

GENERATING INDUCTIVE INFERENCE: PREMISE RELATIONS AND PROPERTY EFFECTS

John D. Coley *and* Nadya Y. Vasilyeva

Contents

1. Introduction	184
1.1. Inference Generation	184
1.2. Induction and Relations among Concepts	187
1.3. Goals of this Chapter	190
2. Effects of Premise Relations on Inference Generation	191
2.1. The Role of Premise Relations in Argument Evaluation	191
2.2. Relative Salience of Conceptual Relations	193
2.3. Study One: Investigating Effects of Premise Relations on Inference Generation	194
2.4. Summary: Effects of Premise Relations on Inference Generation	202
3. Effects of Property on Inference Generation	203
3.1. Property Effects in Argument Evaluation	203
3.2. Study Two: Investigating Effects of Property on Inference Generation	205
3.3. Summary: Effects of Property on Inference Generation	216
4. Inference Generation: Conclusions and Implications	217
4.1. What Have We Learned About Inference Generation?	217
4.2. Implications	219
4.3. Conclusions	224
Acknowledgments	224
References	224

Abstract

Categorical inductive inference is the process by which we project features believed to be true of one class to another related class. Traditional approaches to studying inductive inference have focused on the evaluation of inductive arguments. In this chapter, we introduce a new approach by examining the way people *generate* inductive inferences. We focus on how relations among premise categories, and the nature of the property being projected, impact the kind

of inferences generated. Participants were taught that two animal species shared a novel *substance, disease, or gene*, and were asked what other species might also have the property, and why. Results show that people attend to salient relations between premise categories, determine their relevance based on the property they are asked to project, and then generate inferences consistent with those relations. Participants drew a broad range of inferences based on taxonomic similarity, contextual similarity, and causal relations. Inference generation was constrained both by salient premise relations and the nature of the projected property. We discuss how these findings expand the list of challenges for the models of induction, question the primacy of taxonomic relations in guiding inductive inference, encourage further investigation into the process by which inductive inferences are generated, and emphasize the knowledge-driven and flexible nature of human inductive reasoning.

1. INTRODUCTION

Categorical induction is the process by which we project features believed to be true of one class to another related class. This process is essentially knowledge-driven; when we learn that A has a novel property, we use what we know about the relations between A and B to compute the likelihood that B will also have the property. Unlike deductively valid arguments like *Socrates is a man, all men are mortal, therefore Socrates is mortal*, inductive arguments like *Coatimundis have disease X, therefore wombats have disease X* are not inherently valid or invalid. Rather, they are strong or weak to the degree that they are supported by relevant knowledge. If we know something about relations between coatimundis and wombats that would warrant a common affliction, this argument may seem strong, but that strength derives entirely from our prior knowledge.

The preceding is a more or less standard opening paragraph for a paper on inductive reasoning. And make no mistake; we stand by it. However, our goal in this chapter is to present a decidedly nonstandard approach to understanding inductive inference. Specifically, instead of looking for regularities in the ways in which people evaluate inductive arguments, we focus on the ways in which people generate inductive inferences.

1.1. Inference Generation

Most previous research on inductive reasoning has involved the evaluation of complete inductive arguments in one form or another. Typically, participants are presented with one or more premises in which a property is attributed to a category, along with a conclusion in which the property is attributed to a different category, and asked to evaluate the argument, or rate the likelihood that the conclusion is true given that the premises are true.

This general approach can take several forms. Participants might be presented with an entire argument (e.g., *Frogs have property X, therefore toads have property X*) and asked to evaluate its strength. Alternatively, participants might be asked to make yes/no judgments about whether one category has a property given that another one does (*Frogs have property X. Do toads have property X? Do raccoons have property X? Fish?*). Another popular format involves presenting a premise and forcing a choice of the better conclusion from among two or more alternatives (*Frogs have property X. Is it more likely that raccoons or toads also have property X?*). Our intention is not to call for the abolition of argument evaluation as a way of understanding inductive reasoning. We have learned a great deal about categorical induction from argument evaluation; it has revealed that inductive reasoning is both systematic and flexible (see [Heit, 2000](#), for a review). We have used argument evaluation ourselves, and plan on continuing to do so. Rather, the point we want to make is that we cannot and should not base a psychology of inductive reasoning solely on argument evaluation.

