mumbauer CC, Odom RD. Variables affecting the performance of preschool children in intradimensional, reversal, and extradimensional shifts. Journal of Experimental Psychology. 1967; 75(2):180–187. [PubMed: 6062958]

Munakata Y, Yerys BE. All together now: When dissociations between knowledge and action disappear. Psychological Science. 2001; 12(4):335–337. [PubMed: 11476102]

Mundy P, Block J, Delgado C, Pomares Y, Van Hecke AV, Parlade MV. Individual Differences and the Development of Joint Attention in Infancy. Child Development. 2007; 78(3):938–954. [PubMed: 17517014]

Nosofsky R. Overall similarity and the identification of separable-dimension stimuli: A choice model analysis. Perception & Psychophysics. 1985; 38(5):415–432. [PubMed: 3831920]

Nosofsky RM. Attention, similarity, and the identification-categorization relationship. Journal of Experimental Psychology: General. 1986; 115(1):39–57. [PubMed: 2937873]

Nosofsky, RM. Exemplar-based approach to relating categorization, identification, and recognition. In: Ashby, FG., editor. Multidimensional models of perception and cognition. 1992. p. 363-393.

O'Reilly RC. Biologically Based Computational Models of High-Level Cognition. Science. 2006; 314(5796):91–94. [PubMed: 17023651]

Owen AM, Roberts AC, Hodges JR, Summers BA, Polkey CE, Robbins TW. Contrasting mechanisms of impaired attentional set-shifting in patients with frontal lobe damage or Parkinson's disease. Brain. 1993; 116:1159–1175. [PubMed: 8221053]

Perner J, Lang B. What causes 3-year-olds' difficulty on the dimensional change card sorting task? Infant and Child Development. Special Issue: Executive functions and development: Studies of typical and atypical children. 2002; 11(2):93–105.

Posner MI, Rothbart MK. Research on Attention Networks as a Model for the Integration of Psychological Science. Annual Review of Psychology. 2007; 58:1–23.

Reynolds JR, Braver TS, Brown JW, Van der Stigchel S. Computational and neural mechanisms of task switching. Neurocomputing: An International Journal.Special Issue: Computational Neuroscience: Trends in Research 2006. 2006; 69(10–12):1332–1336.

Roberts AC, Robbins TW, Everitt BJ. The effects of intradimensional and extradimensional shifts on visual discrimination learning in humans and nonhuman primates. The Quarterly Journal of Experimental Psychology B: Comparative and Physiological Psychology. 1988; 40(4-B):321–341.

Rogers RD, Andrews TC, Grasby PM, Brooks DJ, Robbins TW. Contrasting cortical and subcortical activations produced by attentional-set shifting and reversal learning in humans. Journal of Cognitive Neuroscience. 2000; 12(1):142–162. [PubMed: 10769312]

Rougier NP, Noelle DC, Braver TS, Cohen JD, O'Reilly RC. Prefrontal cortex and flexible cognitive control: Rules without symbols. PNAS Proceedings of the National Academy of Sciences of the United States of America. 2005; 102(20):7338–7343.

Rueda MR, Posner MI, Rothbart MK. The Development of Executive Attention: Contributions to the Emergence of Self-Regulation. Developmental Neuropsychology. 2005; 28(2):573–594. [PubMed: 16144428]

Ruff, HA.; Rothbart, MK. Attention in early development: Themes and variations. New York, NY, US: Oxford University Press; 1996. p. 294

Ruskin EM, Kaye DB. Developmental differences in visual processing: Strategy versus structure. Journal of Experimental Child Psychology. 1990; 50(1):1–24. [PubMed: 2398327]

Schöner G, Dineva E. Dynamic instabilities as mechanisms for emergence. Developmental Science. 2007; 10(1):69–74. [PubMed: 17181702]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Shepard RN. Attention and the metric structure of the stimulus space. Journal of Mathematical Psychology. 1964; 1(1):54–87.

