Children's Inductive Inferences are Influenced by Some Features More than Others

Nicholaus S. Noles (n.noles@louisville.edu)
Department of Psychological & Brain Sciences, 317 Life Sciences, University of Louisville
Louisville, KY 40292 USA

Judith Danovitch (j.danovitch@louisville.edu)
Department of Psychological & Brain Sciences, 317 Life Sciences, University of Louisville
Louisville, KY 40292 USA

Abstract

The present study integrates ideas and approaches from studies of psychophysics and conceptual development. Preschool age children were presented with creatures that either fit into a prototypical category or had features/labels from competing categories. Children were asked to complete classification (e.g., identify a creature's category) and induction (e.g., identify a creatures missing features, given a category label) trials. The feature dimensions tested, shape and color, were selected in order to evaluate the relative salience of different kinds of features. When presented with prototypical category members, children's classifications and inductions were accurate. When presented with creatures with counter-predictive features, neither feature was found to unduly influence children's classifications. In contrast, when children were asked to make inductions about creatures with features that conflicted with their category labels, one feature – color – was found to significantly influence inductions while the other feature – shape – did not. This finding indicates that consideration of psychophysical properties is required in order to accurately interpret studies exploring children's conceptual development.

Keywords: categories; labels; concepts; development; induction;

Relatively early in development, children exhibit an impressive ability to organize the world around them. With relatively sparse input, as limited as a verbal label from a peer or adult, children can form categories that guide their classifications of objects, individuals, and events. Moreover, children use these categories to make projective inductions. They apply their knowledge of known category members to make educated guesses about the features of novel category members. So, if a child knows that their pet cat is safe to approach when it is purring and dangerous to approach when it wags its tail, then they can apply that knowledge to draw inferences about new cats that they encounter (for a review, see Murphy, 2002).

The fact that children adeptly navigate categorization and induction is uncontroversial. However, the source of these competencies is a perpetual source of debate within the field of child development. Although there are a number of competing theories and models that might be used to explain how children use and acquire categorization and induction behaviors, these explanations can be roughly condensed into two approaches. Similarity-based approaches focus primarily on associative, perceptual, and statistical factors when investigating and characterizing children's competencies (e.g., Jones & Smith, 1993; McClelland & Rogers, 2003; Rakison, 2004; Sloutsky & Napolitano, 2003). In contrast, theory-based approaches focus on children's knowledge and intuitive theories to address their understanding of intentions, causality, and other nonobvious features of entities (e.g., Gelman, 2003; Gopnik & Sobel, 2000; Opfer & Bulloch, 2007; Wellman & Phillips, 2001).

These two approaches have remained at odds for decades because they have complementary advantages and disadvantages. Similarity-based approaches are elegant because of their simplicity, deriving explanatory power from domain-general cognitive processes related to attention and perception. At some level, these processes, and their related brain areas, must be recruited in perceiving and learning concepts and categories. However, hypotheses and models grounded in a similarity-based approach have had difficulty addressing some basic characteristics of concepts and conceptual development. For example, children's inductive inferences are commonly guided by category membership, even when perceptual features are available and informative (Gelman & Markman, 1986). Also, similarity-based approaches sometimes struggle to explain some basic findings related to categories and concepts, including that some features are more important for categorization than others (i.e., feature centrality, see Gelman & Wellman, 1991) and that the salience of a given feature can vary across categories (i.e., context sensitivity, see Medin & Shoben, 1988).

Theory-based approaches have little difficulty addressing these complications because they attribute deep, and sometimes complex, naïve or intuitive theories to individuals, including young children. Thus, where similarity-based approaches are elegant and efficient, theory-based approaches presuppose substantial innate conceptual acumen. Conversely, theory-based accounts address, and in some cases predict, a wide range of behaviors and intuitions in children and adults that are problematic for similarity-based accounts.

