

Children's Use of Relevance in Open-Ended Induction in the Domain of Biology

Anna Z. Vitkin (vitkin.a@neu.edu)

John D. Coley (j.coley@neu.edu)

Ruiwen Hu (hu.r@neu.edu)

Department of Psychology

360 Huntington Avenue

Boston, MA 02115 USA

Abstract

Induction refers to the process by which we use our existing knowledge to make predictions in the face of uncertainty. Recent research in the domain of biology suggests that existing similarity-based models of categorical induction cannot adequately account for the performance of populations with domain-relevant experience. Other research has suggested that the relations among the premises of an inductive argument may provide relevant information that further constrains inductive inferences. This study uses an open-ended task to examine induction children with differing experience with the biological world. 85 school-aged children from diverse environments in Massachusetts were given 18 pairs of premise animal categories and taught that each pair shared a property (disease or "stuff inside"). Pairs were either taxonomically close or far and were either related by predation, shared habitat, or unrelated. We asked participants to project what other kinds of things would share the property and explain their answers. Responses were coded as similarity-based (related to class inclusion or shared features) or interaction-based (related to contact through environment). Results suggest that children use inductive selectivity, making more similarity-based responses when reasoning about "stuff inside" and more interaction-based responses when reasoning about a disease. Children from rural environments also used the relations among premises to constrain their reasoning whereas urban children did not. Reliance on similarity-based reasoning decreased with age for all children and rural children showed an increase in the use of interaction-based reasoning. Taken together the results suggest that children's experience plays a role in the development of sensitivity to relevant factors that constrain inductive reasoning.

Introduction

In our everyday life we are often unaware of how confidently we navigate a world full of novelty and uncertainty. One way we do this is to extend our existing categorical knowledge about the world to new kinds of things and we do so in a principled way. This process is commonly called category-based induction and such an argument generally takes the form "A's all have property X therefore B's must have property X." How we construct such arguments and how we determine the set of B's to generalize property X over has been the focus of research for many years.

The study of category-based induction has often focused on the domain of biology. The natural world is ubiquitous and well known to children (just browse a handful of children's books). There are several reasons why this

domain is well suited for the study of induction. For one, biological categories are naturally structured in a hierarchical taxonomy so a poodle called "Fluffy" is necessarily also a dog, a mammal and an animal. Another attractive aspect of the domain is the existence of naturally orthogonal alternative category structures; natural categories can be organized by the relations between them, such as habitat (e.g. polar animals, jungle animals, etc.) or biological niche (e.g. predation).

Researchers have attempted to model category based induction. One attempt, the Similarity Coverage Model (Osherson et al. 1990) proposed a specific mechanism of similarity to explain induction. In this case similarity refers to the idea that inductive arguments are strong as a function of the similarity between the premise and conclusion categories. For example, if you know that robins have sarca then you might be more likely to conclude that sparrows might have sarca rather than penguins since robins are more similar to sparrows than penguins. On the other hand the principle of coverage concerns the representativeness of a premise category as a function of the conclusion category. In this example if you know robins have *sarca* you might be more willing to conclude that all birds have *sarca* than if you knew that penguins have *sarca*. According to the SCM this is a result of the fact that robins are more similar as a whole to all instances of the category bird (robin *covers* the category) than penguins. The SCM has had a great deal of success predicting the reasoning of university undergraduates, who are the primary population of this research, however it fails to model results looking at induction in alternative populations. In fact research looking at the Itza' Maya of Guatemala (Lopez et al. 1997), Chicago landscapers (Medin et al. 1997), and commercial fisherman in Boston, among others, has shown that factors like expertise can influence category-based induction to diverge from reasoning predicted by the SCM.