Shepp BE, Barrett SE. The development of perceived structure and attention: Evidence from divided and selective attention tasks. Journal of Experimental Child Psychology. 1991; 51(3):434–458. [PubMed: 2072086]

Shepp, BE.; Burns, B.; McDonough, D. The relation of stimulus structure to perceptual and cognitive development: Further tests of a separability hypothesis. In: Wilkening, F.; Becker, J.; Trabasso, T., editors. Information integration by children. Hillsdale, NJ: Erlbaum; 1980. p. 113-145.

Shepp BE, Swartz KB. Selective attention and the processing of integral and nonintegral dimensions: A developmental study. Journal of Experimental Child Psychology. 1976; 22(1):73–85. [PubMed: 1003096]

Simmering VR, Schutte AR, Spencer JP. Generalizing the dynamic field theory of spatial cognition across real and developmental time scales. Brain Research.Special Issue: Computational Cognitive Neuroscience. 2008; 1202:68–86.

Smith LB. Development of classification: The use of similarity and dimensional relations. Journal of Experimental Child Psychology. 1983; 36(1):150–178.

Smith LB. A model of perceptual classification in children and adults. Psychological Review. 1989; 96(1):125–144. [PubMed: 2928416]

Smith LB, Evans P. Similarity, identity, and dimensions: Perceptual classification in children and adults. Shepp, Bryan E. 1989

Smith LB, Gasser M, Sandhofer CM. Learning to talk about the properties of objects: A network model of the development of dimensions. Goldstone, Robert L. 1997; 36:219–255.

Smith LB, Kemler DG. Developmental trends in free classification: Evidence for a new conceptualization of perceptual development. Journal of Experimental Child Psychology. 1977; 24(2):279–298.

Smith LB, Kemler DG. Levels of experienced dimensionality in children and adults. Cognitive Psychology. 1978; 10(4):502–532. [PubMed: 699515]

Strutt GF, Anderson DR, Well AD. A developmental study of the effects of irrelevant information of speeded classification. Journal of Experimental Child Psychology. 1975; 20(1):127–135.

Thompson LA. Dimensional strategies dominate perceptual classification. Child Development. 1994; 65(6):1627–1645. [PubMed: 7859546]

Tighe, T. Subproblem analysis of discrimination learning. In: Bower, GH., editor. The psychology of learning and motivation: Advances in research and theory. Oxford, England: Academic Press; 1973.

Tighe TJ, Glick J, Cole M. Subproblem analysis of discrimination-shift learning. Psychonomic Science. 1971; 24(4):159–160.

Tighe TJ, Tighe LS, Schechter J. Memory for instances and categories in children and adults. Journal of Experimental Child Psychology. 1975; 20(1):22–37.

Towse JN, Redbond J, Houston-Price CMT, Cook S. Understanding the dimensional change card sort: Perspectives from task success and failure. Cognitive Development. 2000; 15(3):347–365.

Unsworth N, Engle RW. The Nature of Individual Differences in Working Memory Capacity: Active Maintenance in Primary Memory and Controlled Search From Secondary Memory. Psychological Review. 2007; 114(1):104–132. [PubMed: 17227183]

Ward TB. Separable and integral responding by children and adults to the dimensions of length and density. Child Development. 1980; 51(3):676–684. [PubMed: 7418505]

Waszak F, Hommel B, Allport A. Task-switching and long-term priming: Role of episodic stimulus-task bindings in task-shift costs. Cognitive Psychology. 2003; 46(4):361–413. [PubMed: 12809680]

Yerys BE, Munakata Y. When Labels Hurt but Novelty Helps: Children's Perseveration and Flexibility in a Card-Sorting Task. Child Development. 2006; 77(6):1589–1607. [PubMed: 17107448]

Zaki SR, Nosofsky RM, Stanton RD, Cohen AL. Prototype and exemplar accounts of category learning and attentional allocation: A reassessment. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2003; 29(6):1160–1173.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Zelazo PD, Frye D, Rapus T. An age-related dissociation between knowing rules and using them. Cognitive Development. 1996; 11(1):37–63.