Today, the work produced by researchers employing each of these approaches continues in parallel, advancing separate theoretical agendas without directly addressing the friction between views. However, there are some notable exceptions. Specifically, Sloutsky and colleagues (see Sloutsky & Fisher, 2004; Sloutsky & Napolitano, 2003) developed a similarity-based theory called *SINC*. Their model characterizes category labels as perceptual features of entities in an attempt to explain circumstances where children employ labels, and not reliable perceptual features, to guide their inductive inferences (Gelman & Markman, 1986). However, the methods used to develop and test this model were methodologically flawed (Noles & Gelman, 2012a; 2012b). In contrast, Waxman and Gelman (2009) took a different approach, suggesting that the tension between similarity- and theory-based approaches is actually founded upon a false dichotomy between perceptual and conceptual factors, noting that both are critical to cognitive development, and that theorists primarily differ in their emphasis of some factors over others.

The goal of the current project is to begin evaluating the claims made by Waxman and Gelman using the psychophysical approaches developed by Sloutsky and colleagues. Doing so requires integrating research from perception and psychophysics into approaches used to study children's conceptual development. This pairing is not new to the field, but the pairing of perceptual methods with developmental topics is particularly challenging because child participants lack the expertise and attention span of adults.

As a first step toward the goal of more tightly integrating research on perception and conceptual development, the current project evaluates whether certain stimulus dimensions influence children's classification and projective induction more than others. There is evidence that some features are more salient than others (Gelman & Wellman, 1991) and that salient features can guide inductions (Deng & Sloutsky, 2012), but these studies mix and match features that lie along a broad continuum of stimulus dimensions, some of which are more or less categorical and perceptually distinguishable (Garner, 1974). Indeed, the contrasts within and between stimulus dimensions are critical to interpreting categorical and perceptual effects because children and adults flexibly attend to more discriminable, higher contrast stimulus dimensions when they are available (Noles & Gelman, 2012b). Thus, the current study is designed to begin the process of evaluating the relative salience of different kinds of features.

As an initial step, this study focuses on two common features typically used in studies of classification and induction: color changes and shape changes. A third variable, the presence or absence of a third feature, was also employed in order to add additional variety and complexity to the test stimuli. The paradigm used in this study is schematically similar to the design developed by Deng and Sloutsky (2012).

Method

Participants

Prior studies of classification and induction, particularly using the approaches employed in the present study, largely focus on children between the ages of three and five. Thus, that age group was the focus of the present study. Fourteen preschool aged children (M = 4.64, SD = .25, *range*: 4.13 to 5.05, 7 female) were recruited from daycares in urban and suburban settings in the greater metro area of Louisville, Kentucky. Children were tested individually by an experimenter at their school, and they received a certificate and sticker as rewards for participation. Data from one additional child was excluded from analysis because the child failed to follow directions.

Materials & Procedure

The materials used in this study consisted of drawings of artificial creatures. Each creature had three features, including color, a "bottom" feature (i.e., a tail), and a "top" feature (i.e., a mouth or horns). The key features in this study were the color and tail feature. As in prior studies, each of these features was set to one of two binary values. Color was either a fully bright and saturated value for red (rgb = 255, 0, 0) or the same color at 50% the brightness. The bottom or tail feature was one of two triangles. Both triangles had the same area, but one was four times the height of the other. The third and final feature was either present or absent on top of each creature. Two artificial prototypes were created by randomly assigning one color and tail to create two prototype creatures for two categories, *flurps* and *jalets*, see Figure 1. The third feature appeared on half of the creatures and was category neutral (i.e., each individual flurp or jalet had a 50% chance of having a top feature). In addition to the prototypical creatures, *mismatch* creatures were generated by mixing features across categories (e.g., a flurp color with a jalet tail, in classification trials) or presenting a creature with a diagnostic feature from one category with a label from the other category (e.g., a creature labeled as a flurp that is the color of a jalet, in induction trials). Other materials included creatures with competing features from both prototypes, color patches depicting both values of color used in the stimuli, black and white line drawings of creatures, disembodied tails, and two audio tracks. The audio tracks consisted of a woman's voice saying, "this is a flurp," and "this is a jalet." Each was approximately four seconds long.

