Two routes to cognitive flexibility: Learning and response conflict resolution in the dimensional change card sort task

Michael Ramscar, Melody Dye, Jessica Witten & Joseph Klein
Department of Psychology, Stanford University,
Jordan Hall, Stanford, CA 94305.

Abstract

There are at least two ways in which response conflict can be handled in the mind: dynamically, so that conflicting response demands are resolved on-line, and discrimination learning, which reduces the amount of on-line response conflict that needs to be resolved in context. While under fours are perfectly capable of discrimination learning, they appear to lack the ability to dynamically resolve response conflict. They can match their behavior to context in remarkably subtle and sensitive ways when they have learned to do so, but if they have not learned to match a response or a behavior to a context, their inability to handle on-line response conflict is their undoing (for example, in the dimensional change card sort task; DCCS). We present an analysis of how learning in context might aid children's performance in the dimensional change card sorting (DCCS) over time, and a training study in which three groups of age matched under fours attempt to complete the DCCS. We find that appropriate training enables children to flexibly switch between their responses in the DCCS. Without training supporting discrimination learning, children's performance is far worse, and when the task contexts are novel, children fail as expected.

Introduction

Thanks to the insight and inventiveness of developmental psychologists, we know that very young children are different. A three-year-old might girl seem simply a slightly smaller version of her four-year-old brother, however, while he will sail effortlessly, through the battery of tasks that psychologists have devised to expose the shortcomings of the very young, his sister will likely fail every one of them. Her 4-year-old brother will switch responses and probability match in binary choice tasks, understand false belief and the conflicting dimensions of appearance and reality, and switch easily between competing rules in dimensional change card sorting (DCCS; Zelazo, 2006) task, whereas our three-year-old will maximize in binary choice tasks (fixating on the most likely response), fail false belief tasks, be unable to switch from describing the appearance of an object to answering questions about what it really is, and fail to switch from one sorting rule to another, even though the rule is clearly stated (see Ramscar & Gitcho, 2007, for a review).

This raises two questions: first, why do children under four fail to switch to the conflicting but more contextually appropriate response in these tasks; and second, given the inflexibility of thought that these tests reveal, why in the normal course of events do children appear to be perfectly capable of changing their responses and behavior according to context?

Many proposals have been made in trying to answer the first of these questions (see e.g., Zelazo, Müller, Frye & Marcovitch, 2003). In what follows, we seek to answer both of them by examining the different ways in which the conflict between potential responses might be resolved, so that an appropriate response can be given in context. We suggest that that there are at least two ways in which response conflict can be handled in the mind: dynamic response conflict resolution, which enables conflicting response demands to be processed and resolved on-line, and discrimination learning, which enables the strengths by which responses are evoked by contexts to be modulated, reducing the amount of on-line response conflict that needs to be processed and resolved. We suggest that while under fours are perfectly capable of discrimination learning, they lack the ability to resolve response conflict on-line. Under fours are able to match their behavior to context in remarkably subtle and sensitive ways when they have learned to do so. If they have not learned to match a response or a behavior to a context, under fours' inability to handle on-line response conflict is their undoing (for example, in the novel contexts psychologists devise for their tests).

In what follows, we describe the neurological and computational bases for these ideas, and present a computational simulation of how discrimination learning and context might affect children's performance in the dimensional change card sorting (DCCS) over time. The model explains the observed failure of under fours at the DCCS as resulting from a lack of discrimination learning in the context of the "games" children play in the task. Further, it predicts that these children are exposed to the game contexts in ways that promote discrimination learning, they should later succeed at the task with relative ease. We then present a training study in which three groups of age matched under fours attempt to complete the DCCS after exposure to the games that promotes discrimination learning, exposure to the games that does not promote discrimination learning, and where the DCCS games are novel contexts. Consistent with the predictions of the model, we find that after appropriate discrimination learning, children are able to flexibly switch between the various responses required by the DCCS in a contextually appropriate manner. Without appropriate discrimination learning, children's performance is far worse, and when the task contexts are novel, children fail as expected.

The Dimensional Change Card Sort Task

In the Dimensional Change Card Sort (DCCS) Task, three and four year-old children are asked to sort cards with two prominent linked dimensions—a color and shape—into bins in which these dimensions have been reversed. For example, if the 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, and the red trucks will go with the blue trucks. When a child is asked to sort by one dimension—say, shape, switching the sort dimension to color will switch the correct sort bins for the card; e.g., red stars match to the truck bin when sorted by color, but the star bin when sorted by shape. For older children and adults, this is a straightforward task.