Argument evaluation depends on the participant recognizing and evaluating specific hypotheses generated by the experimenter. The participant is required to compare premise and conclusion categories, notice the relation(s) under investigation or lack thereof, evaluate their relevance to supporting an inference about the property—which may or may not be informative—and then render a judgment. As such, this approach is akin to a recognition memory task or a multiple-choice exam where people choose the best answer from a pool of prefabricated alternatives. We believe that this approach potentially misses an entire spectrum of inductive phenomena in which people generate plausible inferences from given information rather than evaluate hypotheses that are given to them.

To begin to explore this spectrum of inductive reasoning, we utilized a novel open-ended induction task.¹ Rather than evaluating the strength of arguments composed of premises and conclusions, or making a forced choice from among a limited set of alternatives, we allowed participants to generate open-ended projections about novel properties from pairs of premise categories (e.g., *A and B have property X. What else do you think would have property X, and why?*). In other words, we present participants with premises and ask them to generate their own conclusions. As such, our approach is more like a recall memory task or an essay exam. We manipulated the relations between A and B, and made corresponding predictions about the ensuing inferences if such relations guide inductive projections. This allowed us to map out the ways in which participants' beliefs about relations between premise categories, and the property we ask them to reason about, influence the kinds of inferences they spontaneously generate. This approach has a number of advantages.

¹ This methodology is not entirely novel. [Proffitt, Coley and Medin \(2000\)](#) in their Experiment 3, report results from a similar method, but present it in a different framework.

First, it represents (as far as we know) the first attempt to look at how people actually construct inductive inferences. Any viable theory or model of induction must explain not only how people evaluate the strength of arguments, but more centrally, how they generate novel inductive inferences. This chapter does not propose such a model or theory, but we believe that it represents a valuable contribution to the stockpile of raw materials for such a theory or model.

Second, asking participants to generate open-ended inferences provides them a greater opportunity to draw on knowledge they spontaneously deem relevant, rather than depending on the experimenter to anticipate potential responses. Likewise, the open-ended method is less likely to artificially force a choice between basing an inference on one of two salient relations when both may seem plausible. Instead, participants are free to base inferences on multiple salient relations, if they so desire.

Third, as we shall see, this approach allows participants to utilize relatively abstract or inchoate knowledge to guide inductive inference. As Keil (2003) has shown, our explanatory knowledge and understanding of causal mechanisms are often much more superficial and vague than we know or would like to admit. Nevertheless, inductive inferences can bridge gaps in specific factual knowledge—indeed, making uncertain guesses about the unknown is what induction is all about. Many researchers have suggested that fairly abstract principles (domain theories, schemata) provide inductive constraints in concept learning (e.g., Heibeck & Markman, 1987; Keil, 1981) language acquisition (Chomsky, 1980), and inductive inference (e.g., Coley, Hayes, Lawson, & Moloney, 2004; Coley, Medin, & Atran, 1997; Goodman, 1955). As demonstrated by Kemp, Perfors, and Tenenbaum (2007), it is possible to learn abstract knowledge from observations before acquiring specific knowledge at lower levels of abstraction. If so, inductive reasoning may be especially likely to rely on abstract ideas in the absence of specific knowledge.

For example, if you learn that coatimundis have a particular disease, you might think it likely that anything that eats coatimundis might also have the disease, even if you have no idea what might do so. In an open-ended response format, you might confidently assert that “Anything that preys on coati-whatevers would also have disease X” and even explain your response along the lines of “because they could contract it from eating tainted coati meat.” However, if presented with an argument like *Coatimundis have disease X, therefore ocelots have disease X*, you might not know that ocelots are potential coati predators, and therefore might simply shrug your shoulders and rate the argument as relatively unlikely. Because you were unable to apply your (relatively abstract) knowledge about disease to evaluate the argument, you rated it as weak when in fact you believed it to be strong, but just didn't realize it.

