Dual Routes to Cognitive Flexibility: Learning and Response-Conflict Resolution in the Dimensional Change Card Sort Task

Michael Ramscar
University of Tübingen

Melody Dye
Indiana University

Jessica W. Gustafson and Joseph Klein
Stanford University

Cognitive control, the ability to align our actions with goals or context, is largely absent in children under four. How then are preschoolers able to tailor their behavior to best match the situation? Learning may provide an alternative route to context-sensitive responding. This study investigated this hypothesis in the Dimensional Change Card Sort (DCCS), a classic test of cognitive control that most under-fours fail. A training intervention based on learning theoretic principles proved highly effective: Three-year-olds who learned about DCCS rules and game contexts in a card-labeling task, subsequently transferred this knowledge to sorting in the DCCS, passing at more than 3 times the rate of controls ($N = 47$). This surprising finding reveals much about the nature of the developing mind.

Although a 3-year-old girl might appear to simply be a smaller version of her 4-year-old brother, studies have revealed some surprising differences between them: While her older brother will successfully pass most of the tasks that have been devised to test the cognitive capabilities of the very young, odds are that she will fail every one of them. At test, a typical 4-year-old can ably navigate the conflicting dimensions of appearance and reality, correctly judge questions of false belief, successfully switch between competing rules in the Dimensional Change Card Sort (DCCS), and when faced with a forced choice task, reliably choose between two alternatives according to their prior probabilities. An average 3-year-old, on the other hand, will maximize in that same task (fixating on the most likely response), fail to distinguish reality from false belief and appearance from reality in standard batteries, and fail to switch from one sorting rule to another in the DCCS, even when the new rule is clearly stated (for reviews, see Diamond, 2002; Hanania & Smith, 2009; Ramscar & Gitcho, 2007).

Over the past two decades, the underlying causes of these phenomena have been much debated, and many conflicting explanations have been offered to account for 3-year-olds' curious performance in these tasks (see, e.g., cognitive control & complexity theory, Zelazo, Müller, Frye, & Marcovitch, 2003; active vs. latent representations, Cepeda & Munakata, 2007; Diamond, Carlson, & Beck, 2005; Morton & Munakata, 2002; Yerys & Munakata, 2006; attentional inertia, Kirkham, Cruess, & Diamond, 2003; redescription, Kloo & Perner, 2005; Kloo, Perner, Kerschhuber, Dabernig, & Aichhorn, 2008; negative priming, Müller, Dick, Gela, Overton, & Zelazo, 2006; Perner, Stummer, & Lang, 1999). While the bulk of these proposals have been aimed at accounting for why 3- and 4-year-olds perform so differently, an equally puzzling question has been accorded rather less attention: Given the inflexibility of thought revealed in these tests, why is this rigidity so difficult to detect in the normal course of events? Why is it that young children appear to be capable of responding in flexible and context sensitive ways in some situations, but not in others (Brooks, Hanauer, Padowska, & Rosman, 2003; Deák, 2003)?

Here, we outline a solution to this puzzle, based on a consideration of the different ways in which a

This material is based on work supported by the National Science Foundation under Grants 0547775 and 0624345 and by the National Science Foundation/IGERT Training Program in the Dynamics of Brain-Body-Environment Systems at Indiana University. We are grateful to Natasha Kirkham, Sharon Thompson-Schill, and Stewart McCauley for discussion of these ideas.

Correspondence concerning this article should be addressed to Michael Ramscar, Department of Linguistics, Eberhard Karls University, Tübingen, Germany 72074. Electronic mail may be sent to michael.ramscar@uni-tuebingen.de.

child might resolve the conflict between competing ways of responding to a given situation, so as to choose the response most appropriate to that context. We propose that there are at least two ways in which response conflict can be handled in the mind: dynamic response conflict resolution, which processes and resolves conflicting response demands online, as they arise, and contextual discrimination learning, which shapes the degree to which responses are evoked in specific contexts and which can serve to reduce (or even eliminate) potential response conflicts that would otherwise need to be resolved online. While older children and adults will in most cases have access to both "routes" of response-conflict resolution, we suggest that children under four are generally limited to one: *learning*. Under-fours are able to match their behavior to context in remarkably subtle and sensitive ways, but only once they have learned to do so. If they have not learned to match a response or a behavior to a context, under-fours' difficulty with handling online response conflict proves their undoing in psychological tests.

Response Conflict and the DCCS Task

In the standard DCCS, children are asked to sort cards with two prominent dimensions—color and shape—into bins in which these dimensions have been reversed (Figure 1). For example, if a child is holding cards with red stars and blue trucks, the bins will be marked with blue stars and red trucks. If the child is asked to sort by color, the red stars will go with the red trucks and the blue stars will go with the blue trucks; if the child is asked to sort by shape, the red stars will go with the blue stars,

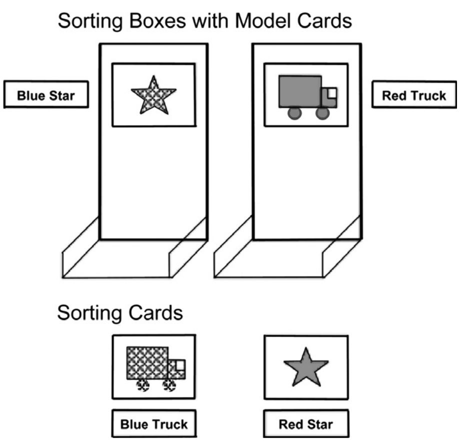

Figure 1. The Dimensional Change Card Sort (DCCS) task. Cards can be sorted by *shape* (in which case, the red star card is sorted into the left bin) or *color* (in which case, the red star card is sorted to the right bin).

and the red trucks will go with the blue trucks. Importantly, when a child is asked to sort by one dimension—say, shape—changing the sort dimension to color will also change the correct sort bin: For example, red stars go in the truck bin when sorted by color, but the star bin when sorted by shape. Successful sorting in the postswitch thus requires the child to actively choose between conflicting responses, suppressing the dominant sorting response (which has been strongly associated with the cards over the initial game trials) in favor of a contradictory alternative (which is now relevant to the task at hand).

Response conflict will arise whenever the requirements of a specific task conflict with an equally or more strongly learned pattern of responding that is prompted by the same context. To successfully resolve this conflict, an individual must be able to effectively override the biased response in favor of a less well-learned (or less well-primed) response that is more appropriate to the context (Novick, Trueswell, & Thompson-Schill, 2010). In adults, the ability to actively and selectively respond in this way can be easily illustrated in relation to the Stroop test, in which subjects are asked to identify the ink color of a conflicting color word. For example, if *green* is printed in red ink, *red* is the appropriate response. Success in this task involves resolving the conflict between the dominant response (reading) and a disfavored but contextually appropriate response (ink naming). Although the undertaking is cognitively demanding (especially when a time component is introduced), most adults are able to pass the Stroop if they are given time to formulate their response.

The DCCS imposes a similar set of task demands, but in a game designed for children. To succeed in the DCCS, a child must be able to resolve response conflict online as the sorting requirements change between games. Unlike adults and older children, however, 3-year-olds not only tend to fail the DCCS but also the broader battery of behavioral measures designed to test various aspects of online response selection (Ramscar & Gitcho, 2007). For the DCCS at least, this poor performance is unlikely to be attributable to the verbal sophistication of the task: Three-year-olds can correctly answer questions about how to sort cards according to the game rules, even as the rules change, and consistently sort correctly in the first game (Zelazo, Frye, & Rapus, 1996; but see Munakata & Yerys, 2001). It is not until the beginning of the second game, when the child is handed a fresh card and asked to sort by the new rule that the difficulty emerges: Three-year-olds who appear to "know" that the rules have changed nevertheless fail to switch their responses, and continue to sort by the first rule. Children make identical errors when observing a puppet sort the cards (Jacques, Zelazo, Kirkham, & Semcesen, 1999), suggesting that the difficulty is not solely one of motor perseveration. This pattern of responding is seen up until around age 4, when children begin to consistently pass the DCCS, successfully matching the cards to the bins both before and *after* the sorting rules are changed (Zelazo et al., 2003).