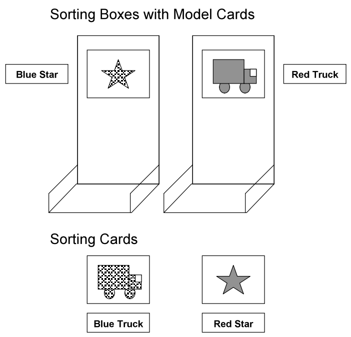

Figure 1: The basic DCCS task. Cards can be sorted by shape (in which case, the red star is sorted into the left bin) or color (in which case, the red star is sorted to the right bin).

When young children are asked to begin sorting by shape, they can easily answer questions regarding the rules for correctly sorting either by shape or by color. In addition, after switching from sorting by shape to sorting by color, children can correctly answer questions about how to correctly sort according to the new rule. However, once children are actually handed a card and asked to sort according to the second rule they have learned, their success in the task varies markedly with age. Generally, 3-year-old children are unsuccessful at this part of the task; they continue to sort the cards according to the first rule (i.e., whatever was learned first, whether it be sorting by shape or color). After age 4, however, children tend to pass the DCCS task and successfully match the cards to the bins both before and after the sorting rules are switched (Zelazo, Frye & Rapus, 1996).

Why do three year olds fail this task? One suggestion is that their poor performance is a related to the late development of prefrontal cortex. Like many other primates, humans are born with an immature brain. In monkeys the post-natal development of the brain occurs at the same rate in all cortical areas (Rakic, Bourgeois, Eckenhoff, Zecevic, & Goldman-Rakic, 1986). In the human cortex, however, while synaptogenesis peaks in visual and auditory cortex within a few months of birth, these developments occur later in prefrontal cortex (Huttenlocher & Dabholkar, 1997; for reviews see Thomson-Schill, Ramscar & Chrysikou, in submission).

One interesting behavioral consequence of this slow prefrontal development is that children appear unable to engage in behaviors that conflict with prepotent responses (see Ramscar & Gitcho, 2007 for a review). The adult ability to select a less well learned, but goal appropriate response is seen in the Stroop Task, in which the subject is asked to identify the ink color of a conflicting color word (e.g., if the word "green" were printed in red ink, red would need to be identified). Performance in this task involves resolving the conflict between the over-learned response (reading) and the appropriate response (ink naming). Adults typically complete the Stroop Task with ease, but young children repeatedly fail similar tasks. In adults, this is made possible by pre-frontal control mechanisms that bias one response over another according to goals or context (Yeung, Botvinick, & Cohen, 2004). The prefrontal cortex functions as a dynamic filter, selectively maintaining task-relevant information and discarding task-irrelevant information (Shimamura, 2000).

If three year olds lack (or are deficient in) the ability to dynamically filter responses in accordance with the demands of a context or goal, this may explain both why they fail at the Stroop Task and why they fail to switch rules in the DCCS. If a card depicts a red star, "red" elicits one response (sorting into the color bin) whereas "star" elicits a different conflicting response (sorting into the shape bin). Thus in the standard DCCS task, successfully switching rules involves changing from one response associated with a given cue—the card—to an alternative, conflicting response. Since this kind of response conflict processing appears to be the preserve of the frontal areas of the brain (Yeung, Botvinick, & Cohen, 2004; Thomson-Schill et al, *in press*), it seems that the failure of three year olds in the DCCS task—that is, their failure to mediate response conflict—may be related to slow pre-frontal development.

Discrimination Learning

If young children lack the ability to resolve conflict on-line, discrimination learning provides another means by which they might still learn to succeed on the DCCS. This is because the games associated with each sorting rule provide cues to the appropriate responses, in addition to the shape and color in the cards themselves. The "shape game" is a cue to the response "sort into the shape bin" and the "color game" is a cue to the response "sort into the color bin." Since children fail the task despite the presence of these cues, it is clear that under ordinary circumstances, the game cues do not provide sufficient extra scaffolding to enable three year olds to pass the DCCS. However, an obvious difference between the cards and the games is that children have a lot of experience with colors and shapes and the various responses they elicit, whereas they have comparatively little experience with sorting games.