One way to address these differences in reasoning is to consider what the characteristics of expert-like reasoning might be. Several converging findings suggest that experts use their specific knowledge about a domain to determine the relevant dimensions of comparison used in their generalizations. For instance, Lynch et al. (2000) found that in a comparison of tree experts and undergraduates reasoning about trees, experts used the dimensions of height and "weediness" as the organizing dimension for the category of trees. Shafto and Coley (2003) found that commercial fisherman used their specific knowledge of

local marine habitats and food chains to make inductive inferences about disease transmission.

Another hallmark of expert-like reasoning is the use of inductive selectivity. Generally inductive selectivity refers to the notion that the property being projected provides an important context for an inference. For instance, if you know that penguins have *hemo* inside, you might infer that this is a property shared by flightless birds. In contrast if you know that penguins have the disease *hemo* you might be more likely to infer that seals might get *hemo* since they live with penguins. The consideration of property of projection as a constraining factor on induction can be seen as a way of integrating specific knowledge within the context of the induction problem. For instance Shafto and Coley also found that commercial fisherman used different relations between marine creatures as a function of the property of projection. Specifically when asked to generalize a blank property across marine categories fisherman tended to rely on taxonomic relations; however, if the property was a novel disease they used their specific knowledge of the local ecology to generalize along ecological dimensions such as food chains.

Apart from expertise effects, several other factors have been shown to contribute to differences in category-based induction. Medin et al. (2003) proposed a framework for investigating induction that centers on the idea of relevance. Briefly, relevance refers to the fact that people are sensitive to the relations between premise and conclusion categories and assume that these relations are informative with respect to strength of the conclusion. Specifically they offer two related principles that might guide induction. First that the projection of a property from premise categories is associated with the most distinctive feature shared by the premise categories. For instance, if you are told that penguins have *andro* you might assume that having *andro* is related to living in cold climates. The second principle concerns the relationship between the premise categories and the conclusion category. Specifically that comparing the premise and conclusion categories should further constrain the relevant dimensions of induction. For instance if you were told that penguins have *andro* therefore zebras have *andro* you might augment your earlier conclusion and believe that *andro* is actually a property related to being black and white. Likewise adding more premise categories such as penguins and pandas have *andro* should make you even more certain of your conclusion.

Using this framework we can look to phenomena in category based induction that are not accounted for by the SCM. Divergence from reasoning predicted by the similarity coverage model has also been shown in children. Ross et al. (2003) compared the performance of urban and rural majority culture and rural Native American children reasoning about animals. Native American and older children tended to give justifications for their projections for items like *bee* and *bear* in terms of the ecological relations between them. Their results suggest that young school age children can also display expert-like use of specific domain

knowledge and inductive selectivity. Furthermore, Vitkin, Coley & Kane (2004) investigated children's inductive reasoning about plants and animals. Using a forced choice triad task. They taught children a property (disease or internal substance) about a premise category and asked children which of the two animal categories (related by common superordinate category or ecology) might share this property. Children were more likely to select ecologically related items when reasoning about diseases and more likely to select taxonomic alternatives when reasoning about internal substances. Taken together these results suggest that young children also use inductive selectivity when alternative relations are familiar.

Research with both children and adults looking at induction has mainly employed forced choice methods. Often this takes the form of an inductive premise with two alternative conclusions. The relations between the premise and the alternatives are manipulated to contrast inferences based on taxonomic similarity and ecological relations. These methods can potentially obscure the complexity of the reasoning process. For instance it is not clear what salient relations determine the choice between any two alternatives. The inference also depends on whether or not participants are aware of the relations between the given alternatives. Results from forced choice methods also necessitate an inverse relationship between inference kinds since choosing one alternative means not choosing the other. Although theories of induction in the biological domain predict such a pattern of results it is not clear if these results are an artifact of task demands.

Little is known about the extent to which children spontaneously use knowledge of the natural world to selectively constrain category-based induction. In the present research we have attempted to address the problems inherent in forced choice tasks by employing an open-ended induction procedure. By providing the premises of an inductive argument and asking children to provide both the conclusion categories and justifications for their projections we hope to capture the robustness and complexity of children's knowledge and their sensitivity to different kinds of relations among biological categories.