Why do 3-year-olds typically fail the DCCS, while 4-year-olds pass it? What aspects of the underlying developmental process motivate this transition? One promising suggestion is that 3-year-olds' poor performance is related to the immaturity of the prefrontal cortex (PFC; Brooks et al., 2003; Deák, 2003; Diamond, 2002; Moriguchi & Hiraki, 2009; Ramscar & Gitcho, 2007). PFC has long been implicated in flexible, goal-directed behavior (Miller & Cohen, 2001), and collectively, its regions are thought to act as a "dynamic filter" that selectively maintains appropriate patterns of neural activation, while gating inappropriate or irrelevant ones (Shimamura, 2000). Moreover, in adults, PFC is known to be a critical component in a functional circuit that appears to deal with the demands of online response-conflict processing (Yeung, Botvinick, & Cohen, 2004; Yeung, Nystrom, Aronson, & Cohen, 2006).

There is strong evidence supporting the idea that prefrontal function is late in developing (Davidson, Amso, Cruess Anderson, & Diamond, 2006) and that the neural circuitry that underpins it follows a protracted, nonlinear developmental trajectory (Ramscar & Gitcho, 2007). For example, while synaptogenesis in visual and auditory cortex peaks a few months after birth, the same developments occur only much later in PFC (Huttenlocher & Dabholkar, 1997). In additional, Von Economo neurons, which appear to play important role in signaling between the areas involved in response-conflict processing, are largely absent in newborns and only reach adult levels in the 4th year (Allman, Watson, Tetreault, & Hakeem, 2005; Ramscar & Gitcho, 2007; Rueda, Rothbart, McCandliss, Saccamanno, & Posner, 2005).

As a consequence of this prolonged period of cortical immaturity, children under four exhibit patterns of impaired behavioral and cognitive control that is similar in many ways to that of patients with neurological PFC damage (Thompson-Schill, Ramscar, & Evangelia, 2009). Young children seem largely unable to act in ways that conflict with dominant prepotent responses and are highly limited in their ability to dynamically filter their behavior or their learning in accordance with the demands of a particular context or goal (Ramscar & Gitcho, 2007). On balance then, the neurological and behavioral evidence suggests that 3-year-olds' difficulties with rule switching in the DCCS stems from the slow functional development of the neural circuitry that underlies response-conflict processing.

Discrimination Learning

While under-fours appear largely incapable of resolving conflict online, there is still another "route" to success available to them in the DCCS: learning. Children can learn that the games associated with each sorting rule provide the contextual information needed to make the appropriate response and then use this (implicit) knowledge to effectively minimize response conflict that the DCCS usually evokes, thereby allowing them to pass the task. This suggestion is supported by the finding that typical patterns of prefrontal activation seen in the structurally similar Stroop task vanish when adult participants engage in the task for a sufficient amount of time, indicating that participants learn context appropriate representations that eliminate the local response conflict that they ordinarily encounter during the course of the task (Erickson et al., 2004; see also Ramscar & Yarlett, 2007; Roediger & Karpicke, 2006). However, as children typically fail the task despite the presence of contextual cues, it also seems clear that in ordinary testing circumstances, the game cues do not provide enough scaffolding to enable 3-year-olds to pass the DCCS.

To explain why these contextual cues might still matter, we first need to consider the way that responses that lead to response conflict in the DCCS could be learned and discriminated in an implicit learning paradigm. In contemporary learning theory, learning is conceptualized as a process of discriminating ever better predictive cues to the relevant events in (or *responses to*) one's environment. Learning models embody the ideas: (a) that learning is driven by "surprise" (prediction error)—which occurs whenever our expectations and reality fail to align—and (b) that learning is attenuated when expectation and reality are aligned, such that there is little remaining uncertainty about the events that will unfold in a given context (for a review, see Ramscar, Yarlett, Dye, Denny, & Thorpe, 2010). Learning serves to reduce uncertainty about how to respond in a given context, by discriminating in favor of more informative cues to responses, and discriminating against less informative cues (i.e., cues which would lead to an inappropriate response; see Rescorla, 1988). As uncertainty diminishes, and a learner's certainty about the likelihood of a given response increases in kind, learning asymptotes.

For example, in the Rescorla and Wagner (1972) learning rule, changes in the value of cues in learning are determined by a discrepancy function ($\lambda - V$), which takes the difference between expectation and reality across trials, and uses this to update expectations by modifying the associative values of the set of cues to a given response. Cue values are strengthened when an appropriate response is underpredicted, and weakened when it is overpredicted (or predicted erroneously; Kamin, 1969; Rescorla & Wagner, 1972). Importantly, associative value lost by one cue can be subsumed by other cues, resulting in a competition for relevance between cues. The outcome of this competition is determined both by positive evidence about co-occurrences between cues and predicted events (reinforcement), and negative evidence about *non*occurrences of predicted responses (error).

Accordingly, the distribution of error associated with the *set* of cues to a particular outcome critically affects the values that individual cues take on in learning. The value of each cue is shaped by the *combined* value of both the positive and negative evidence provided by that cue, *and* the value of the evidence supplied by other *competing* cues. This results in patterns of learning that are very different from those that would result if learning were limited to simple co-occurrence tracking (a common misconception of learning; see Rescorla, 1988). Competitive discrimination learning uncovers the predictive structure of the environment, allowing learners to discover the cues that best predict appropriate outcomes (Ramscar et al., 2010; Smith, Colunga, & Yoshida, 2010).

It is worth adding that while much of the impetus for the development of these models came from behavioral experiments in animals, there is now good evidence for their neurobiological basis (see Schultz, 2006, for a review). The learning rule we have described accurately predicts patterns of synaptic firing in midbrain dopamine neurons in learning tasks (Waelti, Dickinson, & Schultz, 2001) and has been productively applied to many aspects of human behavior and cognition, such as decision making, executive function, habitual learning, and response selection (Montague, Hyman, & Cohen, 2004), demonstrating considerable predictive and explanatory power.

Method

Discriminative learning offers a mechanism that can allow young children to learn to match specific responses to specific contexts. A child can learn to better respond in context by implicitly learning to discriminate the cues to an appropriate response from unhelpful or distracting cues by monitoring the error associated with them. This process will invariably highlight contextual cues (because these will be most informative about context-appropriate responses) and devalue many context-independent cues (which may provide support for inappropriate responses). In principle, the outcome of this process should be better tuned, more context-sensitive responding.

In the DCCS, the "games" associated with each sorting rule offer just the kind of contextual information children need to help them succeed at the task. If a child was to learn to attend to (assign high values to) Conjunctive cues that take into account both the game context and the relevant sort dimension (e.g., *star* + *shape game* and *red* + *color game*), and to ignore (assign low values to) any inappropriate cues, then switching ought not to be a problem. However, to learn *what* the correct cues are, children will first have to discriminate them from cues that are less relevant to this context but that nevertheless carry valuable information, either in other contexts or across contexts: in particular, the dimensional cues, color (*red*) and shape (*star*). In what follows, we present a computational model that shows formally how perseverative sorting on switch trials is a natural result of cue competition in the early parts of this process. Before detailing how this model works, however, it is worth considering the conceptual picture that underpins our simulations of learning in the DCCS, in which the process of learning the game rules is seen as a one of tuning the expectations that "sculpt attentional biases" (Colunga & Smith, 2008, p. 180).

Imagine a 3-year-old watching her brother complete the DCCS. Although she has never played before, her brother has played many times, and he always sorts correctly. As she watches him sort cards, the girl (or, more accurately, her implicit learning mechanisms, which process information in the way we describe here) tries to guess each of his upcoming moves, placing bets on the probability that he will sort one way or the other. In an attempt to optimize her success rate, she successively refines her predictions over the course of the game based on how each of her bets pays off.

Suppose the girl's brother begins the DCCS with a color sorting game. Initially, she might hedge her bets because she will not know whether the color cue (sort by red) and the shape cue (sort by star) predict the correct sort. However, because *red* always succeeds in predicting the correct sort in the color game, whereas *star* never does, she will place more and more of her chips on *red*, until she is no longer betting on *star* at all.

By the end of the first color game, she will have learned that she can maximize her returns by betting all her chips on *red* all the time. However, when the game changes—and her brother begins to sort by shape—her luck will change. She will still be placing her bets on *red*, but now *star* is proving to be the more reliable cue. In response to this, she will begin to divest some chips from *red* and put them on *star* instead. As the shape game continues, she will move more and more of her assets to *star*.

As we have described it so far, the process of reallocating chips will simply reverse itself on every switch trial: the girl will learn to bet on one cue, then another, and then back to the first. However, she can learn to predict her brother's sorts more accurately if she incorporates some contextual information into her strategy. In the DCCS, each sort takes place within the context of the shape or color game. Instead of placing her bets on only the two cues, she can hedge her bets by considering other possible cues. For example, in addition to the two-dimensional cues (*red* and *star*), she might consider placing her bets on conjunctive cues (Gluck & Bower, 1988) that integrate these dimensions with the contexts provided by the games (e.g., *color game* + *red* and *shape game* + *star*).