To explain why this might matter, we need to consider the way that responses that lead to response conflict in the DCCS are learned and discriminated. Discrimination learning is a process by which information is acquired about the probabilistic relationships between important regularities in the environment (such as objects or events) and the cues that allow those regularities to be predicted (see e.g., Rescorla & Wagner, 1972; Gallistel & Gibbon, 2000).

Crucially, the learning process is driven by discrepancies between what is expected and what is actually observed in experience (termed "error-driven" learning). The learned predictive value of cues produces expectations, and any difference between the value of what is expected and what is observed produces further learning. The predictive value associated with cues is strengthened when relevant events (such as events, objects or labels) are under-predicted, and weakened when they are over-predicted (Kamin, 1969; Rescorla & Wagner, 1972). As a result, cues compete for relevance, and the outcome of this competition is shaped both by positive evidence about co-occurrences between cues and predicted events, and negative evidence about non-occurrences of predicted events. This produces patterns of learning that are very different from those that would be expected if learning were shaped by positive evidence alone (a common portrayal of Pavlovian conditioning). Learners discover the predictive structure of the environment, and not just simple patterns of correlations in it.

To briefly illustrate how discrimination learning works, imagine a child learning to play the games associated with the DCCS. We shall first consider a case where the experimenter shows the child the card, and is asked to sort them by color.

We can assume that previously the child has heard objects referred to before in terms of both their shape and their color because, though they usually fail to sort using these dimensions, they can reliably name the shapes and colors on the cards (Kirkham, Cruess & Diamond, 2003). The problem, therefore, seems to be that children experience more response-conflict with regards the correct dimension to attend to in order to sort by the rule than they do when it comes to selecting an appropriate dimension for naming (this is perhaps unsurprising, since children will have more experience with names than sorting). That is, when children are asked to *sort* the cards, both shape and color appear to be active as relevant dimensions to sort by.

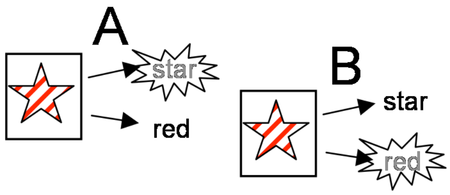

Figure 2. If a child has learned that a card with a red star on it might be sorted by *red* or *star*, when the card is presented she will expect to sort by *red* and *star*. In sorting by *red* (A), the child's expectations will weaken the association between the card and *star* in this context. The converse is true in the (B).

If the cards cause a child to expect both dimensions to be relevant, but only one is used in sorting, there will be a violation of expectation (Figure 2). Given that a response to the relevant dimension event didn't occur, she will begin to adjust her expectations accordingly. This may then cause problems when the child is asked to sort by the other dimension, because the child will have learned to *ignore* the now relevant dimension on the earlier sort trials.

This is because in the *color game* the *red star card* is sorted by "red." Because the *red star card* has been previously associated with both "red" and "star", it also incorrectly cues "star." As a result, the value of the association between *red star card* and "star" will decrease ("star" will be learned about even though it is not heard). Further, because the context *color game* has been introduced, in subsequent color game trials, a conjunctive cue *red star card* + *color game* (e.g., Gluck & Bower, 1988) can compete with *red star card* (and *color game*) for associativity to "red".

The converse will occur if the child switches to the shape game. Because all of the dimensions of the *red star card* will be present in both the color and the shape games, *red star card* alone will prove to be a less useful cue than the conjunctive cues *color game* + *red star card* and *shape game* + *red star card*.

To formally test these ideas, we simulated the competition between conjunctive cues representing *color game* + *red* and *shape game* + *star* and the individual cues *red* and *star* across repeated DCCS trials using the Rescorla & Wagner (1972) model. [1] The

[1] In the Rescorla-Wagner (1972) model the change in associative strength between a stimulus *i* and a response *j* on trial n is defined to be:

$$\Delta V_{ij}^{n} = \alpha_i \beta_j (\lambda_j - V_{total})$$

simulation assumes that the output is the appropriate sorting response, and that *red* and *star* have been previously learned as sorts for the *red star card* "red" 50% of the time each, and that *color game* + *red* will predict "red" 100% of the time. The individual cue was initially trained on with color and shape as alternate labeling events, and then the *color game* was introduced, and that *color game* was present on all color trials (there are two colors, equally represented).