Fourth, an open-ended format increases the ecological validity of induction research by capturing the previously unacknowledged open-ended

nature of everyday inductive reasoning. Seldom in everyday life are we asked to evaluate fully formed arguments; rather, we have some facts, and are free to generalize from those facts as we see fit. For example, I may know that butter and bacon are to be avoided on a low cholesterol diet; it seems more ecologically valid for me to generate inferences about “what else should be avoided” than to evaluate a series of arguments such as *Butter and bacon have high cholesterol, therefore lettuce has high cholesterol. Butter and bacon have high cholesterol, therefore lamb has high cholesterol. . .*

In sum, by focusing on how people generate, rather than evaluate, inductive inferences, we allow participants to draw upon any knowledge that they deem relevant in the context of an ecologically valid inductive problem. Our goal is to use an open-ended inductive inference task to take a detailed look at how people spontaneously recruit different kinds of knowledge to generate inductive inferences.

1.2. Induction and Relations among Concepts

A central problem in inductive reasoning is deciding what knowledge should inform a given inference. There is a vast amount of information associated with any given concept that can be used in induction. Some information is more likely to support inferences than other information; it is safe to say that a person who knows that a frog is an amphibian and who has seen a frog yesterday is more likely to project a novel property learned about a frog to other amphibians than to other things seen yesterday. In the following we will discuss three kinds of conceptual relations that have been shown to be relevant to evaluating the strength of inductive arguments: taxonomic (intrinsic) similarity, contextual (extrinsic) similarity, and causal relatedness.

1.2.1. Taxonomic Similarity

Similarity is an intuitively appealing candidate for guiding induction. It plays an important role in many natural categories, and categorizing novel instances based on their similarity to known category exemplars or to a category prototype provides a basis for forming novel expectations about them. And indeed, taxonomic, or intrinsic, similarity—especially in the form of common category membership—has repeatedly been shown to be an especially strong predictor of the strength of an inductive argument. For example, the argument *Frogs have property X, therefore toads have property X* might seem strong because frogs and toads are both amphibians; since they are similar kinds of animals that share many known features, it is likely that they will share a newly learned feature as well. Reasoning of this kind is well described by models of category-based induction that emphasize the importance of similarity, shared features and/or common category membership in guiding inductive inferences (e.g., [Osherson, Smith, Wilkie, Lopez, &](#)

Shafir, 1990; Rips, 1975; Sloman, 1993; Sloutsky & Fisher, 2004). According to such models, arguments are perceived as strong to the degree that premises and conclusions are generally similar, share specific features, or belong to a common taxonomic class.

However, it turns out that this notion of similarity may be both too broad and too narrow to provide a sufficiently detailed and exhaustive account of the range of inductive inferences people make. The construct is too broad because people make use of more precise subtypes of intrinsic similarity. One influential demonstration of this was done by Heit and Rubinstein (1994) who found that people prefer to base inferences on anatomical similarity (e.g., projecting a property from a bat to a mouse) or on behavioral similarity (e.g., projecting a property from a bat to a sparrow) depending on the property being projected. As the authors point out, these results logically eliminate the possibility that a single similarity measure can predict inferences about animals. At the same time, the construct is too narrow because intrinsic taxonomic similarity, even broken down into different subtypes, does not exhaust the types of similarity available for guiding inductive inferences.

1.2.2. Extrinsic Similarity

In addition to being a member of a certain taxonomic class, frogs participate in a number of contextual relationships with other species that could potentially provide a basis for induction. For example, a property known to be true of frogs might reasonably be extended to fish; in other words, the argument *Frogs have property X, therefore fish have property X* might seem strong not because frogs are generally similar or taxonomically related to fish, but because they share one potentially important extrinsic contextual feature—an aquatic environment. That is, if one is aware of a specific contextual relation between premise and conclusion, that relation can increase the perception of argument strength to the degree that it is deemed relevant.

There is evidence that adults utilize extrinsic similarity to evaluate inductive arguments. For example, when making categorical inferences about food, people use both taxonomic categories (like *fruit* or *meat*)—based on shared intrinsic features or composition—and script categories (like *breakfast foods*)—based on time, location, or setting in which the foods are eaten—to evaluate potential inductive projections (Nguyen & Murphy, 2003; Ross & Murphy, 1999; Vitkin, Coley & Feigin, 2005). Lin and Murphy (2001) demonstrated that participants view a wide range of thematic relations (e.g., camel–desert, cat–litter box, bees–honey, Michael Jordan–basketball, Hawaii–beach) as plausible bases for inferences. Shafto, Coley, and Baldwin (2007) showed that when people reason about animals, inferences are strengthened by extrinsic relations—specifically shared habitat—between premise and conclusion species, as well as by common membership in taxonomic categories.