To this end we employed premise pairs that differed on two orthogonal dimensions: *ecological relatedness* (predatory/prey, habitat, no relation) and *taxonomic distance* (same or different superordinate category). We asked children to reason about a novel disease or "stuff inside" and subsequently analyzed the basis of their projections. Using this method we were able to investigate whether under little task constraint children would spontaneously use specific knowledge of biological relations to guide their inferences. Specifically, we examine the following questions: (1) Do inferences about disease differ systematically from inferences about insides? (2) Do children spontaneously utilize relevant relations among premise pairs to constrain inferences? (3) Do these sensitivities change with age or with opportunities for direct experience with plants and animals?

Methods

Participants

A total of 85 children in kindergarten through 6th grade were recruited through elementary schools and after-school programs in 9 communities in Massachusetts. Participants ranged in age from 5-11 to 12-8 with a mean age of 8-10. Communities ranged in population density from 22 to 13488 people per square mile.

Materials and Design

Stimuli consisted of 36 5 x 7 in. laminated cards. Each card depicted detailed color line drawings of two animals, arranged vertically. We manipulated the relations between the base pair of animals on two dimensions, taxonomic distance and ecological relatedness. Pairs were either taxonomically *close* (same superordinate category) or *far* (different superordinate category). Each pair of animals was also ecologically related by *predation* (one of the animals eats the other), *habitat* (both animals are found in the same place), or *unrelated* (have no discernable ecological relationship). See Table 1 below for examples. These relations were varied within subjects and each participant saw 3 pairs from each cell in Table 1, for a total of 18 items. Animals and relations among them were carefully chosen to be familiar to a wide range of children. Two sets of items were used; one depicted local species, and the other depicted exotic species. Children were randomly assigned to one of these conditions.

For each item children were taught a property about the premise pair of animals and were asked what other kinds of things might also share this property and why. We varied the property children were asked to project between subjects. One group reasoned about a novel physio-anatomical property (“have a stuff inside called *cyto*”) and the other group reasoned about a novel disease (“have a sickness called *cyto*”).

Procedure

Children who received parental consent were interviewed individually at their school or after-school program. They began with a warm-up task asking children about their hobbies and interests. Children then performed this and two related tasks in counterbalanced order across subjects. Overall children only saw any given species or relation once across all tasks.

Table 1: Example of Manipulated Relationships Among Premise Pairs

Ecological Relations	Taxonomic Relations	
	Close	Far
Predatory/Prey	<i>Spider / Fly</i>	<i>Salmon / Black Bear</i>
Habitat	<i>Moose / Porcupine</i>	<i>Lobster / Tuna</i>
No Relation	<i>Newt / Box Turtle</i>	<i>Raccoon / Pelican</i>

For this task children were shown 18 cards containing inductive pairs in random order. For each pair children were told, “There’s this stuff / sickness called X. Lots of things can have X. A’s and B’s have X. What other kinds of things do you think might also have X and why?” where A and B are animals and X is the property of projection (disease or stuff inside). Responses were tape recorded and transcribed for coding.

After the data was collected four trained coders classified each child’s responses to each item. For each response coders evaluated the kinds of relations that formed the basis of the projection from the premises as being. For purposes of this paper, we consider two broad classes of responses. Projections were coded as *similarity based* if the property was projected on the basis of shared features, overall similarity, or shared category membership. Projections were coded as *interaction based* if the property was projected on the basis of spatio-temporal interactions among species. See Table 2 for sample responses. A single response could receive multiple codes.