Because these conjunctive cues will only be present in the contexts in which they actually occur, they will not lead to losses: Each bet on the appropriate conjunctive cues will be rewarded. By contrast, the dimensional cues will only pay dividends half the time (*red* is a good bet in color games but not shape games, and vice versa for *star*). As evidence of the unreliability of the dimensional cues grows, the girl will move more and more of her chips away from them, as experience teaches her to focus her attention on the conjunctive cues instead.

A Model of Error-Driven Learning

To formally illustrate how this process leads to the pattern of perseveration seen in the DCCS, we simulated the competition between conjunctive (Contextual + Dimensional) cues representing *color game* + *red* and *shape game* + *star* and the individual cues *red* and *star* across repeated DCCS trials using the Rescorla and Wagner (1972) model of error-driven learning (Siegel & Allan, 1996; for other applications to learning in early childhood, see Ramscar et al., 2010; Ramscar, Dye, Popick, & O'Donnell-McCarthy, 2011).

In Rescorla–Wagner, learning is simulated by changes to the associative strengths between individual cues (C) and an outcome based on the results of discrete learning trials. If the presence of a cue or outcome X at time t is defined as present (X, t), and its absence as absent (X, t), then the predictive value V of a cue C_i for an outcome O after a learning event at time $t + 1$ can be stated as:

$$V_i^{t+1} = V_i^t + \Delta V_i^t, \tag{1}$$

while the change (Δ) to the predictive value of V_i^t can be defined as:

$$\Delta V_i^t = \begin{cases} 0 & \text{if ABSENT}(C_i,t) \\ \alpha_i\beta_1\left(\lambda - \sum_{\text{PRESENT}(C_j,t)} V_j\right) & \text{if PRESENT}(C_j,t)\ \&\ \text{PRESENT}(O,t) \\ \alpha_i\beta_2\left(0 - \sum_{\text{PRESENT}(C_j,t)} V_j\right) & \text{if PRESENT}(C_j,t)\ \&\ \text{ABSENT}(O,t) \end{cases} \tag{2}$$

Learning is thus governed by a discrepancy function in which λ is the total value of a predicted event O (the maximum amount of associative strength that the event can support; here it is simply set to 100, indicating that an event can be fully anticipated) and V_j is the predictive value for O given the set of cues C_j present at time t.

Importantly, there is no notion of explicit feedback in the model. Rescorla–Wagner learns—that is, adjusts cue weights—in response to the occurrence or *non*occurrence of relevant events. In this, it samples the implicit predictive structure of the learning environment. The learning rule is thus particularly well suited to capture the patterns of learning of young children, who are not yet able to selectively attend to their environment in the same manner as adults (Thompson-Schill et al., 2009).

In trials in which there is positive evidence—that is, in which expected responses *do* occur—the learning rule produces a negatively accelerated learning curve (the result of responses being better predicted, which reduces the discrepancy between what is expected and what is observed) and asymptotic learning over repeated trials (as responses become fully predicted).

In trials that result in negative evidence—that is, in which an expected response fails to occur—λ_j (the expected response) takes a value of zero because it did not occur. In such cases, the discrepancy function ($\lambda_j - V_j$) produces a negative value, resulting in a reduction in the associative strength between the cues present on that trial and the absent response *j*.

The total amount of predictive (cue) value any given outcome can support in learning is finite. Informally, this captures the idea that if predictive confidence keeps rising, it must eventually reach a point of certainty. As a result, cues *compete* with one another for relevance, and this produces learning patterns that often differ greatly from those that would arise by simply recording the correlations between cues and responses (Rescorla, 1988).

The rate of change (Δ) at *t* is determined by two factors: the overall learning rate β (where $0 \leq \beta \leq 1$) and the individual saliency of cues, α_i (where $0 \leq \alpha \leq 1$). In the simulation we conducted, λ = 100% (when *j* was present in a trial) or 0% (when *j* was not present in a trial), $\alpha_i = 1$ and $\beta_j = 0.3$. By setting the α_i parameter to 1, we effectively eliminated its influence on our simulations. These parameter settings were held constant throughout all the simulations reported below.

Simulating Competitive Learning in the DCCS

Simulation 1 models competitive learning in the DCCS over three successive games. The simulation predicts that if a child observes the correct sort being made on every trial, she will learn to value conjunctive (context + dimension) cues, allowing her to successfully switch rules in the DCCS. The initial implementation makes three straightforward assumptions:

1. The appropriate sorting response is always made; that is, the child learns in response to what *ought* to happen.
2. That the dimensional cues will each predict the correct sort on half the trials; for example, *red* will predict the correct sort in the color game but not the shape game.
3. That the conjunctive cues will always predict the correct sort; for example, *color game + red* will predict the correct sort in color game trials and will not be present on any trial in which shape is matched (i.e., shape game trials).

The model thus captures the state of a child who is given the opportunity to sort the cards, and either (a) correctly sorts each one on each trial, despite other possible responses, or (b) observes another participant make the correct response on every trial. Given this assumption, it is important to keep in mind the (somewhat counterintuitive) idea that even when the model predicts on one trial that the child will give an incorrect response, learning on the next trial proceeds *as if she had not*. The simulation thus makes predictions about what the child's response behavior will look like at any given point along the learning trajectory, after having observed—or made—a certain number of correct sorts and switches. (This assumption is discussed in further detail in the following section.)

Training began with a color game (Trial 1), followed by a shape game (appearing in Trial 7). To reflect children's experience in DCCS tests, the model was trained for six trials in each game, with training alternating between games.

In the first DCCS game shown in Figure 2, the dimensional cue *red* and the conjunctive cue *color game + red* gain in associative value. At the rule switch, the model pegs *red* as the favored cue to

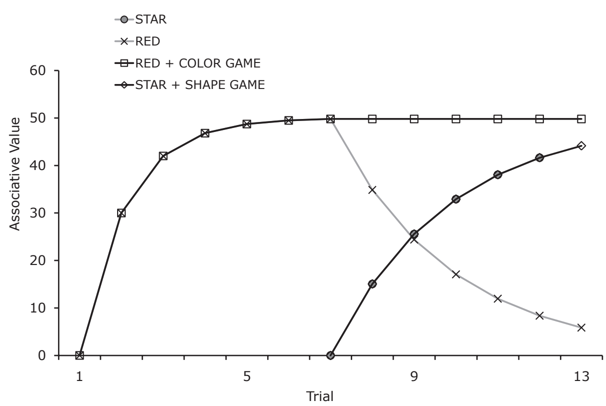

Figure 2. Rescorla–Wagner simulation of cue competition in two Dimensional Change Card Sort (DCCS) trials. Each line represents the association between a given cue and the correct response. Trials 1–6 represent a color game and Trials 7–12 a shape game (the six trials in each DCCS game correspond to six trials in the simulation). The value of *red* increases during the color game trials, meaning that in the first trial of the shape game, *red* is more highly valued as a sorting cue than either *star* or the conjunctive cue *star + shape-game*, leading to perseverative responding.

sorting. However, following the switch, *red* begins to lose associative value as it now fails to predict the correct sort. Over time, cue competition causes *red* to be effectively dissociated from correct sorting in this situation. The error generated by *red* over shape trials (where *red* continues to predict a color sort) results in a gradual shift of associative value from the dimensional cue *red* to the conjunctive cue *color game + red*, which does not lose value in this way because it is not present on shape trials.

It is important to note here that even though the dimensional cue (*red*) and the conjunctive cue (*color game + red*) co-occur with correct sorts with exactly the same frequency—that is, on half of all trials—the dimensional cues also co-occur with incorrect sorts, whereas the conjunctive cues do not. This means that the distribution of error (the combined evidence) will favor the conjunctive cues over the long run. Thus, the child will learn to ignore the unreliable cues over switch trials, thereby diminishing any bias toward perseverative responding.