In sum, although extrinsic contextual similarity has received less attention than intrinsic taxonomic similarity from researchers studying categorical induction, it represents an important alternative basis for inductive inference.

1.2.3. Causal Relations

Although adding extrinsic similarity to our list of potential bases for induction is a step in the right direction, it is important to point out that similarity, however flexibly construed, does not exhaust the kinds of knowledge potentially relevant to guiding inductive inference. For example, we also possess causal knowledge about the way frogs interact with other species and their environment. For instance, if you learned that frogs have a property, you might infer that raccoons would also have this property, knowing that because raccoons eat frogs, they could potentially contract the property through ingestion. That is, if one is aware of a causal chain linking premise to conclusion, such as a food chain relation, it can inform evaluation of an inductive argument. In other words, the argument *Frogs have property X therefore raccoons have property X* is potentially strong not because frogs and raccoons are similar in any way, but because we have knowledge of a causal chain that links the two and is potentially relevant to property projections. In support of this idea, [Medin, Coley, Storms, and Hayes \(2003\)](#) demonstrated sensitivity to causal relations between premises and conclusions in a number of ways. For example, participants rated arguments where premise and conclusion were taxonomically dissimilar but shared a salient causal relation (e.g., *Bananas have property X therefore monkeys have property X*) to be as strong as arguments where premise and conclusion were taxonomically more similar but causally unrelated (e.g., *Mice have property X therefore monkeys have property X*). Salient causal relations also lead people to commit the conjunction fallacy ([Tversky & Kahneman, 1973](#)) by rating arguments with a conjunctive conclusion emphasizing a causal chain (e.g., *Grain has property X therefore mice and owls have property X*) as stronger than arguments with a single constituent category as a conclusion (e.g., *Grain has property X therefore owls have property X*). [Feeney, Shafto, and Dunning \(2007\)](#) replicated this inductive conjunction fallacy effect, and showed that causal relations led to stronger and more persistent fallacies than taxonomic relations.

Reliance on causal relations in reasoning has been shown to increase with relevant expertise. For example, [López, Atran, Coley, Medin, and Smith \(1997\)](#) found that Itza' Maya, indigenous people of Guatemala who rely on hunting and agriculture and live in close contact with nature, when asked to evaluate inductive arguments about local species, appeal to specific causal ecological relations between animals. [Proffitt, Coley, and Medin \(2000\)](#) demonstrated a similar effect with North American tree experts who were asked to reason about inductive problems involving disease distribution among trees. Rather than appealing to overall or categorical similarity of tree types, tree experts used their knowledge to construct

sophisticated explanations of how diseases might be transmitted from one tree to another. Likewise, [Shafto and Coley \(2003\)](#) showed that when projecting novel diseases among local marine species, commercial fishermen used causal knowledge of food webs to evaluate arguments.

However, even relative novices (undergraduates) actively use causal relations to evaluate arguments when tested about familiar categories (e.g., [Feeney et al., 2007](#); [Medin et al., 2003](#)) or when specifically trained about novel causal systems ([Shafto, Kemp, Bonawitz, Coley, & Tenenbaum, 2008](#)). Moreover, the expectation that causal relations provide a useful basis for inferences is present early; [Muratore and Coley \(2009\)](#) showed that 8-year-old children, when they have necessary knowledge about ecological interactions between animals, use causal information to make inferences. As demonstrated by [Sloman \(1994\)](#), inductive arguments can spontaneously trigger causal reasoning. When participants could construct a single explanation of why both premise and conclusion have a property, arguments were seen as more plausible than when two separate explanations were required to connect property to the premise and to the conclusion.

In sum, people use a variety of conceptual relations to evaluate categorical inductive arguments. Taxonomic similarity—based on shared category membership and/or shared intrinsic features—is one common metric, and it has been widely studied and modeled. However, extrinsic similarity—based on shared context, or common links to the outside world—and causal relatedness—coherent causal pathways that could explain how or why a property is shared by premise and conclusion categories—are also potentially powerful guides for inductive inference. In this chapter, we examine factors that impact the frequency with which people generate inferences based on these three kinds of relations.