Table 2: Sample Similarity-based and Interaction-based Responses

Type of Response	Examples
Similarity-based	-- <i>Monkeys because monkeys come from the same family as a gorilla.</i> -- <i>Foxes in the arctic because they have hair like a polar bear.</i>
Interaction-based	-- <i>Fish because newts and box turtles bring the disease to the water.</i> -- <i>Field mice because owls might eat them.</i>

Results

For each subject there were 3 items of each type. We computed the number of similarity-based and interaction-based responses for each item type. Scores could be higher than 3 because of multiple subcategories of similarity-based and interaction-based projections. Preliminary analyses revealed no differences between response patterns for local versus exotic pairs, so we do not analyze that factor further.

For clarity of presentation, children were divided into urban and rural groups based on population density; rural children came from communities ranging from 22-245 people per square mile, and urban children came from communities ranging from 6851 to 13488 people per square mile. We also performed a median split on age to derive older and younger age groups. The younger group ranged from 5-11 to 8-4 with a mean age of 7-7, and the older group ranged from 8-4 to 12-8 with a mean age of 10-0. We analyzed patterns of similarity-based and interaction-based projections separately for urban and rural children, conducting four separate 2 (Age: younger, older) x 2 (Property: disease, stuff inside) x 2 (Taxonomic Distance:

close, far) x 3 (Ecological Relation: habitat, predation, unrelated) mixed ANOVAs on frequency of response.

If children exhibit inductive selectivity, we would expect relatively high levels of similarity-based inferences from children reasoning about *insides*, and relatively high levels of interaction-based inferences for children reasoning about *disease*. If children are sensitive to relevant relations among premise pairs, we expect more similarity-based projections from taxonomically close than far pairs, and more interaction-based projections from ecologically related than unrelated pairs. If greater potential for interaction with living things increases context-sensitive induction, then rural children might be more sensitive to manipulation of relations among premise pairs than urban children. Finally, we are also interested in differences between younger and older children.

Urban Children

Similarity-based Projections

Urban children showed use of inductive selectivity, they made more similarity-based projections in the *insides* ($M = 3.26$) than *disease* ($M = 1.94$) condition ($F(1, 37) = 10.09$, $MSe = 120.86$, $p = .003$).

We also found a main effect of age ($F(1, 37) = 4.25$, $MSe = 50.93$, $p = .05$) such that younger children relied on similarity-based inference ($M = 2.86$) more than older children ($M = 2.05$).

We observed no effects of taxonomic distance between premises or ecological relatedness on similarity-based projections.

Interaction-based Projections

Urban children's interaction-based projections were also selectively sensitive to the kind of property being projected ($F(1, 37) = 22.54$, $MSe = 164.59$, $p < .0001$). Interaction-based projections were more common for the *disease* condition ($M = 1.83$) than the *insides* condition ($M = .26$).

We observed no developmental effects of age on the frequency of interaction-based reasoning. We also observed no effects of taxonomic distance between premises or ecological relatedness for interaction-based projections

Rural Children

Similarity-based Projections

Rural children also showed evidence of inductive selectivity; there was a main effect of property condition, *disease* or *insides*, ($F(1, 40) = 17.82$, $MSe = 45.25$, $p = .01$). Rural children's similarity-based inferences were also constrained by the kind of property they were reasoning about with more similarity-based projections for the *insides* ($M = 3.30$) than *disease* ($M = 1.92$) condition.

Like urban children, rural children also showed a decline in the use of similarity-based reasoning with age ($F(1, 40) = 7.07$, $MSe = 45.26$, $p = .01$). Younger children made more

similarity-based projection ($M = 3.22$) than older children ($M = 2.28$).

Unlike urban children, we observed a main effect of taxonomic distance between premises, as predicted by the relevance framework. for rural children ($F = 4.61$, $MSe = 5.59$, $p = .04$). Rural children made more similarity-based projections for taxonomically close items ($M = 2.88$) than taxonomically far items ($M = 2.53$). This main effect was further qualified by an interaction between taxonomic distance and age ($F(1, 40) = 4.19$, $MSe = 5.08$, $p = .05$) revealing that older children's sensitivity to manipulation of taxonomic distance was driving the effect (See Figure 1). Rural children's similarity-based projections were not affected by our manipulation of ecological relatedness among premises.