Zelazo PD, Müller U, Frye D, Marcovitch S. What do children perseverate on when they perseverate? Monographs of the Society for Research in Child Development. 2003; 68(3)

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

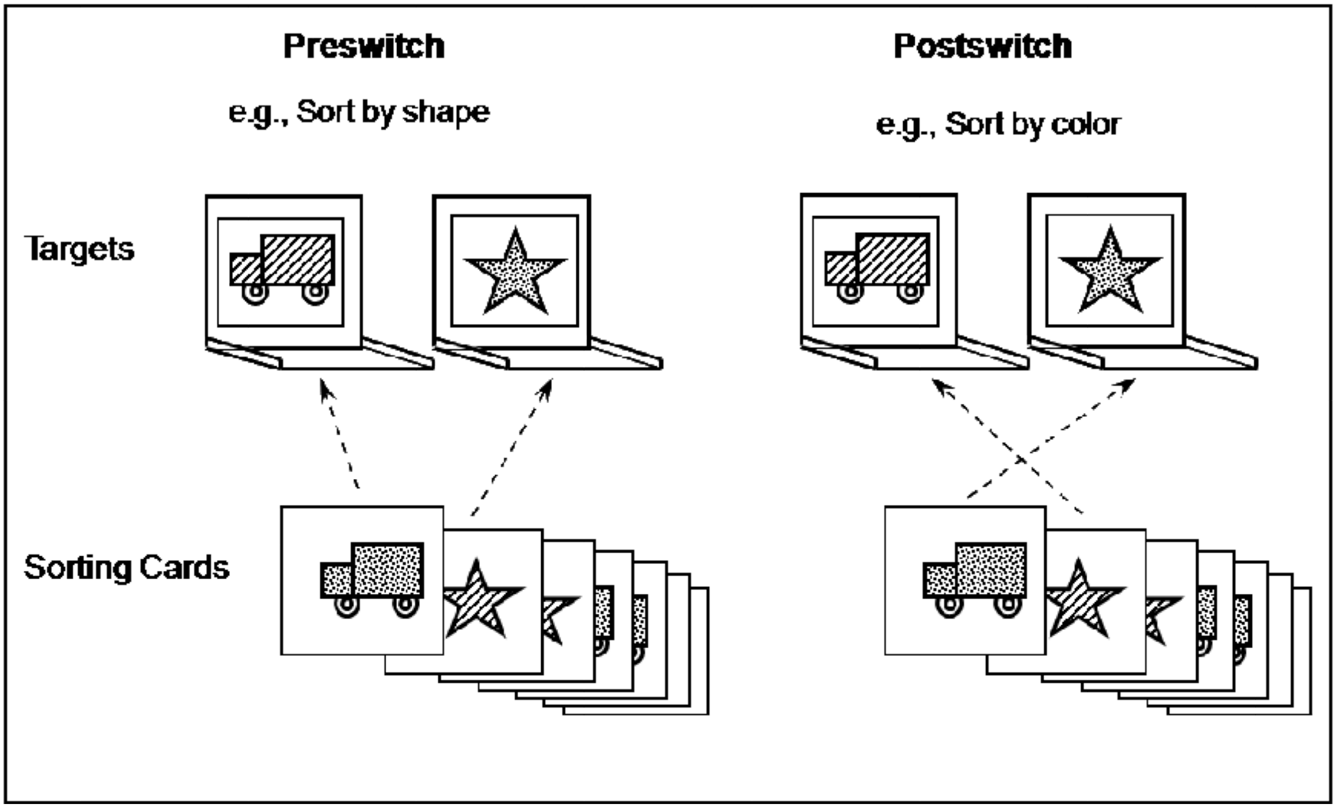


Figure 1.