1.3. Goals of this Chapter

We have argued that people use taxonomic, contextual, and causal relations among categories to evaluate inductive arguments. What factors guide recruitment of these relations during inference generation? In this chapter we utilize a novel open-ended induction task to examine how people generate inductive inferences about plants and animals. We chose the domain of folk biology because there is a rich literature on inductive reasoning using biological categories, and because the different kinds of salient and potentially orthogonal relations among living things (e.g., biological family, behavior, shared habitat, ecological niche, predator–prey) naturally lend themselves to supporting a range of different kinds of inferences.

In particular, we examine how people use intrinsic taxonomic similarity, extrinsic ecological similarity, and causal relations to generate inferences about what animal species are likely to share novel properties. We focus on the following questions: First, to what degree do salient relations among

premise categories determine the nature of inferences generated from those categories? In general, we expect that participants will compare premise categories, extract salient relations, and generate inferences consistent with those relations. Thus, premises with salient taxonomic relations will yield taxonomic inferences, whereas premises with salient spatiotemporal or causal relations will yield corresponding inferences. Second, to what degree does the property being projected influence the process of inference generation? Previous research suggests that property can have a large effect on how arguments are evaluated. We know little about the effects of property on how arguments are generated; of particular interest is whether property serves as an overall biasing factor, or whether it changes the salience of particular premise relations in guiding inferences.

Overall, our goal is to extend the knowledge base available to inform theories and models of inductive reasoning. We have no theoretical axe to grind; rather, we seek to expand the range of phenomena on inductive inference that any theory or model must account for by examining the nature of spontaneous inductive projections and explanations for those projections. In the next section, we consider the effects of relations among premise categories on inference generation. We review previous studies suggesting that salient relations—especially taxonomic relations—among premise and/or conclusion categories influence perceived strength of inductive arguments. We then present evidence that premise relations are important for guiding inference generation, and that taxonomic relations are less privileged than we might expect. In the following section, we consider property effects on inference generation. We first review research showing systematic effects of property being projected on argument evaluation, and then present evidence that property has a number of important effects on inference generation, including overall biasing toward a particular basis of inference, and changing the salience of relations among premise categories. Finally, in the last section, we summarize our findings and draw conclusions.



2. EFFECTS OF PREMISE RELATIONS ON INFERENCE GENERATION

2.1. The Role of Premise Relations in Argument Evaluation

Previous research gives us some reason to expect relations among premise categories to be an important influence on the kinds of inferences people make. For example, [Medin et al. \(2003\)](#) have shown that salient relations among premise categories can lead to violations of normative logic or similarity-based predictions. Consider the diversity principle which suggests that dissimilar premise categories should provide stronger evidence for a generalization to an inclusive category than similar premise categories.

Medin et al. showed that arguments with taxonomically diverse yet causally linked premise categories (e.g., *Robins have property X and worms have property X therefore goldfish have property X*) were rated as weaker than arguments with less diverse yet unrelated premise categories (e.g., *Robins have property X and iguanas have property X therefore goldfish have property X*). Likewise, arguments with taxonomically diverse premise categories that also share a salient and potentially relevant feature (e.g., *Polar bears have property X and penguins have property X therefore all animals have property X*) were rated as weaker than arguments with less diverse yet unrelated premise categories (e.g., *Polar bears have property X and antelopes have property X therefore all animals have property X*). In both cases, comparison of premise categories yielded a salient linking relation, be it causal (robins eat worms) or featural (polar bears and penguins are both found in cold climates). This relation provided an explanation for the presence of the property in both species, and therefore weakened the general argument relative to a less diverse yet unrelated pair of premise categories (but see [Heit & Feeney, 2005](#)). Likewise, salient relations among premise categories also resulted in nonmonotonicity, whereby arguments with fewer premise categories that are proper members of the same superordinate as a conclusion category (e.g., *Brown bears have property X therefore buffalo have property X*) were rated as stronger than arguments with more such premise categories (e.g., *Brown bears have property X, polar bears have property X, and grizzly bears have property X therefore buffalo have property X*) if the additional premises reinforced a relation shared by premise categories but not the conclusion category.