Participants were tested individually in a quiet space at their preschool. Stimuli were displayed using presentation software on a laptop with a 15-inch screen, and the experimenter determined the pace of the session, ensuring that children were looking at the screen before beginning each trial. Three kinds of trials were presented to participants, including training trials, classification trials, and induction trials.

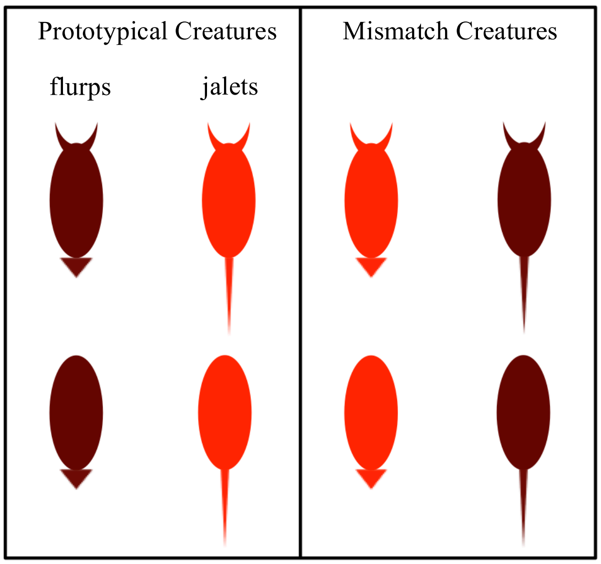

Figure 1: Prototypical and mismatch creatures. Mismatches have one feature from each category.

Procedure

Training Trials The experiment began when the experimenter said, "I'm going to show you two kinds of creatures, flurps and jalets. I want you to look at them and remember what they look like because I will ask you about them later." Children then saw each prototypical creature six times, for a total of 24 training trails. These trials were split evenly between the two to-be-learned categories, and they depicted prototypical jalets and flurps. Each was accompanied by a four second audio track that both labeled each creature and indicated when the experimenter should progress to the next trial (i.e., exposure to each trial was approximately 4 seconds). Training trials were presented in a random order for each participant.

Classification Trials Classification trials began with the following instructions: "I'm going to show you some more creatures that look a lot like the ones before. This time I'm going to show you a creature and ask you to tell me if it's a flurp or a jalet." Classification trials were identical to training trials except that the creatures were not labeled. After each trial, the experimenter asked if the creature was a jalet or a flurp. Participants were presented with eight warm-up trials, divided evenly between categories so that each prototypical creature appeared twice. These trials were presented with feedback. If a creature was incorrectly classified, the experimenter said, "Oops, that's a _____." Test trials without feedback began immediately after the 8th warm-up trial, and consisted of 16 trials, including two instances of each prototypical creature and two instances of each mismatch creature. Mismatch creatures were constructed using one feature from each category. Classification trials were presented pseudo-randomly. The first two trials were always prototypical creatures, but the remaining trials were intermixed and presented randomly.

Induction Trials After the final classification trial, the experimenter said, "Now I'm going to show you some more creatures that look a lot like the ones before. This time I'm going to tell you if it's a jalet or a flurp, but I'm going to ask you to guess what's missing." In each induction trial, participants were presented with labeled creatures that were missing a target feature. Because this experiment focused on color and shape changes, these features were missing from each creature, and participants were prompted to select which of two possible features was missing from the creature (e.g., "Which color/bottom is missing?" see Figure 2). Induction trials featured a warm-up and test phase with the same number, disposition, and pseudo-randomization as the previously displayed classification trials