Example of cards and trays in a standard dimensional-change card-sorting (DCCS) task. Sorting cards are presented in random order. Note: stripes indicate red; dots indicate blue.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

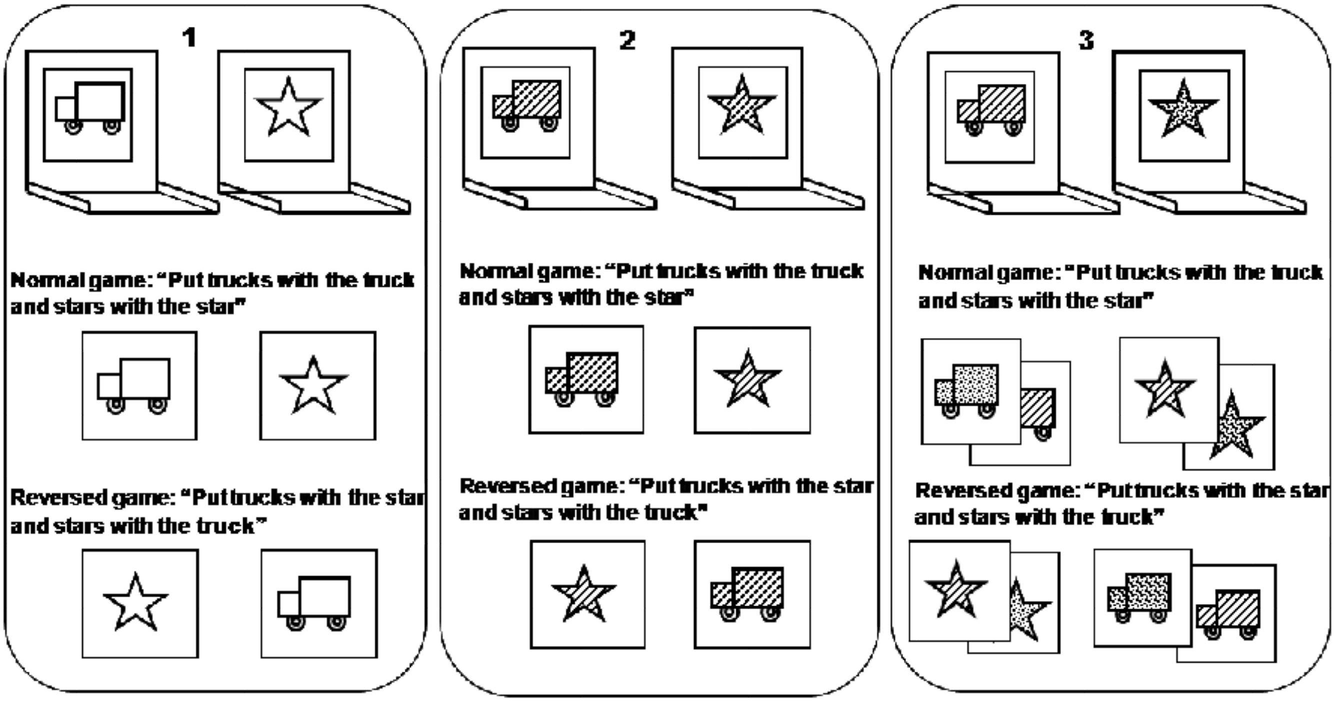


Figure 2.

The structure of three experiments manipulating the number of varying dimensions (all illustrated as red and blue, trucks and stars for ease of comparison). Only one dimension--shape--is relevant throughout; responses are reversed postswitch without changing the relevant dimension. Panel 1: black/white line-drawings (Brooks et al, 2003, exp 1); Panel 2: single-color line-drawings (Perner and Lang, 2002, reversal shift); and Panel 3: line-drawings with irrelevant color variation (Brooks et al, 2003, exp 2). Note: white indicates no color; stripes indicate red; dots indicate blue.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