In sum, [Medin et al. \(2003\)](#) demonstrate that salient shared features or causal relations among premise categories can have a marked effect on inductive inferences. By rendering specific categories and/or relations such as *bears*, *arctic animals*, or *predator-prey* highly salient, these manipulations serve to overcome more general default approaches to evaluation of inductive arguments. These findings suggest that participants are comparing premise categories and extracting salient relations between them in order to inform their evaluations of categorical inductive arguments.

Consistent with this view, [McDonald, Samuels, and Rispoli \(1996\)](#) proposed that subjects view the premises of a categorical argument as evidence and the conclusion as a hypothesis. As such, the argument should be perceived as weak to the degree that competing hypotheses are brought to mind by the evidence (that is, the premises). To test this, McDonald et al. asked one group of participants to imagine they were scientists who had just discovered “substance X” in certain sets of organisms. Using this “evidence,” they were asked to construct plausible conclusions as to “general categories of organisms that might reasonably be expected to contain substance X” (p. 204). Another group rated the strength of arguments taking each set of organisms as premises and a more general class as a conclusion. Results showed a strong relation between responses to the

two tasks; the more competing hypotheses were generated in the first task, the weaker the argument was rated in the second task. These results imply that people may spontaneously generate alternatives to a given conclusion based on salient relations among premise species, and utilize these alternatives to evaluate the strength of an argument.

Indeed, Feeney, Coley, and Crisp (2010) showed that while reading premises of an inductive argument, participants actively construct hypotheses about which relations among premise categories might be relevant; when a premise is inconsistent with a current hypothesis, that premise takes longer to read, and has a larger effect on ratings of argument strength, than when the same premise is consistent with a current hypothesis. For example, participants were faster to read the third premise of the argument *Magpies have property X, panda bears have property X, zebras have property X, therefore. . .* than to read the third premise of the argument *Brown bears have property X, panda bears have property X, zebras have property X, therefore. . .* despite the fact that the third premise and the preceding second premise are identical in both arguments. This suggests that people form a hypothesis about likely relevant relations based on comparing the categories in the first two premises. In the first argument, this is likely something like, *black and white animals*, so that *zebras* is consistent with the hypothesis and therefore processed quickly. In the latter argument, the likely hypothesis is probably *bears*, and because it is inconsistent with this hypothesis, *zebras* takes longer to process. Together, this evidence suggests that people may compare premise categories to glean information about likely conclusions, and that salient relations among premise categories may serve to bias or constrain the kinds of inferences generated from those premises.

2.2. Relative Salience of Conceptual Relations

We have argued that inductive projections from one category to related categories can be based on a number of different kinds of intercategory relations, including common membership in a taxonomic category and intrinsic similarity, extrinsic similarity with respect to some shared contextual or environmental relation, or causal relations between categories. We have also argued that salient relations among premise categories are among the factors that influence what knowledge is used to guide an inductive inference. However, there is reason to believe that knowledge of these relations may differ systematically in baseline salience; in particular, taxonomic knowledge may be especially salient and accessible, and correspondingly privileged for guiding inductive inference, at least for North American university undergraduates reasoning in domains in which they lack expertise (Coley, Medin, Proffitt, Lynch, & Atran, 1999; Coley, Shafto, Stepanova, & Barraff, 2005; Shafto, Coley, & Vitkin, 2007). For example, verification of taxonomic category membership of foods (e.g., *apple is a fruit*) is faster

than verification of script category membership (*apple* is a *snack*), and priming facilitates script category verification but not taxonomic verification (Ross & Murphy, 1999; Vitkin, Coley, & Feigin, 2005). Moreover, priming taxonomic category membership inhibits script categorization, whereas priming script categories has no effect on taxonomic categorization (Vitkin et al., 2005). Likewise, time pressure inhibited inductive projections among animals that shared habitat, but had no effect on projections among taxonomically related animals (Shafto, Coley, & Baldwin, 2007). Inferences based on shared category membership are present from early in development (Gelman, 2003; Gelman & Coley, 1990). Novice adults widely utilize taxonomic relations to guide inferences (López et al., 1997; Osherson et al., 1990; Shafto & Coley, 2003) as do experts when the property being projected provides little information about what other relations might be relevant (Shafto & Coley, 2003).