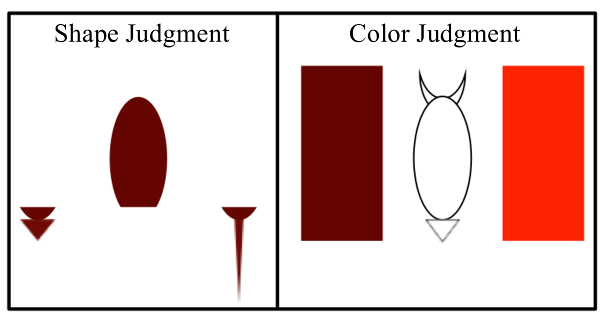

Figure 2: Example induction trials. Each trial was labeled, and participants were asked to indicate the missing feature.

Results

Classification Children were very accurate when classifying prototypical creatures, making correct identifications on 80% of trials, which was significantly greater than chance responding, $t(13) = 4.67$, $p < .001$. Since the informative features on each mismatch creature were counter-predictive, accuracy for mismatch trials was expected to be random if the tested features were equivalently salient. For the purpose of detecting differences in salience, accuracy on these trials was therefore arbitrarily defined as a classification that was consistent with each mismatch creature's tail. As can be seen in Figure 3, participants responded rationally and randomly (M = 49%) when presented with conflicting features.

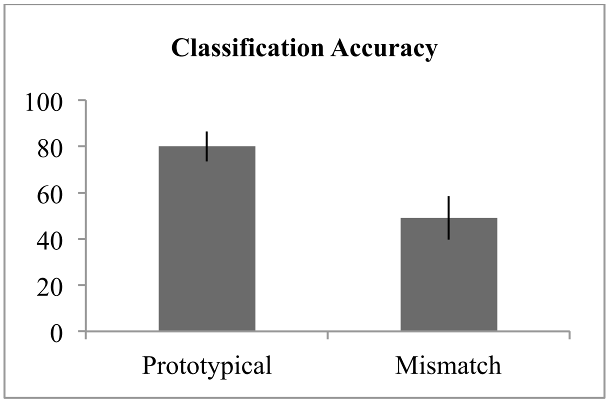

Figure 3: Children accurately identified prototypical creatures, and their classifications were not unduly influenced by either Feature-type.

Induction For the purpose of evaluating inductions, label-consistent responses were defined as "accurate." In the case of prototypes, this means that participants chose the missing feature that matched both the label and the visible informative feature. In the case of mismatches, responses were labeled as accurate when they aligned with labels (as in Gelman & Markman, 1986), and not the visible feature from the opposite category (e.g., if a creature with a jalet color is labeled as a flurp, and then the participant is accurate if s/he selects the feature that matches the label and not the visible feature).

Participants' inductive inferences were evaluated with a 2x2 repeated measures ANOVA with Feature-type (color vs. shape) and Creature-type (prototype vs. mismatch) as within-subjects factors. This analysis revealed a significant main effect of Creature-type, $F(1,13) = 7.32$, $p < .05$, $\eta_p^2 = .36$, indicating that accuracy for prototypes was significantly greater than mismatches. Planned comparisons (Bonferroni-corrected) further revealed that the difference between prototype and mismatch creatures was only significantly different when color was presented in opposition to labels, $p < .05$. To put this finding in perspective, the means for prototypes and mismatches across both Feature-types were compared to chance responding using one-sample t-tests (chance = .5). Every set of responses tested in this study significantly exceeded chance responding (p's < .05) except when labels were placed in competition with visible colors.

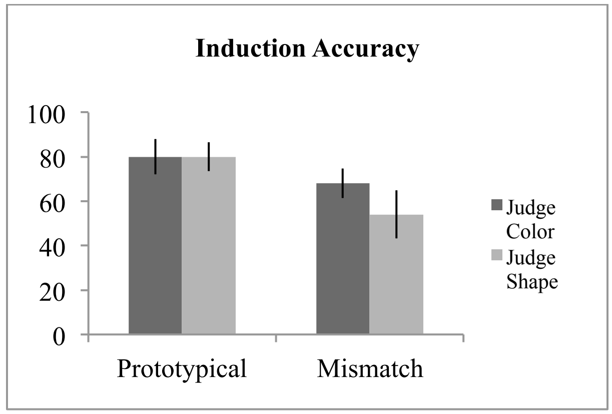

Figure 4: Children made accurate inductions when presented with prototypical creatures. When presented with mismatches, their inductions were label-consistent when judging color, but random when judging shape.