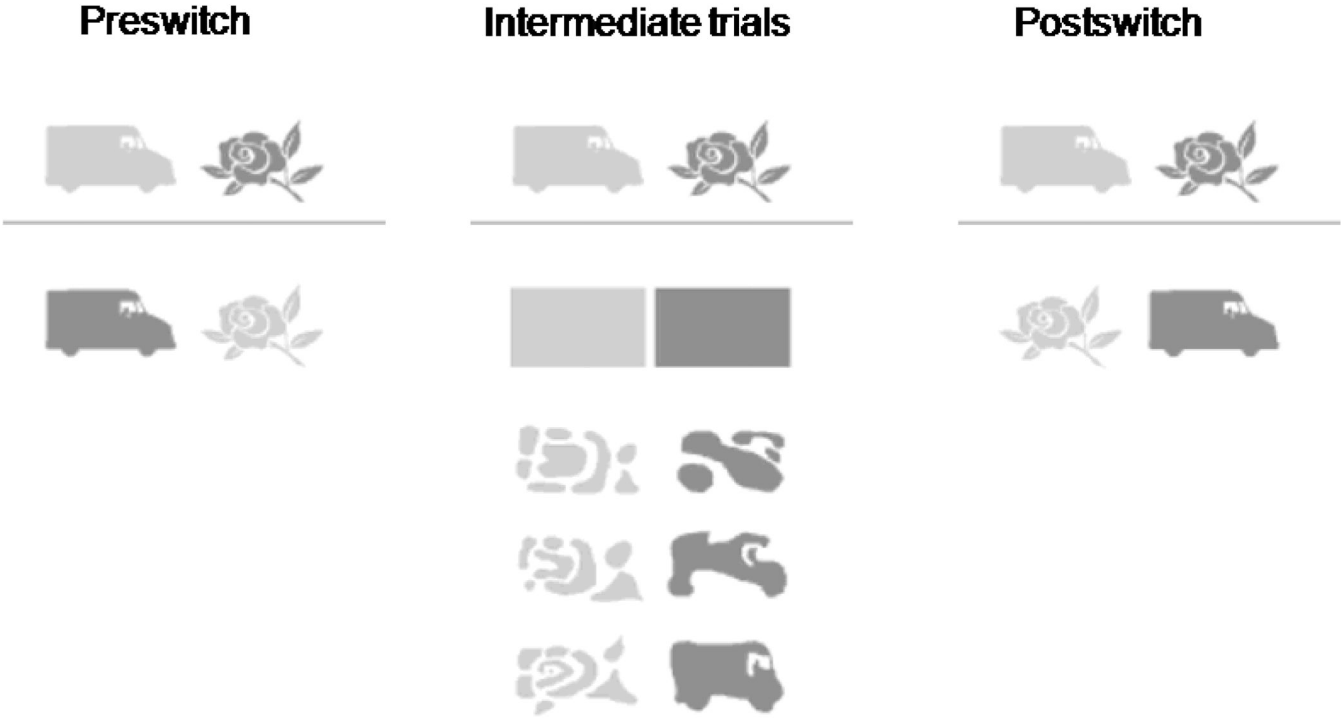


Figure 3.

Adapted from Brace, Morton and Munakata (2006), Figure 1. In their task, children sorted by shape on preswitch trials and by color postswitch, with an intermediate phase involving uni-dimensional sorting cards (with only the postswitch relevant dimension) which were then gradually morphed back to include the irrelevant dimension of shape.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

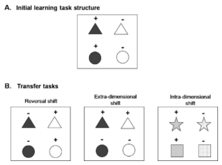


Figure 4.
Discrimination learning task. Initial learning task (panel A) and three transfer learning tasks (panel B). Plus sign indicates a “winning” item, minus sign indicates a “losing” item.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