This pattern of learning is clearly evident in the simulation (Figure 3): Depending on the game being played, the predictive value of a given conjunctive cue will either flatline if it is not present (e.g., *red + color game* on shape game trials, Trials 7–12) or increase if it is (e.g., *star + shape game* on shape game trials, Trials 19–24). Meanwhile, the value of each of the isolated dimensional cues *red* and *star* will continue to rise and fall depending on whether they are being reinforced in the game being played (e.g., *star* will rise on shape game trials and fall on color game trials). However, the

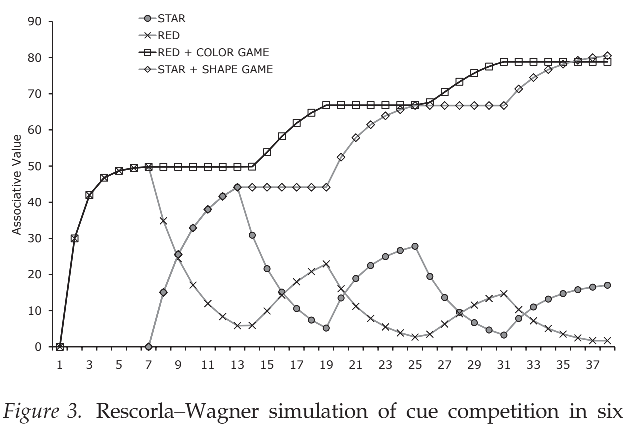

Figure 3. Rescorla–Wagner simulation of cue competition in six DCCS games. Each peak can be taken to represent a rule switch. The simulation assumes that the correct sorts are seen in each Trial. Prediction error (negative evidence) causes the individual cues to lose associative value to the conjunctive (context + dimension) cues over successive trials.

overall trend for the dimensional cues is downward, as they will bleed associative value to the conjunctive cues. Indeed, after the second switch trial (Trial 13), the value of the dimensional cues will fall below that of the contextual cues, never to recover.

Behaviorally, the model suggests that perseverative sorting will occur whenever cues supporting the incorrect response are valued *over* those supporting the correct response, and, response conflict will occur whenever cues supporting the wrong response are strong relative to those supporting the correct response, even if they are not necessarily preferred. Given this, we can see that the model predicts a perseverative response on the first switch Trial (7): The value of the *red* cue far outstrips that of the *star* cue. However, the balance of power soon shifts. If we now turn to the third switch trial (19), we can see that while the value of the *red* cue outstrips that of the *star* cue, the value of the *star + shape game* cue far outstrips both dimensional cues.

Compatible Accounts

In Simulation 1, initial perseveration resulted from children failing to discriminate the contextual cues appropriate for correct sorting from competing cues that lead to error. This formal account of the way learning influences performance in the DCCS is compatible with a number of theoretical proposals about the factors that affect children's ability to successfully complete the task. For example, it has been suggested that perseverative responding in the task arises from attentional inertia: that once 3-year-olds have "focused their attention on a particular dimension, their attention gets stuck there, and they have extreme difficulty redirecting it" (Kirkham et al., 2003, p. 451). The problem is thought to be one of perspective: Once a child has grasped one sort dimension, she has trouble flipping her "mental focus" and thinking about it along the other (Diamond & Kirkham, 2005; Kloo & Perner, 2005). Our model of learning formalizes the idea of "mental focus," allowing the way attention shifts as a function of experience to be more precisely accounted for (see also Colunga & Smith, 2005).

Similarly, our account shares certain similarities with cognitive complexity and control theory (CCC; Zelazo et al., 2003), which suggests that successfully passing the DCCS requires that a child formulate a higher order "if . . . if . . . then" rule that integrates two conflicting pairs of rules within a propositional structure (e.g., "if a1, then if b1, then c1" while "if a2, then if b1, then c2"). Our model recharacterizes these rules in terms that allow us to examine the degree to which children's learning about them is implicit and contextual (Towse, Redbond, Houston-Price, & Cook, 2000). If children are given the DCCS without the aid of contextual training, they will still need to be able to resolve response conflict to pass it, as CCC theory suggests (indeed, most theories agree on this; see Diamond et al., 2005). However, formalizing the way that rules and representations are learned not only explains why children ordinarily fail the DCCS but also explains why children are able to respond flexibly in some tasks, even as they fail to do so in others. Furthermore, it allows us to predict the kind of learning experiences that should improve children's performance on the task.

Generalizing From Naming to Sorting

Simulation 1 (Figures 2 and 3) shows how contextual learning can serve to reduce response conflict in the standard DCCS, thereby offering an alternate route to success in the DCCS. If children are trained to associate sorting by shape with a "shape game" and sorting by color with a "color game," they can eliminate the response conflict normally associated with the DCCS by learning to value context-dependent cues to sorting (*red star card* + *color game* = *sort by red*) over context-independent cues, which frequently lead to errors.

This simulation provides a solution to a puzzle in the current literature: although under-fours routinely fail the DCCS, they can often reliably *name* the dimensions of the cards, even when they fail to correctly sort them (Kirkham et al., 2003). While on its face this finding may seem surprising, it need not be. Children have far more experience with naming than with sorting, and they get far more opportunities to watch other people use names in context. Given that children regularly hear objects referred to in terms of their shape and color, and so on, they are likely to have done a great deal of learning (including observational learning) about pairing shape words with shape questions and color words with color questions and so forth. (Landau, Smith, & Jones, 1988; Sandhofer & Smith, 1999). When it comes to a card-*naming* task, this kind of contextual discrimination learning will have diminished the degree to which color words suggest themselves as responses to shape-naming questions, thereby minimizing response conflict and enabling children to flexibly name the appropriate dimensions of the cards according to context.

Simulation 1 thus suggests that learning context-*dependent* cues could enable children to effectively bypass response conflict in the DCCS. Importantly, however, our computational simulation of *sorting* assumed that learners made the correct response throughout. Given that over the normal course of the DCCS, under-fours make frequent sorting errors, it is highly unlikely that a child playing the DCCS will actually experience the ideal learning conditions assumed in our simulation.

However, a considerable body of evidence supports the idea that learning generalizes across contexts (Gluck, 1991; Shepard, 1987). Within a discriminative learning framework, generalization results from a *lack* of discrimination. For example, in Simulation 1, when the initially poorly discriminated cues support perseverative responding, these perseverative responses are a form of (over-) generalization of learning from earlier trials (see Ramscar & Yarlett, 2007; Ramscar et al., 2010). Accordingly, if a child was to learn a particular configuration of cues and how they support (or failed to support) the correct response in one type of game context, what she learned about that context should transfer to other, similar contexts. This suggests that children's performance on the DCCS could be successfully enhanced by providing them with an opportunity to learn about a *color shape card* + *color game* context (or *color shape card* + *shape game* context) while playing a game at which they could succeed on every trial, even after a rule switch.

Because 3-year-olds can reliably and flexibly name the dimensions of the cards in the DCCS even when they fail to correctly sort by them, a naming game provides precisely this kind of learning opportunity. The results of Simulation 1 suggest that asking children to *name* the conflicting dimensions of the cards in the context of the games should lead to them learning the context-dependent cue configurations that lead to the correct naming responses. If children then generalize this knowledge to their *sorting* responses—using the same conjunctive cues they learned in the labeling game—this should substantially reduce the response conflict they encounter in the DCCS, markedly improving their chances for successful rule switching.

Simulating Noncompetitive Learning in the DCCS

Our simulation predicts that competitive discrimination learning about the game contexts in a labeling context will help children's performance on the DCCS. However, not all labeling tasks promote competitive learning. Depending upon the temporal sequence in which objects and their labels are encountered, learning can be made more or less competitive (Ramscar et al., 2010), thus allowing us to isolate the effects of competitive learning on DCCS performance.

In the naming game, presenting cards and then labeling them (feature-to-label, or FL-learning; Ramscar et al., 2010; Ramscar et al., 2011) allows the features of the cards and game contexts to serve as *cues* to labeling *responses*. This facilitates competitive learning, in which the various cues "compete" for relevance in predicting the appropriate labeling response, losing associativity to other, more reliable cues, and gaining associativity from other, less reliable cues. Cue competition thus allows a learner to discriminate the cues that most *reliably* predict the correct response (such as the conjunctive cues) from unreliable cues that result in error (such as the dimensional cues; see Figure 3).

Importantly, simply reversing the predictive (temporal) relation between labels and features can have a dramatic impact on cue competition. Whereas FL-learning involves predicting an individual label from a complex, multidimensional set of cues (the features of the cards), label-to-feature (or LF-) learning involves just the opposite: predicting a complex, multidimensional perceptual item (the card) from an individual cue (a symbolic label). Because verbal labels are sparse, they do not provide a learner with the raw materials required for discrimination learning (a matrix of feature dimensions that covary across learning contexts; Ramscar et al., 2010). In effect, LF-learning provides learners with just a single cue (a label) to learn from, thereby eliminating cue competition (if there are no other cues present, there are no other cues to lose associative value to). As a result, when a label acts as a cue, its associative value rises and falls according to whether it successfully or unsuccessfully predicts a given dimensional feature, effectively tracking the frequency with which labels and card features co-occur (for a formal analysis and extended discussion of the effects of information structure on learning, see Ramscar et al., 2010).