Together, these findings suggest that taxonomic knowledge may be more accessible than other kinds of knowledge, and that taxonomic similarity may be an especially important foundation for inductive inferences. However, as argued above, people are willing and able to recognize that inferences based on other kinds of relations, including extrinsic similarity and causal relatedness, can also be inductively strong. This raises the question about the frequency with which people spontaneously utilize these different relations to guide inductive inferences. In this study, we examine how participants with no special expertise in a domain recruit knowledge of taxonomic, causal, and contextual relations to generate inductive inferences.

2.3. Study One: Investigating Effects of Premise Relations on Inference Generation

To investigate the degree to which relations among premise categories guide open-ended inductive inferences, we constructed pairs of local animals that varied in their taxonomic and ecological relatedness. Participants were told that each pair shared a novel internal substance and were asked what other animals they thought might have the substance and why. We reasoned that a novel “substance” was sufficiently ambiguous to avoid overly constraining the nature of the inference; if construed as an innate physiological substance (e.g., analogous to serotonin), it could plausibly be projected along taxonomic lines, whereas if construed as an environmentally transmitted substance (e.g., analogous to DDT), it might plausibly be projected on the basis of extrinsic similarity or causal interaction. Finally, we assessed participants’ beliefs about the relatedness of the premise pairs, which allowed us to directly examine how such beliefs predicted patterns of inference.

In general, we hypothesized that salient relations among premise categories would influence the basis of inductive inference; we expect participants to compare premise categories, extract salient relations, and generate inferences consistent with those relations. If so, then salient taxonomic relations among premises should yield taxonomic inferences, whereas premises with salient spatiotemporal or causal relations should yield corresponding inferences.

We were also interested in examining the degree to which taxonomic similarity may be a privileged basis for guiding inference generation. A taxonomic bias might manifest itself in several ways. We might simply observe that inferences based on taxonomic similarity are generated more frequently than those based on extrinsic similarity or causal relatedness. Alternatively, taxonomic relations among premise categories might exert a stronger influence on inferences than ecological relations. In the following, we explore these possibilities.

2.3.1. Method

2.3.1.1. Research Design and Procedure Thirty-one Northeastern University undergraduates were recruited from introductory psychology classes and participated for course credit. We chose 12 pairs of animal species native to Massachusetts to independently manipulate the presence of salient taxonomic and ecological relations (see Table 1). Pairs were either taxonomically near (drawn from the same or a closely related superordinate biological class) or taxonomically far (drawn from different superordinate biological classes). Orthogonally, pairs were either ecologically related—via predatory/prey relation, shared habitat, or ecological niche—or ecologically unrelated. Results of posttests (see below) confirmed that participants viewed the relatedness of the premise pairs in the manner in which we intended.

Table 1 Stimulus Pairs, Study One.

Ecological relatedness	Taxonomic distance	
	Taxonomically near	Taxonomically far
Related	Coyote/Bobcat	Beaver/Spotted Turtle
	Water Snake/Green Frog	Red-tailed Hawk/ Field Mouse
	Heron/Duck	Garter Snake/Owl
Unrelated	River Otter/Deer	Chipmunk/Bullfrog
	Hummingbird/Canada Goose	Chickadee/ Salamander
	Box Turtle/Gray Tree Frog	Muskrat/Woodpecker

Participants were tested individually or in small groups in the laboratory. They were presented with a packet containing 12 pairs of animal names; each pair was presented on a separate page. Participants reasoned about a substance found in the bloodstream. Pairs were presented in random order. Instructions read “On each page of this packet you will find a pair of local animals which have been discovered to have a specific, naturally occurring substance in their bloodstreams. All you know about the substance is that these two kinds of animals have it. You will be asked to list other animals or kinds of animals you think might also have that substance, as well as reasons for your answers.”

For each pair, participants wrote down other species they expected would share the property and an explanation for why they projected the property from the premise pair to these animals. Following the inference generation task, participants completed a belief posttest in which they were asked directly about how they thought each premise pair was related. Specifically, for each item, participants were asked the following: Do these animals belong to the same biological category? Do these animals live in the same habitat? Does one of these animals eat the other? Questions were presented in this fixed order for all participants, who simply checked “yes” or “no.”