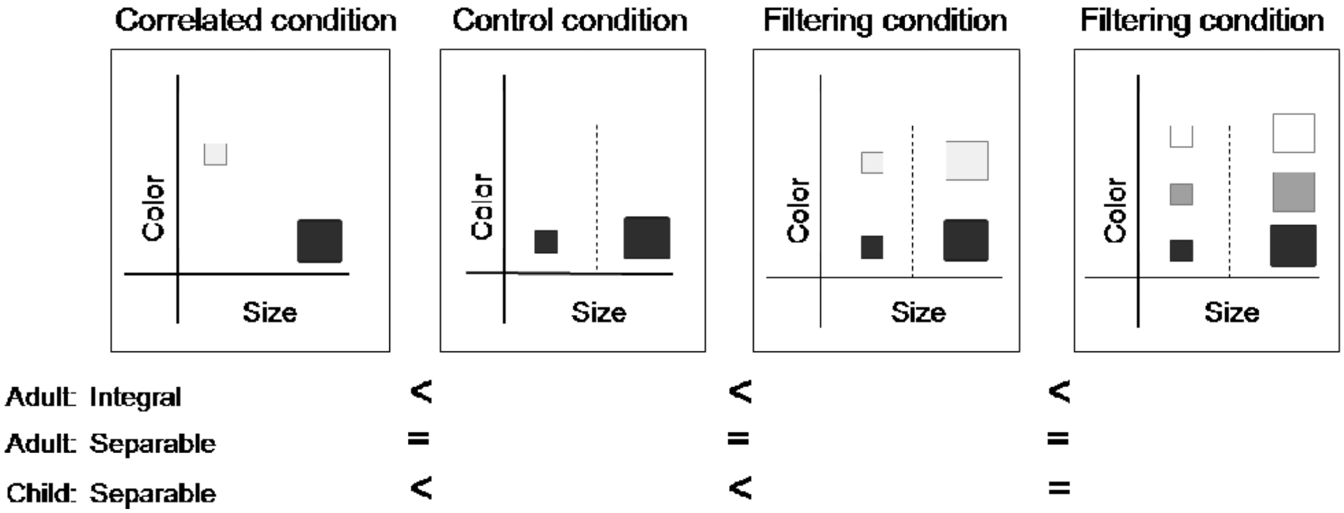
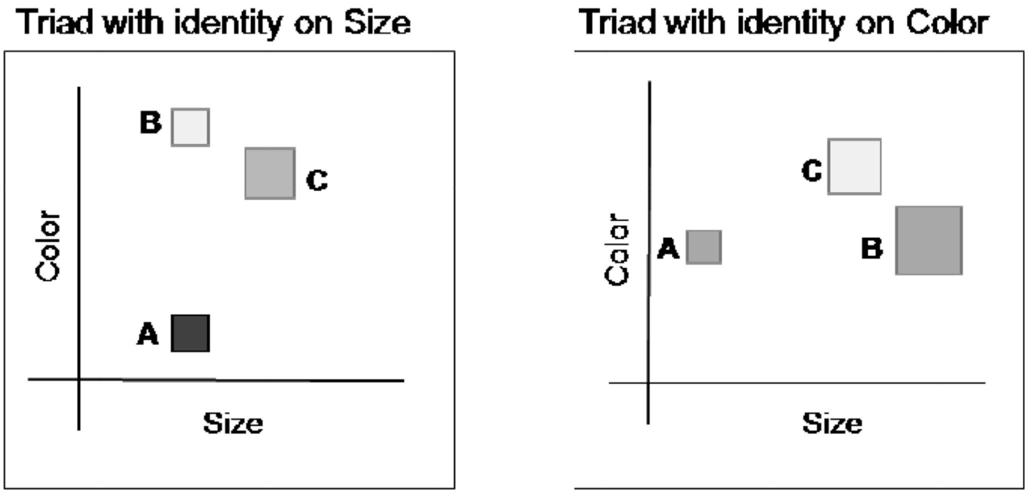


Figure 5.
Sample speeded sorting stimuli, represented in stimulus space, for each of the sorting conditions. Indicated below is the relative performance on each condition for children and adults (for separable and integral dimensions).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript



Adults: Integral	BC	BC
Adults: Separable	AB	AB
2 – 5-year-olds: Separable	BC	BC
5 – 7-year-olds: Separable	AB	BC
> 8-year-olds: Separable	AB	AB

Figure 6.

Two sample triads represented in stimulus space. The one on the left has a match on size, the other matches on color. Listed below are the typical classes formed for each age group and for integral versus separable dimensions.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