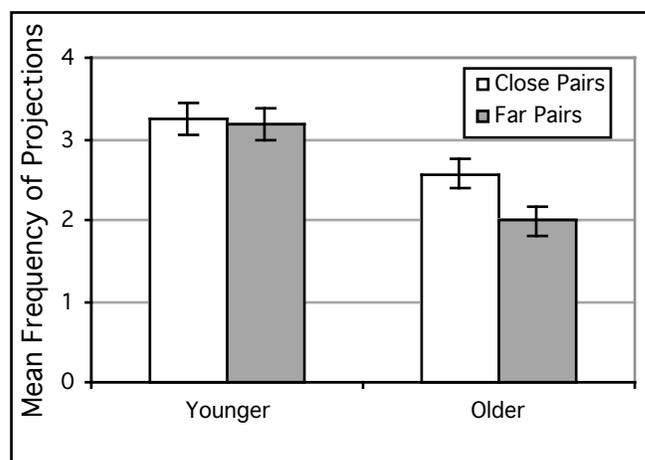


Figure 1. Mean frequency of Similarity-based Projections by Age and Taxonomic Distance for Rural Children

Interaction-based Projections

Rural children's interaction-based projections also showed evidence of inductive selectivity. There was a main effect of property condition ($F(1, 40) = 8.83$, $MSe = 61.86$, $p = .01$) such that rural participants made more interaction-based responses in the *disease* ($M = 1.77$) than *insides* ($M = 0.71$) condition.

There was also a main effect of age ($F(1, 40) = 6.21$, $MSe = 43.53$, $p = .02$) such that older children made more interaction-based projections ($M = 1.56$) than younger children ($M = 0.71$).

We found a main effect of taxonomic distance for interaction-based projections ($F(1, 40) = 6.30$, $MSe = 3.82$, $p = .02$). Rural children made more interaction-based projections for taxonomically far items ($M = 1.32$) than taxonomically close items ($M = 1.02$).

Unlike urban children, we observed a main effect of ecological relations among premises for rural children ($F(2, 80) = 3.40$, $MSe = 1.82$, $p = .04$). As expected rural children showed more sensitivity to the ecological relations among

premises such that items containing predator / prey relations had significantly more interaction based-projections ($M = 1.35$) than both items related through habitat ($M = 1.08$) and unrelated items ($M = 1.08$).

Discussion

The purpose of this study was to use a new methodology to explore the role of relevance-based reasoning in children's category based induction. Overall our findings suggest that children are in fact sensitive to the context provided by an inductive argument to constrain their reasoning. Furthermore this sensitivity develops as a function of experience afforded by a child's local environment.

Our results support the notion that children use inductive selectivity to guide their reasoning. Effects of condition (reasoning about disease or "stuff inside") were seen for both similarity and interaction based projections for both urban and rural children. Specifically children reasoning about disease used fewer similarity-based justifications than those reasoning about an internal substance. Conversely more interaction-based projections were observed for children reasoning about disease than those reasoning about internal substances. These results suggest that children can constrain their inferences based on the nature of the property of projection. Reasoning about disease seems to make ecological relations relevant to the inference whereas internal substance provides cues for the relevance of taxonomic relations.

The relevance framework offers another constraint on category-based induction. It suggests that people are sensitive to relations among a set of premises and that these relations might further constrain their conclusion. With this in mind we manipulated taxonomic and ecological relations among premise pairs. We found that rural children in particular showed a systematic sensitivity to the relations among premises while urban children did not.

In particular we found that taxonomic distance between premises was a relevant cue for rural children's responses. Similarity-based responses were more common for taxonomically close than taxonomically far premise pairs suggesting that taxonomic closeness was a salient cue for making generalizations over this dimension. We also found that rural children made more interaction-based responses for taxonomically far than close pairs suggesting that in the absence of strong taxonomic cues rural children relied on relevant ecological knowledge to guide their projections.