To illustrate the effect this manipulation has on discrimination learning, Figure 4a depicts a variant of our earlier simulation in which cards depicting objects are alternately presented and then named in color and shape-naming games. As can be seen, FL-learning results in the devaluing of the individual dimensional cues (*red* and *star*), and learning of the conjunctive cues ("rules") appropriate to each naming game. However, in LF-learning (Figure 4b), because the cards' features do not compete as cues,

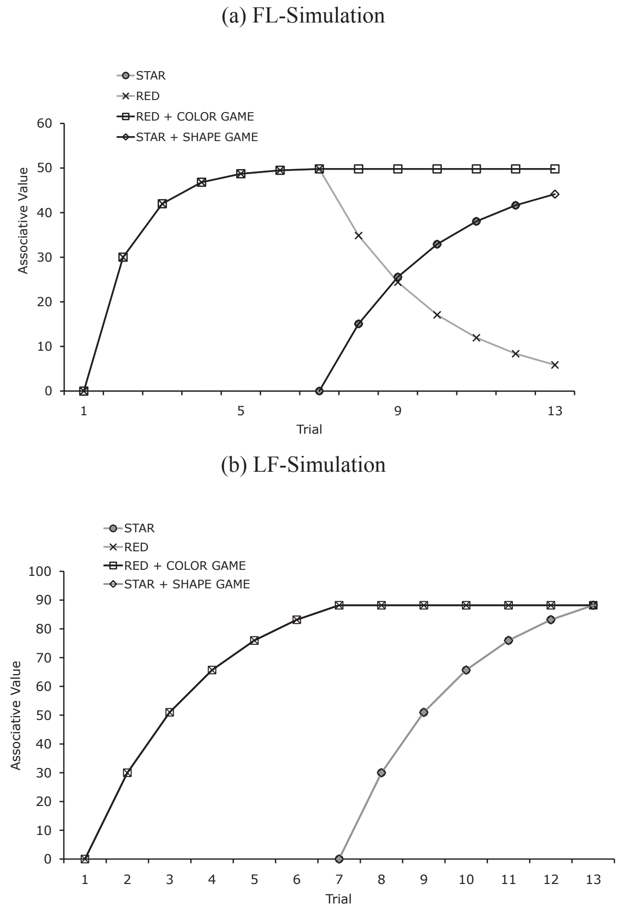

Figure 4. Simulations of cue competition in naming games in feature-to-label (FL) learning (left panel) and label-to-feature (LF) learning (right panel). Each line represents the association between a given cue and the correct response. Trials 1–6 represent a color-naming game, trials 7–12 a shape-naming game. FL-learning causes the *devaluing* of the individual dimensional cues (*red* and *star*), whereas LF-learning results in a uniform pattern of association between the dimensional and contextual cues.

the associative value accorded to the conjunctive cues will simply track the *frequency* with which dimensional features are present after labeling in each labeling game context.

As the dimensions occur with equal frequency (every card has both a color and shape), LF-learning will fail to discriminate between them. This means that although there will be learning about the game contexts in LF-presentations, LF-learning will provide identical evidence for the conflicting cues (Figure 4b). Thus, although a child asked to switch in a sorting task after LF-training will not face the same predicament as a child asked to switch dimensions in an ordinary DCCS task (where the available evidence actually *favors* perseveration; see Figure 2), she will still face the problem of selecting among competing responses.

The different outcomes promoted by FL- and LF-training allow for our predictions about both the influence of discrimination learning in context, and the mechanisms that give rise to it, to be formalized and tested in detail. Figure 5 presents the results of two simulations that model performance on two DCCS tests. The simulations were given the same information—structured in the same way—as the children in our training experiment. In the first simulation, test performance was assessed following two FL naming games, and in the second simulation, following two LF naming games. In the games, two sets of cards that varied on two dimensions (shape and color) were presented and labeled in either an FL- or LF-sequence. In the color game, the color dimension was labeled, and in the shape game, shape was labeled.

Although the temporal sequence in which cards were presented and labeled varied across the two training conditions, the sequence of presentation and elicitation in DCCS testing was the same across conditions (children in all three experimental conditions were tested on the standard DCCS). Notably, because cards are presented prior to sorting in the DCCS, the information structure in test trials resembled the FL-training trials.

As can be seen in the FL-trained simulation (Figure 5a), by the first switch in sorting at Trial 19, the associative value of the appropriate conjunctive cue (*star + shape game*) is appreciably stronger than the competitor *red* cue, and by the second switch at Trial 31, the gap between reliable cue and competitor has more than doubled. This suggests that FL-training should do much to minimize response conflict, helping most children sort correctly across both tests.

By contrast, LF-training is far less helpful, at least initially (Figure 5b). At the critical first switch at Trial 19, the associative value of the conjunctive cue (*star + shape game*) is the same as that of the competitor *red* cue, which means that both of two conflicting responses are supported. However, by the second switch at Trial 31, there is a significant gap between the appropriate conjunctive cue (*red + color game*) and the competitor (*star*), which predicts reliable sorting. At first glance, this would seem to suggest that while many LF-learners should fail on the first switch, they may well pass on the second. Tempting as this interpretation is, it is also misleading, because the simulation assumed that *the correct sort is made throughout*. However, in the context of the experiment, this is only necessarily true of the

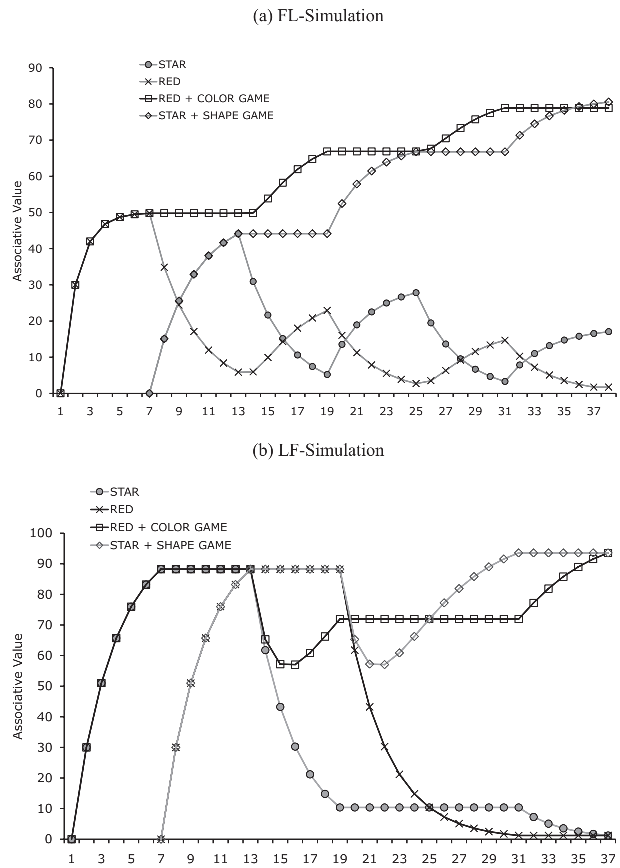

Figure 5. Simulations of Dimensional Change Card Sort (DCCS) testing following feature-to-label (FL) training (a), and label-to-feature (LF) training (b). In both simulations, Trials 1–12 depict training (naming games), while testing occurs over Trials 13–24 (first DCCS test) and 25–36 (second DCCS test). The important switch trials in testing occur on Trials 19 and 31. Breaking it down further, Trials 1–6 represent a color *naming* game, Trials 7–12 a shape-*naming* game, Trials 13–18 a color *sorting* game, and Trials 19–24 a shape *sorting* game. Because the initial sort dimension is reversed in the second DCCS, Trials 25–30 also represent a shape *sorting* game, with another switch back to color *sorting* in Trials 31–36. In interpreting these figures, and comparing them with our empirical results, it is important to note that the curves of the model represent the learned strength of a response, not the probability of a response. In situations in which there is response conflict, young children—unlike adults—tend to maximize, going with the most likely response every time, rather than probability match, alternating between responses based on their respective probabilities (Hudson-Kam & Newport, 2005).

training trials (1–12), in which children are labeling the cards. If, in testing, subjects are unable to resolve the response conflict on the first switch trial—and end up sorting incorrectly over Trials 19–24—they will gain none of the benefits from learning on subsequent trials that the simulation suggests.