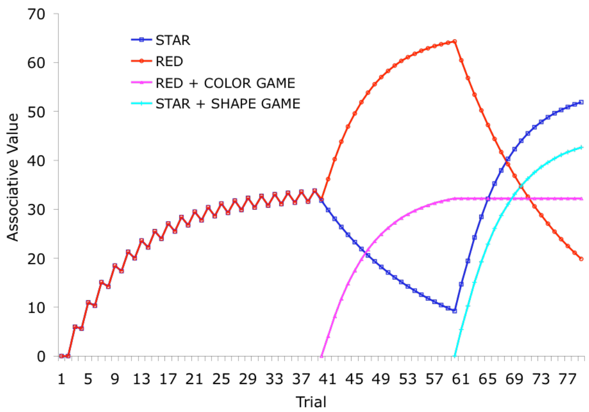

Figure 3: Rescorla-Wagner simulation of cue competition in two DCCS trials. The erroneous expectations *shape* produces in *color game* trials cause it to be unlearned, resulting in red is being a far more active cue on the switch trial (trial 61).

In the first DCCS game shown in Figure 3, *red* and the conjunctive cues the *color game* + *red* gain in associative value as a result of the diminishing value of the *star* cue. Importantly, even though all of the cues co-occur with exactly the same frequency with "red," learning effectively dissociates *red star card* and *color game* from "red" in this situation.

As can be seen in Figure 4, assuming correct sorting, the erroneous expectations produced by *red* and *star* cause them to lose out in competition with the conjunctive cues that embody the games as contexts, such that the dimensional cues alone are effectively unlearned in this context, even though they co-occur with the appropriate responses with exactly the same frequency as the conjunctive cues. This is because in error-driven learning predictive power, not frequency or simple probability, determines cue value. Thus, as long as the cards are correctly labeled in each context, a child will learn to ignore the ambiguous cues, thereby improving response discrimination.

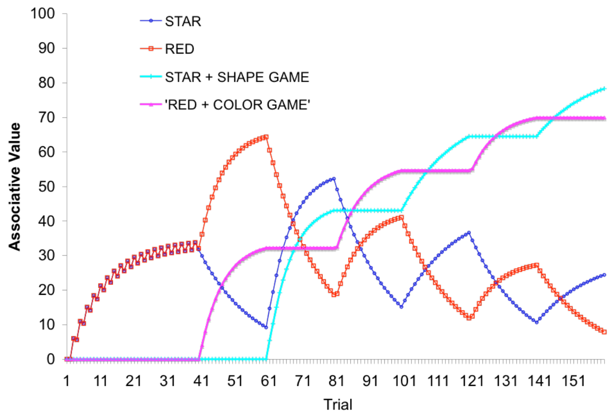

Figure 4: Rescorla-Wagner simulation of cue competition in six DCCS trials. Each peak represents a rule switch.

Cue competition devalues the cues that result in prediction error and increases the value of those that do not, emphasizing the value of reliable cues. To illustrate the importance of cue competition to discrimination learning, it is useful to consider the effect of learning in the absence of cue competition.

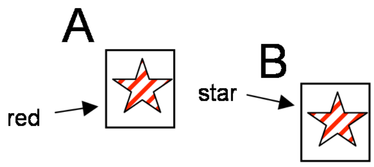

Figure 5: When labels precede the cards as discrete events, there may be no opportunity for cue competition. Each cue will simply come to predict the card to asymptote.

We call learning in the situation just described, where card Features predicted Labels, *FL-learning*. We can then define the situation in which Labels predict Features as *LF-learning* (Ramscar, Yarlett, Dye, Denny & Thorpe, in submission). In this situation, something very different will happen in learning. To explain why, we need to consider how the structure of cues and predicted events conspire to produce cue competition. In the FL-learning scenario described above, the labels for the relevant dimensions are discrete, and only one occurs at any one time. This results in prediction error if cues present on trials when "red" is subsequently labeled are present on trials when "star" is subsequently labeled. Cues not present on one or the other type of trial come to be favored as a result of cue competition. However, if the labels (or the labels in context) are presented prior to the cards (Figure 5), because the labels are discrete as events and as stimuli (whereas the dimensions of the cards in context are not), they cannot compete as cues, so no discrimination learning will take place.