With regard to ecological relations among premises we observed that rural children made more interaction-based projections when the premises were related by predation than both habitat-related and unrelated items. Ecological relations among premises did not have an effect on similarity-based projections. Ecological relations among premises were only relevant in the context of reasoning based on interaction through ecology and did not affect similarity-based responses.

As previously stated urban children did not show any sensitivity to manipulation of relations among premises. There was no difference in similarity- or interaction-based reasoning as a function of either taxonomic or ecological relatedness of the premise pairs. This motivates questions about the contribution of experience in the local environment to the use of relevant relations to guide reasoning. We found significantly different patterns of reasoning for urban and rural children. In particular, as we predicted the opportunity to directly experience relatively intact ecosystems did promote sensitivity to relevant relations among biological kinds.

Finally, we also found evidence for the developmental emergence of reasoning predicted by the relevance framework. Generally older children relied less on similarity-based explanations for their projections, they made fewer similarity-based projections than younger children. It may be the case that as children get older they move away from a reliance on similarity cues to guide their reasoning. However, only rural children showed an increase in the use of interaction-based explanations as they get older. This suggests that experience may play a role in the emergence of ecologically informed reasoning.

Previous accounts of category-based induction may have overestimated children's reliance on similarity in their reasoning. This study suggests that children are sensitive to both taxonomic similarity and ecological relations and take into account the properties they are reasoning about to determine which relations are relevant to their inferences. Both urban and rural children showed evidence of inductive selectivity, however only rural children were sensitive to taxonomic and ecological relations among premises. The emergence of this sensitivity to relevant information used to inform category-based induction appears to develop differently in children as a function of their environment. Using an open-ended task we hope to have captured a more veridical picture of how children use existing contextual factors to facilitate their reasoning about the domain of biology.

Acknowledgments

This research is funded by National Science Foundation Grant # BCS-0236338 to J.D. Coley. We would like to thank Robin Kane, Claire Seaton, Jennelle Yopchick, Konstantin Feigin, Allison Baker, and the rest of the Categorization and Reasoning Lab for their tireless work and incalculable contributions to this project. We would also like to acknowledge the support of participating administrators, teachers, parents, and children, without whom this research would not be possible.

References

- Lopez, A., Atran, S., Coley, J., Medin, D., Smith, E. (1997). The tree of life: Universals of folk-biological taxonomies and inductions. *Cognitive Psychology*, 32, 251-295.
- Lynch, E. B., Coley, J. D., & Medin, D. L. (2000). Tall is typical: Central tendency, ideal dimensions, and graded

- category structure among tree experts and novices. *Memory & Cognition*, 28, 41-50.
- Medin, D. L., Lynch, E. B., Coley, J. D. & Atran, S. (1997). Categorization and Reasoning among Tree Experts: Do All Roads Lead to Rome? *Cognitive Psychology*, 32, 49-96.
- Medin, D. L., Coley, J. D., Storms, G., Hayes, B. K. (2003). A Relevance theory of induction. *Psychonomic Bulletin & Review*, 10(3), 517-532.
- Osherson, D. N., Smith, E. E., Wilkie, O., Lopez, A., & Shafir, E. (1990). Category-based induction. *Psychological Review*, 97, 185-200.
- Ross, N, Medin, D., Coley, J. D., Atran, S. (2003). Cultural and experiential differences in the development of folkbiological induction. *Cognitive Development*, 18, 25-47.
- Shafto, P., Coley, J. D. (2003). Development of Categorization and Reasoning in the Natural World: Novices to Experts, Naïve Similarity to Ecological Knowledge. *Journal of Experimental Psychology*, 29, 641-649.
- Vitkin, A., Coley, J.D. & Kane, R. (2004). *The Role of Experience in the Development of Biological Reasoning in Children*. Paper presented at the 45th Annual Meeting of the Psychonomic Society, Minneapolis, MN.