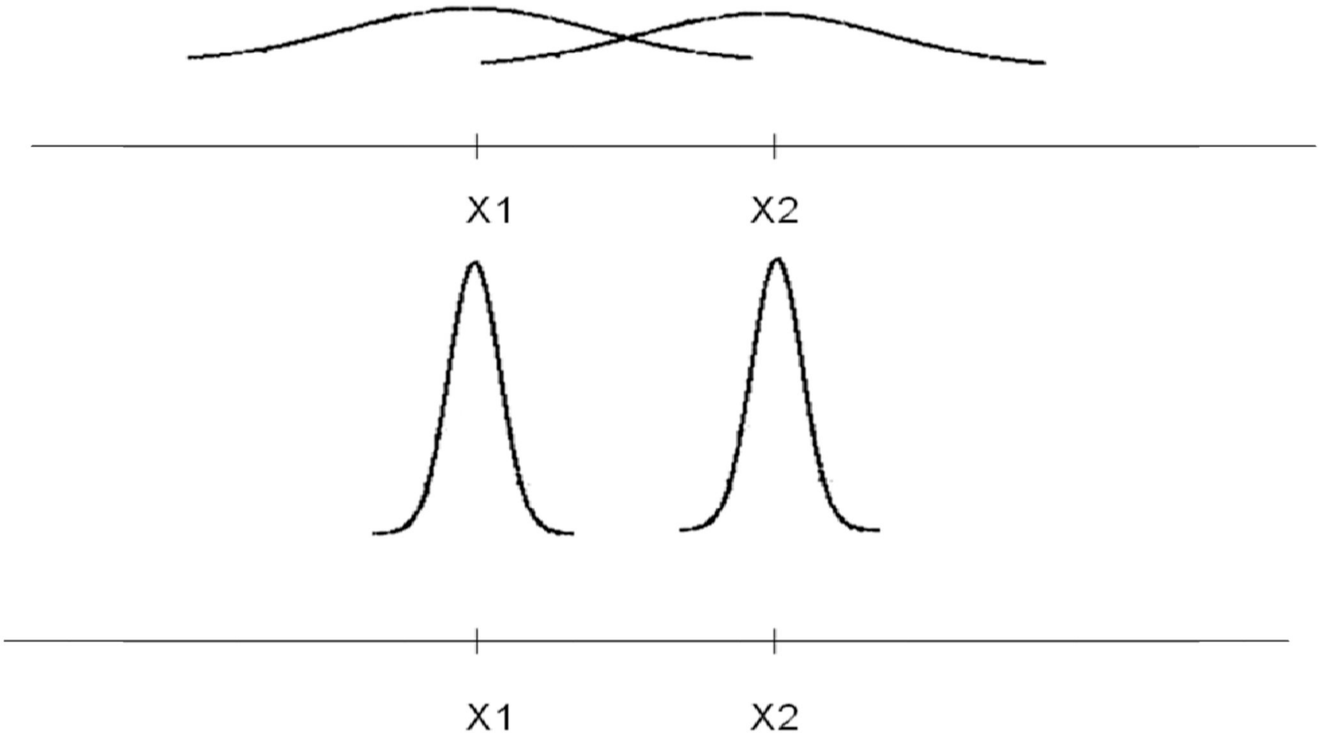


Figure 7.

Stimulus discrimination curves for two stimuli, X1 and X2. The top curves represent more graded discrimination functions, as younger children exhibit, and the bottom curves represent more sharply defined discrimination functions, as older children exhibit.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 1

Sample results from DCCS experiments

Reference	Description	Measure[a]	Correct Sorts[b]		
			3 yr	**4 yr**	**5 yr**
Frye et al.,1995; Experiment 1	Standard – geometric shapes	Trials	20%	26%	82%
Frye et al.,1995; Experiment 2	Standard – various pairs of dimensions	Children	42%	72%	
Zelazo et al.,1996; Experiment 1	Standard	Children	40%	90%	
Perner & Lang , 2002 (estimated from Figure 2)	Standard	Trials	46%	84%	
Kirkham et al., 2003	Standard	Children	42%	92%	
Perner & Lang , 2002; (estimated from Fig 2)	Reversal shift (shape always relevant) – standard cards	Trials	88%	98%	
Perner & Lang , 2002 (estimated from Fig 2)	Puppet target cards	Trials	88%	94%	
Brooks et al., 2003; Experiment 1	Reversal shift (shape always relevant) – uncolored line drawings	Children	79%		
Brooks et al. , 2003; Experiment 2	Reversal shift (shape always relevant) – irrelevant color variation	Trials	29%	58%	
Zelazo et al.,2003; Experiment 9	Redundant condition: sorting cards match targets on preswitch trials	Children	69%		
Zelazo et al.,2003; Experiment 7	Total Change condition: new shapes and colors used postswitch	Children	79%		
Kirkham et al., 2003	Children label relevant dimension	Children	78%	85%	
Kloo and Perner 2005; Experiment 1	Fully separated dimensions (on both target and test cards)	Trials	85%		
Brace, Morton, Munakata 2006; Scaffolding condition	Uni-dimensional (color) sorting cards are morphed back to include shape	Children	81%		