The model thus specifies how the associative strength (V) between a conditioned stimulus (CS_i) and an unconditioned stimulus (US_j) changes as a result of discrete training trials, where n indexes the current trial. $0 \leq \alpha_i \leq 1$ denotes the saliency of CS_i, $0 \leq \beta_j \leq 1$ denotes the learning rate of US_j, λ_j denotes the maximum amount of associative strength that US_j can support, and V_{total} is the sum of the associative strengths between all CSs present on the current trial and US_j. Learning is governed by the value of (λ_j - V_{TOTAL}) where λ_j is the value of the predicted event and V_{total} is the predictive value of a set of cues. In the simulation, all λ = 100%, α_i=0.2 and β_j=0.3.

Because there are no other labels (cues) to compete for associative value, there can be no loss of potential associative value to other cues over the course of learning. Because of this, the effect of prediction-error on cue value will be very different. In the absence of cue competition, the cue value of a label will simply come to represent the proportion of successful predictions it has made relative to the proportion of unsuccessful predictions; the cue value of a label will simply approximate the conditional probability of a feature given the label (in the DCCS, where cards vary in color or shape, this variance will be represented probabilistically after LF-learning). LF-learning thus provides little help when it comes to learning about situations in which response conflict is inherent (Ramscar et al, in submission).

Error-Driven Learning and the DCCS

The analysis above suggests that if children correctly respond to the appropriate dimensions in the early stages of the DCCS, contextual learning will reduce response conflict in later trials. Children trained to associate sorting by shape with a "shape game" and sorting by color with a "color game" can eliminate the response-conflict normally associated with the DCCS by learning context-dependent rules; for example, *"**red** star card* + ***shape game*** *sort by red. "*

Given stimulus generalization (Shepher, 1987), one might expect that these will generalize to a degree to, *"**color** shape card* + ***color game*** *sort by color"* Similarly, we might expect that if children learn to attend to one dimension in learning about a response in context, they might transfer that learning to another response. Since children can *name* the appropriate dimensions of the cards in the DCCS before they can sort them, we expected that if children were taught to associate naming the appropriate contexts with the game rules in an FL-training configuration, they would learn the high predictive value of these specific cue configurations and that this contextual learning might then enable them to successfully sort in the same contexts in the DCCS task.

Since we would expect that similar training in LF configuration would result only in the learning of the transitional probabilities between the dimension labels and the cards (as described above), the lack of cue competition in this condition ought to result in far less reduction in the amount of response conflict in the task than FL-Learning. To test these ideas, we examined the effect this kind of off-line discrimination training on children's on-line performance in the DCCS.

Training Experiment

Participants

47 children between 3- and 4-years-old (M = 3 years, 6.8 months) participated in this study.

Methods and Materials

Two groups of children received either Label-Second (FL) or Label-First (LF) training on the cards, before completing standard DCCS tasks (Zelazo, 2006). A control group was tested on the DCCS without training.

In the XL (label-second) condition, children were introduced to the shape and color games prior to the DCCS. They were 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 6 cards in the new game.

While children in the FL-condition saw the card and labeled it, children in the LF-condition were asked to say the label first and then saw 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 structure of the LF-training was the same as the FL-training: naming 6 cards by one dimension and then switching to the other dimension.

The two training groups (FL and LF) then completed two standard DCCS tasks, with the first testing dimension (either shape or color) counterbalanced across children. There were 12 test trials completed by each child (six consecutive trails for the first dimension and six for the second dimension). Children were required to correctly sort six cards in the pre-switch, and before each trial, children 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.

Once a child had sorted six cards along the pre-switch dimension, the sorting dimension was switched. Exactly six cards were sorted in the post-switch test. After the first DCCS task, the children completed a second standard DCCS task with new cards.

Results

All the children in the two training conditions correctly labeled the cards. Children were considered to have "passed" the DCCS task if they sorted at least 5 out of 6 of the post-switch cards correctly. 69% of the FL-trained children successfully switched rules in the first DCCS task, and 75% in the second DCCS task. By contrast, in the 33% LF trained children completed the first rule switch, and 40% the second. 19% of the control children switched rules in each test (Figure 4).

Chi-square ($\chi 2$) tests revealed that children in the FL (Label-Second) condition had significantly higher passing rates (11/16 children passed) in the first DCCS as compared to children in the LF (Label-First) condition (5/15); χ2 [1, N = 31] = 9.7, p = 0.005; second test, label first, 12/16 children passed as compared to 6/15 in the label second condition, χ2 [1, N = 31] = 17.0, p = 0.001). Against the control group (3/16), the comparisons with the FL (Label-Second) group were, first switch, χ2 [1, N = 33] = 14.9, p = 0.001; second switch, χ2 [1, N = 33] = 23.7, p = 0.001.