Figure 5 thus confirms our prediction that FL-structured naming games ought to provide children with considerable help in avoiding perseverative responding (Figure 5a), whereas LF-training will only be of help *if* children can overcome the response conflict that arises on the critical first switch trial (Figure 5b).

Training Experiment

To test these ideas, we examined the effect that providing discrimination training to children would have on their performance in the DCCS.

Method

Participants. A total of 47 English-speaking children between 3 and 4 years old (M = 3 years 6.8 months) participated in this study, with a near-even balance between genders (25 girls and 22 boys). Participants were recruited from Stanford and the surrounding community.

Methods and materials. Two groups of children received either FL or LF training on the cards (Ramscar et al., 2010), before completing standard DCCS tasks (Zelazo, 2006). A control group was tested on the DCCS *without* training. All three groups were age and gender balanced.

In the two training conditions, children were introduced to shape and color games prior to the DCCS using 12 sorting cards (six yellow flowers and six green boats). In the FL-condition, children were first told, "In the shape game, we name the different shapes on these cards." The experimenter then presented the first card to the child and asked the child to label it. After children correctly labeled the first 6 of the 12 cards, the experimenter said, "We're going to play the color game. In the color game, we are going to say what colors are on these cards." Children then labeled the remaining six cards in the new game.

Whereas children in the FL condition saw the card and labeled it, children in the LF condition were asked to say the label *before* seeing the card. They were told, "In the shape game, we name the different shapes on these cards. The first card is going to be a flower—can you say 'flower'?" The experimenter showed the card to the child only after the child had repeated the label. The basic structure of the LF-training was the same as the FL-training: naming six cards by one dimension and then switching to the other dimension.

The two training groups and the control group then completed two standard DCCS tasks. The first DCCS task used the same sort cards as those in training (yellow flowers and green boats), whereas the second DCCS task used new cards, with different shapes and colors (blue trucks and red stars). The second DCCS task thus tested children's ability to generalize what they had learned, as it required them to sort by newly introduced colors in the context of the color game, rather than simply select yellow versus green.

The first testing dimension (shape or color) was counterbalanced across children, and children were required to correctly sort six cards in the preswitch. Once a child had done this, the sorting dimension was switched. Exactly six cards were sorted in the postswitch test. Before each trial, the subjects were either reminded of the current game's rules or asked to answer "knowledge questions," such as, "Where do the flowers go? Where do the boats go?" Children were given *no feedback* about their sorting of the cards. After the first DCCS task, children completed the second DCCS task with new cards, and the switching dimension was reversed from the first task.

Results

All children in the two training conditions correctly labeled the cards. In testing, children were considered to have "passed" the DCCS task if they sorted at least five of the six postswitch cards correctly. Sixty-nine percent of the FL-trained children passed the first DCCS task and 75% the second. By contrast, 33% of the LF-trained children passed the first task and 40% the second. Nineteen percent of the control children passed in each test (Figure 6).

Chi-square tests (the usual method of analysis for the DCCS) revealed significantly higher rates of successful rule switching in the FL-condition (11/16

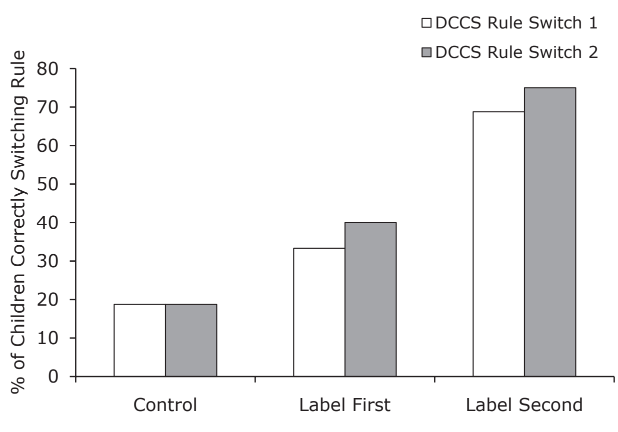

Figure 6. Percentage of children switching correctly in the first and second Dimensional Change Card Sort tasks by condition.

children succeeded) compared to 5/15 in the LF-condition, $\chi^2(1, N = 31) = 9.7$, $p = .005$; second switch, FL-condition (12/16 succeeded), LF-condition 6/15, $\chi^2(1, N = 31) = 17.0$, $p = .001$. Against the control group (3/16 in each), the comparisons with the FL-condition were, first switch, $\chi^2(1, N = 32) = 14.9$, $p = .001$; second switch, $\chi^2(1, N = 32) = 23.7$, $p = .001$. (These findings were confirmed in an equivalent series of parametric analyses.)

A series of one-sample *t* tests revealed that, overall, the FL-trained children had managed to sort correctly at above chance rates in the switch trials, $t(31) = 2.666$, $p < .02$, confirming that their training had enabled them to overcome the perseverative bias induced in the DCCS, while the baseline children sorted correctly at below chance rates, $t(31) = -3.161$, $p < .005$, confirming that they had perseverated on their prior sorts in the switch trials, just as previous studies of the DCCS in children of this age suggested they would.

Taken as a whole, the LF-trained children performed at chance in the switch trials, $t(29) = -0.586$, $p = .56$. This was consistent with the predictions of our LF-trained simulation, which suggested that while these children would learn something about the contextual rules, their use of the rules would be subject to considerable interference from the irrelevant cues (Figure 5b). However, inspection of the LF-trained children's results revealed that individually, the children's performance was not at chance; rather, some children were consistently sorting correctly, whereas others were perseverating.

Given that our maturational account suggests that response-conflict processing is acquired gradually, it follows that we should expect that, on average, LF-training ought to be of more benefit to the older children in our sample (older 3-year-olds are more likely to be in a position to process the limited conflict that remains after training than younger 3-year-olds). This should be especially true in comparison to the baseline children, who will still face a perseverative bias after the first sort game. In contrast, FL-learners ought to be *less* affected by age, as FL-learning leads to more distinct representations of the sort rules than LF (Figure 5).

To examine the extent to which these predictions were borne out in our results, we performed a median split by age on the groups of children in the study (for a breakdown of these groups, see Table 1), subjecting children's sorting to analysis in a 3 (training condition) × 2 (older vs. younger) × 2 (first vs. second DCCS trial) analysis of variance. This revealed that while there was no effect of the first versus the second DCCS test on children's performance, $F(1, 94) = 0.120$, $p > .7$, both training, $F(2, 93) = 9.199$, $p < .001$, and age, $F(1, 94) = 14.458$, $p < .001$, did have a significant effect on sorting. Furthermore, consistent with the predictions of our model, there was a significant interaction between age and training in children's performance, $F(2, 93) = 3.229$, $p < .05$ (see Figure 7). As expected, while the discrimination learning brought about by FL-training helped children pass the DCCS across the full age spectrum tested in our experiment, the overall reduction in response conflict in LF-training was *only* of benefit to the older children tested.

Notably, there may be other possible explanations for why the LF-trained children failed to benefit from training in the same way as their FL-trained counterparts. In particular, the FL-structure of testing may have made it easier for participants trained FL to generalize their knowledge to the sorting task. Given the design of the standard DCCS, this possibility cannot be completely ruled out. However, it should be noted that in previous studies of number and color learning in children, and relative pitch and complex-category learning in adults, FL-trained participants consistently outperform LF-trained participants, even when testing is structured LF (Ramscar et al., 2010; Ramscar et al., 2011). The consistency of these temporal ordering effects across the visual and auditory domains, and

Table 1
Mean Age and Population Size of the Groups in the Analysis of Age on Training and Performance

		FL	LF	Control
Younger	Mean age	3 years 2.5 months	3 years 2.5 months	3 years 2.3 months
	N	8	8	8
Older	Mean age	3 years 8.2 months	3 years 8.1 months	3 years 8.2 months
	N	8	7	8

Note. FL = feature-to-label; LF = label-to-feature.