Note. Above the dark line are measures from standard DCCS tasks. Below that line are variations on the task, cited in the paper for demonstrating selective attention.

[a] Measure: children = % children passing the postswitch phase (i.e. no more than 1 error); trials = % of correctly sorted cards in postswitch phase.

[b] Some percentages listed in this table were estimated from graphs.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 2

Mean proportion AB (dimensional identity) and BC (overall similarity) classifications

Reference	Description	Age	BC	AB
Smith & Kemler, 1977, exp 1	Triad; size-shade			
		5–6 yrs	0.56	0.34
		7–8 yrs	0.40	0.50
		10–11 yrs	0.22	0.68
Smith & Kemler, 1977, exp 2	Tetrad: Size-shade			
		5–6 yrs	.58	.34
		10–11 yrs	.18	.80
Ward, 1980, exp 1	Triad: length-density			
(columnar condition)		4–6 yrs	.56	.28
		Adults	.34	.61
Shepp, Burns & McDonough, 1980, exp 1	Triad: size-shade			
		4 yrs	0.80	.13
		6–7 yrs	0.45	.48
		11–12 yrs	0.21	.79
Shepp et al., 1980, exp 1	Triad: size-angle			
		4 yrs	0.78	.15
		6–7 yrs	0.43	.50
		11–12 yrs	0.26	.67
Kemler, 1982, exp 1	Triad: size-shade			
		4–5 yrs	.70	.20
Smith, 1983, exp 1	Tetrad: size-shade			
		4–5 yrs	.66	.29
		5–6 yrs	.66	.30
Smith, 1989, exp 1[a]	Triad: size-shade			
		2 yrs	.48	.20
		3 yrs	.55	.24
		4 yrs	.52	.28
		5 yrs	.62	.28
		8yrs	.30	.70
		Adult	.03	.97
Ruskin & Kaye, 1990	Triad: size-shade			
		5–6 yrs	*.33*	*.60*
		11–12 yrs	.16	.82

Note. Overall similarity proportions for some experiments are calculated from dimensional identity and haphazard grouping.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

[a] In Smith's (1989) experiment, items could be grouped freely, so there were more than 2 possibilities. (Numbers were estimated from Smith, 1989, Figure 11).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 3

Know	Don't Know
In the DCCS task, both older and younger children succeed on the preswitch phase. In Garner speeded classification and free classification tasks, there is a protracted developmental course for selective attention. In the free classification task, the developmental trend appears to be from non-selective attention to perseveration to flexible switching.	Do perseverators and non-perseverators solve the preswitch task in same way? Do children who perseverate in the DCCS task also also show poorer selective attention in the Garner tasks? Is performance in the free classification task predictive of performance in the DCCS task? Is perseveration in the free classification task the same as in the DCCS task? Does one predict the other?
Dimension words cue selective attention.	Do these words cue younger and older children in the same way? Are there measurable weaknesses in younger children's word knowledge that would predict poorer selective attention in Garner tasks and/or more perseveration in DCCS tasks? Is it knowledge of dimensions (and experience in dimensional comparison) that fosters selective attention and attention switching rather than knowledge of attributes? Do dimension words help selective attention in the free classification task in the same way as in the DCCS task?
There is a developmental trend from more graded discrimination of stimulus differences to sharper and more finely tuned discriminations. Smaller stimulus differences and harder stimulus discriminations are associated with poorer selective attention.	Does the magnitude of stimulus differences or degree of discriminability in the DCCS task matter? Is more graded or noisier discrimination and generalization associated with increased perseveration in the DCCS task?