Discussion

Children's classifications were very accurate when they were presented with prototypical creatures. This result indicates that they learned the target categories and remembered them. When presented with mismatch creatures, children responded randomly but rationally. Because mismatch creatures were constructed of features from competing categories (e.g., a jalet's tail paired with a flurp's color), children had only conflicting evidence to guide their classifications. If children had a predisposition to only learn a single feature, or to view one feature as more informative for classification, then children's responses would have been non-random. Instead, their random responding indicates the quality of their learning and their lack of biases with respect to feature dimensions when making classification judgments.

Children's inductions revealed a different pattern of behaviors. As in classification trials, children were accurate when making inductions about features missing from prototypical creatures. Thus, when the label and the visible feature were category-consistent, children effectively identified the missing feature, regardless of which feature they were judging.

In contrast, children's inductions diverged by Feature-type when they were presented with mismatch creatures. When they were asked to judge color, children exhibited a bias to make label-consistent judgments, as indicated by responses that significantly differed from chance but did not differ from judgments of prototypical creatures, which were always label-consistent if accurate. However, when asked to judge shape, children's responses were significantly less accurate (i.e., less label-consistent) than their responses to prototypical creatures, and their responses did not differ from chance.

Recall that mismatch creatures were constructed from a label from one category and a visible feature from a competing category. Thus, there was no "right" answer, and either the label or the feature might have guided children's intuitions. In prior studies (e.g., Gelman & Markman, 1986), children tended to make label-consistent judgments, and indeed, theorists have posited that they did so because labels are category-referring, and thus more informative than other features of entities.

In the current study, children's responses revealed that they found labels to be more salient cues to category membership than the shape of a part, but they found labels and colors to be equally salient. This pattern of results indicates that there were important differences in salience between the stimulus dimensions of color and shape. Broadly, these results represent evidence that stimulus dimensions are not all equally salient or informative. Deng and Sloutsky (2012) presented a similar finding when they reported that manipulating the salience of a single feature (e.g., by making it move) provoked children into making more feature-consistent and fewer label-consistent inductions. The current results extend this finding further, indicating that even in the absence of explicit manipulations of salience, some stimulus dimensions were inferred to be more informative or salient than others.

The pattern of results recorded here both supports and undermines similarity-based accounts of representation and conceptual development. On one hand, these data reveal that results focusing on salience may actually be tapping into differences between stimulus dimensions, and not manipulations of attention. For example, Deng and Sloutsky's effect might be attributable to the part that they selected to make more salient, and not to the manipulation of salience that they applied. On the other hand, these results support claims that feature salience is important and influences inductions. However, these results also provide a mechanism by which certain kinds of comparisons can be combined with specific features in order to provoke findings that support a false dichotomy between perceptual and conceptual factors, as suggested by Waxman and Gelman (2009).

Perhaps the most important outcome of this study is to highlight a fundamental problem with studies that probe conceptual representations. Specifically, broad claims about the role of labels or feature salience cannot be evaluated without a strong understanding of the between- and within-feature contrasts that they represent. It is important to acknowledge that decisions made while designing auditory and visual stimuli may influence or provoke certain response patterns in young participants. Historically, these factors have received little attention, even though they have recently been identified as powerfully influencing both child and adult participants (e.g., Noles & Gelman, 2012a; 2012b).