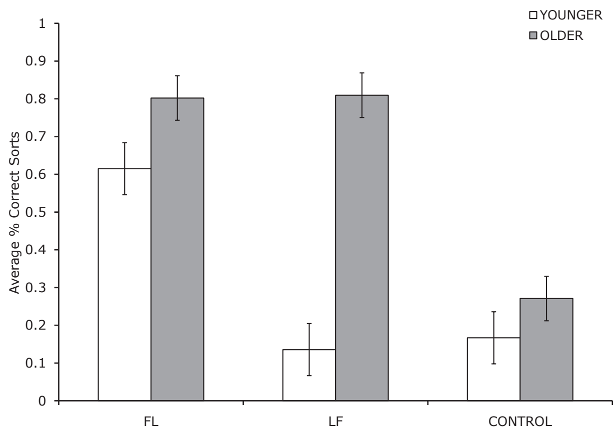

Figure 7. Percentage of correct sorts by the older and younger children by training condition. Error bars are SEM.

across a diverse array of tasks, argues for the consistency of the underlying learning mechanism.

Discussion

We predicted that playing naming games that promoted discrimination learning about the sorting rules used in the DCCS would improve children's performance on this task. Consistent with these predictions, younger children given training that was designed to promote contextual discrimination learning were consistently able to switch rules and pass the DCCS. When a less effective training program was administered, only the older children derived any benefit; children given the task without any prior training failed as expected. Taken together, these findings are consistent with our suggestion that the mind has multiple ways of dealing with conflicting response demands, and with evidence that online processing of response conflict appears to develop over the course of childhood (Rueda et al., 2005) and may thus be largely unavailable to children under four (Ramscar & Gitcho, 2007; Thompson-Schill et al., 2009). That under-fours are often able to behave in ways that are flexible and contextually appropriate appears to be due to discrimination learning. As the results of the FL-trained children show, under-fours can match their behavior to context in remarkably subtle and sensitive ways, but only once discrimination learning has enabled them to do so (for comparable effects of training in "theory of mind" tasks, see Amsterlaw & Wellman, 2006; Slaughter & Gopnik, 1996).

Applying a Dual Route Account to Other Findings

Our results and analysis suggest that two factors influence the way under-fours deal with potentially conflicting task demands, such as those imposed by the DCCS, namely, in terms of discrimination learning on one hand, and response-conflict processing on the other. There are at least half a dozen theoretical accounts of the effect that response conflict has on DCCS performance. However, given under-4s' overall weakness when it comes to response-conflict processing, it seems clear that an increased understanding of the way that children learn as they perform the task, and the contexts in which this learning occurs, can offer insight into a range of theoretical debates in the literature.

For example, in recent years, there has been much interest in "negative priming" (NP) effects in DCCS tasks. In a classic NP study, Zelazo et al. (2003) ran two variants of the DCCS: a partial change version and a total change version. In the partial change task, the dimension that was irrelevant to sorting in the preswitch was changed in the postswitch (e.g., children might sort *red* rabbits by shape during the preswitch and *yellow* rabbits by color during the postswitch). In the total change task, by contrast, both dimensions changed from preswitch to postswitch (e.g., children might sort *red rabbits* by shape during the preswitch and *blue trucks* by color in the post-switch). Zelazo et al. found that while most 3-year-olds continued to perseverate in the partial change task, most passed in the total change task. Zelazo and colleagues attributed these results to negative priming—that is, to the "disruption or slowing of a response to a stimulus that has previously been ignored" (Müller et al., 2006)—and concluded that to successfully sort during the preswitch trials, children needed to "suppress" attention to the irrelevant dimension, a tactic that then proved troublesome when that dimension become relevant again in the postswitch.

This theoretical stance is both in keeping with our suggestion that perseverative responding in the postswitch is the result of how dimensional cues are devalued in the preswitch and, too, with our suggestion that 3-year-olds generally lack the ability to process response conflict online. However, framing the different "values" of children's representations in terms of formal learning models, allows us to give a more nuanced explanatory account of both how performance develops over learning trials and how performance will be affected by various changes in the task paradigm.

Take the partial change version of the DCCS: In the preswitch shape game, a child learns that cards are sorted by *rabbit* but not by *red*; then, in the postswitch, a child is expected to sort *yellow rabbits* by color. In this version of the task, the child learns on the preswitch that the dimensional cue *rabbit* is a strong predictor of a correct sort. Then on the postswitch, rabbit competes with a new cue—*yellow*—which has no predictive value whatsoever, as it has not been learned about; indeed, if anything has been learned about *yellow*, it is that it is not an informative cue to sorting. That children perseverate, and continue to sort by *rabbits*, is readily captured by a learning account.

The results of the total change version are similarly predictable: In the preswitch, the child learns about a specific set of cues and their outcomes (red rabbits sorted by shape); whereas, in the postswitch, the child learns about an entirely new set of cues and their outcomes (blue trucks sorted by color). While we would expect some form of cue generalization even after one switch, and therefore some learning about color—which may explain why fully one fourth of 3-year-olds still do not pass this version—we would also expect that there should be much *less* interference when the sets of cues change completely, and correspondingly higher overall pass rates.

Separating out the influences of learning and response conflict can reconcile a number of other notable findings regarding the various factors that improve 3-year-olds' performance on DCCS tasks. For example, having the child—rather than the experimenter—label the card to be sorted (Towse et al., 2000) facilitates discrimination learning, whereas having the bins marked by cartoon characters (who desire certain cards) instead of target cards (with conflicting dimensions; Perner & Lang, 2002) reduces the response conflict inherent in the task. Likewise, separating out the relevant dimensions (color and shape) and having the child characterize them in relation to different objects (Diamond et al., 2005; Kloo & Perner, 2005) does both: reducing the task demands, while facilitating learning.

Considering both of these issues in tandem may help explain why even 4-year-olds encounter difficulty with flexible switching when strong situational cues are absent (Blaye, Paour, Bernard-Peyron, & Bonthoux, 2006), while even some of the youngest 3-year-olds in our experiment were able to pass the DCCS, given FL-training. In addition, the marked split between the older and younger children in our LF condition is a telling illustration of how cognitive maturity and learning interact in performance. These findings underscore the idea that while changes in response conflict processing around age 4 noticeably begin to facilitate flexible switching, the process of change is gradual, and subject to an extended developmental timetable (Anderson, 2002; Davidson et al., 2006; Rueda et al., 2005).

The Advantages of Cognitive Immaturity

In all of this, one final question remains: Why is the ability to dynamically process response conflict comparatively weak in young children? What possible benefit might this confer? We suggest that delayed cognitive maturation may actually be advantageous to children when it comes to learning behavioral conventions (Ramscar & Gitcho, 2007; Thompson-Schill et al., 2009). Both linguistic and social knowledge are, in essence, conventional. For symbols (such as words) to have "symbolic value," their values must be both conventionalized and internalized, and it may be that inflexible learning is actually useful in this regard (Ramscar et al., 2010).

The question of *how* whole communities come to share symbolic conventions is rarely posed, let alone answered. We propose that if young learners are unable to actively select what they attend to during the course of learning, then they will tend to sample the environment in much the same way: Given similar input to learn from, they will acquire similar representations of that input. If, on the other hand, those same learners were able to choose what they attended to in learning, the potential for conventions to be acquired in this way would decrease dramatically. It may be that the less children are able to direct their attention in learning, the more what they learn will be shaped by the statistical regularities of their physical, social, and linguistic environments (see also Hudson-Kam & Newport, 2005; Singleton & Newport, 2004). It follows, then, that poor performance on the DCCS may be part of the price children pay for their extraordinary abilities when it comes to acquiring social and linguistic conventions.

Finally, the ability of under-fours to "hide" their cognitive inflexibility, such that it requires careful experimentation to be revealed, and the speed with which children in our experiment were able to utilize contextual cues to succeed at the DCCS, may have important implications for our understanding of everyday cognitive processing in adults. The evidence presented here suggests that young children make extensive use of context in their interactions with the world. It may be that much of adult cognition works in the same way and is far more contextually embedded than is often supposed.

References

Allman, J. M., Watson, K. K., Tetreault, N. A., & Hakeem, A. Y. (2005). Intuition and autism: A possible role for Von Economo neurons. *Trends in Cognitive Science*, *9*, 367–373.

Amsterlaw, J., & Wellman, H. (2006). Theories of mind in transition: A microgenetic study of the development of false belief understanding. *Journal of Cognition and Development*, *7*, 139–172.

Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychology*, *8*, 71–82.

Blaye, A., Paour, J. P., Bernard-Peyron, V., & Bonthoux, F. (2006). Categorical flexibility in children: Distinguishing response flexibility from conceptual flexibility: The protracted development of taxonomic representations. *European Journal of Developmental Psychology*, *3*, 163–188.