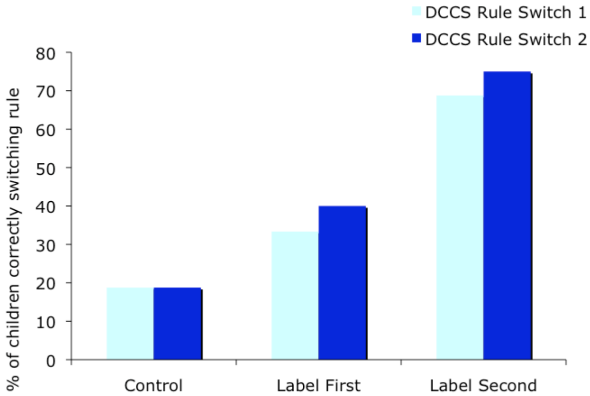

Figure 4: Percentage of children successfully switching rules in the first and second DCCS tasks by condition.

Discussion

We suggested that the observed failure of under fours in the DCCS might result from a lack of discrimination learning about the contexts provided by the "games" children play in the task. We predicted that if children were exposed to the game contexts in ways that promote discrimination learning, they would later succeed at the task with relative ease. Consistent with these predictions, we found that after appropriate discrimination learning, children were to flexibly switch between the various responses required by the DCCS in a contextually appropriate manner. With less appropriate discrimination learning, children's performance was far worse, and when the task contexts were novel, children failed as expected.

This finding is consistent with our suggestion that that there are at least two ways in which response conflict can be handled in the mind: dynamic response conflict resolution, which enables conflicting response demands to be processed and resolved on-line, and discrimination learning, which enables the strengths by which responses are evoked by contexts to be modulated, reducing the amount of response conflict that needs to be processed and resolved. It appears that while under fours are perfectly capable of discrimination learning, they lack the ability to resolve response conflict on-line (see also Ramscar & Gitcho, 2007; Thomson-Shill et al, in submission). As the children who received FL-Training show, discrimination learning allows under fours to match their behavior to context in remarkably subtle and sensitive ways once they have learned to do so. However, as the performance of children in the LF-training and control groups shows, if children have not learned context appropriate behavior, their inability to resolve response conflict dynamically causes problems when dealing with the demands of responding flexibly in ambiguous situations.

Acknowledgments

This material is based upon work supported by NSF Grant Nos. 0547775 and 0624345 to Michael Ramscar.

References

Gallistel, C.R., & Gibbon, J. (2000). Time, rate and conditioning. *Psychological Review*, 107, 344-389

Gluck and Bower (1988). Evaluating an Adaptive Network Model of Human Learning, *Journal of Memory and Language*, 27, 166-195

Huttenlocher, P. R., & Dabholkar, A. S. (1997). Regional differences in synaptogenesis in human cerebral cortex. *Journal of Comparative Neurology*, 387(2), 167-178

Kamin L.J. (1969). *Predictability, surprise, attention, and conditioning.* In: Campbell B, Church R (eds). Punishment and Aversive Behaviour. Appleton-Century-Crofts: New York.

Kirkham, N.Z. , Cruess, L.M., & Diamond, A. (2003). Helping Children Apply their Knowledge to their Behavior on a Dimension-Switching Task. *Developmental Science*, 6(5), 449-467.

Ramscar, M. Yarlett, D., Dye, M., Denny, K., & Thorpe, K. (in submission) The Feature-Label-Order Effect In Symbolic Learning

Rescorla R.A. and Wagner A.R. (1972). A Theory of Pavlovian Conditioning: *Variations in the Effectiveness of Reinforcement and Nonreinforcement.* In Black & Prokasy (Eds.), Classical Conditioning II: Current Research and Theory. New York: Appleton-Century-Crofts.

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

Thompson-Schill, S. L., Ramscar, M., & Evangelia, C. G., (in submission) Cognition without control: When a little frontal lobe goes a long way

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(4), 931-959

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. (2006). The dimensional change card sort (DCCS): A method of assessing executive function in children. *Nature Protocols*, 1, 297-301.