Studies of psychophysics focusing on child participants are relatively rare. Preschoolers especially are not well suited to the number of trials or the kinds of manipulations usually employed by such approaches. However, it is increasingly obvious that the mechanics of perception and attention are playing important and under-represented roles in our modern understanding of children's conceptual development. The purpose of the current project to address this gap in our understanding, and to begin to contextualize and understand how past design decisions and experimental methodologies have shaped our modern understanding of children's categorization and induction behaviors.

Future studies should focus on exploring other popular manipulations, such as changes in size and positioning, as well as the presence or absence of features, which was a random factor in the present study but is currently being studied in ongoing projects. Contrasting values of these features are employed in many studies, but they are not well understood. There are also findings and theories in perception that are highly relevant to interpreting findings using different stimulus dimensions (e.g., the work of Wendell Garner), but that are not widely integrated into modern studies of child development. More generally, these studies may hold the key to addressing ongoing debates between theory- and similarity-based approaches to understanding conceptual development.

Acknowledgements

We thank the staff, children, and parents at Keneseth Israel Preschool, Newburg Kindercare, and St. Paul School in Louisville, KY for their support. Thank you to Emma Groskind, Jacob Messmer, Kristain Pile, Kayla Renner, and Gretchen Santana for their assistance.

References

Deng, W., & Sloutsky, V.M. (2012). Carrot Eaters or Moving Heads : Inductive Inference Is Better Supported by Salient Features Than by Category Labels. *Psychological Science, 23,* 178-186.

Garner, W.R. (1974). *The processing of information and structure.* New York, NY: Lawrence Erlbaum Associates.

Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought.* New York, NY: Oxford University Press.

Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition, 23,* 183–209.

Gelman, S. A., & Wellman, H. M. (1991). Insides and essences: Early understanding of the non-obvious. *Cognition, 38,* 213–244.

Gopnik, A., & Sobel, D. M. (2000). Detecting blickets: How young children use information about novel causal powers in categorization and induction. *Child Development, 71,* 1205–1222.

Jones, S. S., & Smith, L. B. (1993). The place of perception in children's concepts. *Cognitive Development, 8,* 113–139.

McClelland, J. L., & Rogers, T. T. (2003). The parallel distributed processing approach to semantic cognition. *Nature Reviews Neuroscience, 4,* 310–322.

Medin, D. L., & Shoben, E. J. (1988). Context and structure in conceptual combination. *Cognitive Psychology, 20,* 158–190.

Murphy, G. L. (2002). *The big book of concepts.* Cambridge, MA: MIT Press.

Noles, N.S., & Gelman, S.A. (2012a). Effects of categorical labels on similarity judgments: A critical analysis of similarity-based approaches. *Developmental Psychology, 48,* 890-896.

Noles, N.S., & Gelman, S.A. (2012b). Preschool-age children and adults flexibly shift their preferences for auditory versus visual modalities but do not exhibit auditory dominance. *Journal of Experimental Child Psychology, 112,* 338-350.

Opfer, J. E., & Bulloch, M. J. (2007). Causal relations drive young children's induction, naming, and categorization. *Cognition, 105,* 206– 217.

Rakison, D. H. (2004). Infants' sensitivity to correlations between static and dynamic features in a category context. *Journal of Experimental Child Psychology, 89,* 1–30.

Sloutsky, V. M., & Fisher, A. V. (2004). Induction and categorization in young children: A similarity-based model. *Journal of Experimental Psychology: General, 133,* 166–188.

Sloutsky, V. M., & Napolitano, A. C. (2003). Is a picture worth a thousand words? Preference for auditory modality in young children. *Child Development, 74,* 822–833.

Waxman, S. R., & Gelman, S. A. (2009). Early word-learning entails reference, not merely associations. *Trends in Cognitive Sciences, 13,* 258–263.

Wellman, H. M., & Phillips, A. T. (2001). Developing intentional understandings. In B. Malle, L. Moses, & D. Baldwin (Eds.), *Intentionality: A key to human understanding* (pp. 125–148). Cambridge, MA: MIT Press.