Brooks, P. J., Hanauer, J. B., Padowska, B., & Rosman, H. (2003). The role of selective attention in preschoolers' rule use in a novel dimensional card sort. *Cognitive Development*, *18*, 195–215.

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development*, *72*, 1032–1053.

Cepeda, N. J., & Munakata, Y. (2007). Why do children perseverate when they seem to know better: Graded working memory, or directed inhibition? *Psychonomic Bulletin & Review*, *14*, 1058–1065.

Colunga, E., & Smith, L. B. (2005). From the lexicon to expectations about kinds: A role for associative learning. *Psychological Review*, *112*, 347–382.

Colunga, E., & Smith, L. B. (2008). Flexibility and variability: Essential to human cognition and the study of human cognition. *New Ideas in Psychology*, *26*, 174–192.

Davidson, M. C., Amso, D., Cruess Anderson, L., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, *44*, 2037–2078.

Deák, G. O. (2003). The development of cognitive flexibility and language abilities. In R. Kail (Ed.), *Advances in child development and behavior* (Vol. 31, pp. 271–327). San Diego, CA: Academic Press.

Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy, and biochemistry. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function* (pp. 466–503). London, UK: Oxford University Press.

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*, 689–729.

Diamond, A., & Kirkham, N. Z. (2005). Not quite as grown-up as we like to think: Parallels between cognition in childhood and adulthood. *Psychological Science*, *16*, 291–297.

Erickson, K. I., Milham, M. P., Colcombe, S. J., Kramer, A. F., Banich, M. T., Webb, A., et al. (2004). Behavioral conflict, anterior cingulate cortex, and experiment duration: Implications of diverging data. *Human Brain Mapping*, *21*, 98–107.

Gluck, M. A. (1991). Stimulus generalization and representation in adaptive network models of category learning. *Psychological Science*, *2*, 50–55.

Gluck, M., & Bower, G. (1988). Evaluating an adaptive network model of human learning. *Journal of Memory and Language*, *27*, 166–195.

Hanania, R., & Smith, L. B. (2010). Selective attention and attention switching. *Developmental Science*, *13*, 622–635.

Hudson-Kam, C. L., & Newport, E. L. (2005). Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development*, *1*, 151–195.

Huttenlocher, P. R., & Dabholkar, A. S. (1997). Regional differences in synaptogenesis in human cerebral cortex. *Journal of Comparative Neurology*, *387*, 167–178.

Jacques, S., Zelazo, P. D., Kirkham, N. Z., & Semcesen, T. K. (1999). Rule selection versus rule execution in preschoolers: An error-detection approach. *Developmental Psychology*, *35*, 770–780.

Kamin, L. J. (1969). Predictability, surprise, attention, and conditioning. In B. Campbell & R. Church (Eds.), *Punishment and aversive behaviour*. New York: Appleton-Century-Crofts.

Kirkham, N. Z., Cruess, L., & Diamond, A. (2003). Helping children apply their knowledge to their behavior on a dimension-switching task. *Developmental Science*, *6*, 449–467.

Kloo, D., & Perner, J. (2005). Disentangling dimensions in the Dimensional Change Card Sorting task. *Developmental Science*, *8*, 44–56.

Kloo, D., Perner, J., Kerschhuber, A., Dabernig, S., & Aichhorn, M. (2008). Sorting between dimensions: Conditions of cognitive flexibility in preschoolers. *Journal of Experimental Child Psychology*, *100*, 115–134.

Landau, B., Smith, L. B., & Jones, S. (1988). The importance of shape in early lexical learning. *Cognitive Development*, *3*, 299–321.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167–202.

Montague, P. R., Hyman, S. E., & Cohen, J. D. (2004). Computational roles for dopamine in behavioural control. *Nature*, *431*, 760–767.

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

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

Müller, U., Dick, A. S., Gela, K., Overton, W. F., & Zelazo, P. D. (2006). The role of negative priming in the Dimensional Change Card Sort task. *Child Development*, *77*, 395–412.

Munakata, Y., & Yerys, B. E. (2001). All together now: When dissociations between knowledge and action disappear. *Psychological Science*, *12*, 335–337.

Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2010). Broca's area and language processing: Evidence for the cognitive control connection. *Language and Linguistics Compass*, *4*, 906–924.

Perner, J., & Lang, B. (2002). What causes 3-year olds' difficulty on the Dimensional Change Card Sorting task? *Infant & Child Development*, *11*, 93–105.

Perner, J., Stummer, S., & Lang, B. (1999). Executive functions and theory of mind: Cognitive complexity or functional dependence? In P. D. Zelazo, J. W. Astington, & D. R. Olson (Eds.), *Developing theories of intention: Social understanding and self-control* (pp. 133–152). Mahwah, NJ: Erlbaum.

Ramscar, M., Dye, M., Popick, H., & O'Donnell-McCarthy, F. (2011). The enigma of number: Why children find the meanings of even small number words hard to learn and how we can help them do better. *PLoS ONE*, e22501.

Ramscar, M., & Gitcho, N. (2007). Developmental change and the nature of learning in childhood. *Trends In Cognitive Science*, *11*, 274–279.

Ramscar, M., & Yarlett, D. (2007). Linguistic self-correction in the absence of feedback: A new approach to the logical problem of language acquisition. *Cognitive Science*, *31*, 927–960.

Ramscar, M., Yarlett, D., Dye, M., Denny, K., & Thorpe, K. (2010). Feature-label-order effects and their implications for symbolic learning. *Cognitive Science*, *34*, 909–957.

Rescorla, R. A. (1988). Pavlovian conditioning: It's not what you think it is. *American Psychologist*, *43*, 151–160.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). New York: Appleton-Century-Crofts.

Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, *17*, 249–255.

Rueda, M. R., Rothbart, M. K., McCandliss, B. D., Saccamanno, L., & Posner, M. I. (2005). Training, maturation and genetic influences on the development of executive attention. *Proceedings of the National Academy of Sciences*, *102*, 14931–14936.

Sandhofer, C. M., & Smith, L. B. (1999). Learning color words involves learning a system of mappings. *Developmental Psychology*, *35*, 668–679.

Schultz, W. (2006). Behavioral theories and the neurophysiology of reward. *Annual Review of Psychology*, *57*, 87–115.

Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science*, *237*, 1317–1323.

Shimamura, A. P. (2000). The role of the prefrontal cortex in dynamic filtering. *Psychobiology*, *28*, 207–218.

Siegel, S., & Allan, L. G. (1996). The widespread influence of the Rescorla-Wagner model. *Psychonomic Bulletin and Review*, *3*, 314–321.

Singleton, J. L., & Newport, E. L. (2004). When learners surpass their models: The acquisition of American sign language from inconsistent input. *Cognitive Psychology*, *49*, 370–407.

Slaughter, V., & Gopnik, A. (1996). Conceptual coherence in the child's theory of mind: Training children to understand belief. *Child Development*, *67*, 2967–2988.

Smith, L. B., Colunga, E., & Yoshida, H. (2010). Knowledge as process: Contextually-cued attention and early word learning. *Cognitive Science*, *34*, 1287–1314.

Thompson-Schill, S. L., Ramscar, M., & Evangelia, C. G. (2009). Cognition without control: When a little frontal lobe goes a long way. *Current Directions in Psychological Science*, *8*, 259–263.

Towse, J. N., Redbond, J., Houston-Price, C. M. T., & Cook, S. (2000). Understanding the Dimensional Change Card Sort. Perspectives from task success and failure. *Cognitive Development*, *15*, 347–365.

Waelti, P., Dickinson, A., & Schultz, W. (2001). Dopamine responses comply with basic assumptions of formal learning theory. *Nature*, *412*, 43–48.

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*, 1589–1607.

Yeung, N., Botvinick, M. M., & Cohen, J. D. (2004). The neural basis of error detection: Conflict monitoring and the error-related negativity. *Psychological Review*, *111*, 931–959.

Yeung, N., Nystrom, L. E., Aronson, J. A., & Cohen, J. D. (2006). Between-task competition and cognitive control in task switching. *Journal of Neuroscience*, *26*, 1429–1438.

Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): A method of assessing executive function in children. *Nature Protocols*, *1*, 297–301.

Zelazo, P. D., Frye, D., & Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cognitive Development*, *11*, 37–63.

Zelazo, P. D., Müller, U., Frye, D., & Marcovitch, S. (2003). The development of executive function. *Monographs of the Society for Research in Child Development*, *68*(Serial No. 274).