2.3.1.2. Coding Each response consisted of a list of participant-generated conclusion categories and an explanation for why a property true of the premise pair was likely to be shared by those categories. In order to systematically quantify these responses, we developed a coding system to characterize the basis of each inference. Responses were coded based on both (1) relations between experimenter-generated premise categories and participant-generated conclusion categories, and (2) the explanation generated by the participant. Coding categories were not mutually exclusive; a given response could receive multiple codes.² Four or five trained coders blind to the hypotheses coded each response independently. Consensus (defined as agreement between N-1 coders) was reached on over 90% of codes. Disagreements were resolved by discussion. The coding scheme is summarized in [Table 2](#).

Inferences based on category membership and similar habitat were most frequent, followed by inferences based on food chain interaction, behavioral similarity, and perceptual similarity. The remaining types of inferences were relatively infrequent; means are presented in [Table 3](#). To examine the effects of relations between premise categories on the three broad classes of inferences discussed above, we collapsed the initial coding categories into those that reflect reasoning based on common category membership, appearance, or other shared intrinsic features (henceforth, we will refer to these as taxonomic inferences), those that reflect reasoning based on extrinsic similarity or shared contextual features like similar diet or habitat (henceforth,

² Multiple-code assignments did not represent ambiguous responses, but rather responses in which participants invoked different kinds of reasons to support a particular inference.

Table 2 Coding Scheme for Characterizing Basis of Inference.

Basis of inference	Definition and example
<i>Taxonomic inferences</i>	
Category membership	P and C belong to the same class or category. [Heron/Duck] <i>Geese and cranes because the disease seems to be related to birds.</i>
Perceptual similarity	P and C are similar with respect to some aspect of superficial surface appearance. [Box turtle/Gray tree frog] <i>Lizards because they are green just like turtles and frogs.</i>
Behavioral similarity	P and C are similar with respect to some aspect of behavior. [Chipmunk/Bullfrog] <i>Squirrel and rabbit because they are fast-moving animals.</i>
Physiological similarity	P and C are similar with respect to specific organs or systems. [Box turtle/Gray tree frog] <i>Rodents and turtles because they have similar genetic makeup.</i>
General similarity	P and C are alike or have similarities without further specifying the nature of the similarity. [Muskrat/Woodpecker] <i>Rodents and birds because they are similar to muskrat and woodpecker.</i>
<i>Extrinsic inferences</i>	
Similar diet	P and C are similar with respect to diet or eating the same kind of thing. [Chickadee/Salamander] <i>Other plant-eating or insect-eating animals because both examples eat plant matter and insects.</i>
Similar habitat	P and C share similar or the same habitat without specification that the property is transmitted via habitat. [Owl/Garter snake] <i>Bears and tigers because they can all be found in the woods.</i>
<i>Causal inferences</i>	
Predatory interaction	P and C interact with respect to predation, that is, one or both Ps eats or is eaten by C. [Field mouse/Red-tailed hawk] <i>An owl, red-tailed hawks could get substance B from eating field mice. Owls eat small animals like field mice.</i>
Habitat interaction	P and C share or pass a property by coming into contact through the same habitat. [Beaver/Spotted turtle] <i>Other animals of their local area; two animals sharing the same bodies of water probably get the substance from it, either by ingestion or absorption, since they spend a lot of time in the water.</i>
Behavioral interaction	P and C interact via some aspect of behavior. [Water snake/Frog] <i>Beaver - beaver has the tendency to fight snakes and if the disease C is contagious, the beaver might end up contracting it.</i>
General interaction	P and C interact without further specification of the nature of the interaction. [River otter/Deer] <i>Ticks, flies, river fish; anything that would come in contact with both the deer and the river otter.</i>

Note: P = (given) premise categories; C = (participant generated) conclusion categories.

Table 3 Mean Relative Frequencies for each Inference Type, Studies One and Two.

Type of inference	Study One		Study Two	
	Substance	Substance	Disease	Gene
Taxonomic	0.55 (0.240)	0.55 (0.185)	0.45 (0.225)	0.74 (0.167)
Category membership	0.34 (0.200)	0.34 (0.177)	0.36 (0.178)	0.26 (0.223)
Perceptual similarity	0.10 (0.129)	0.12 (0.131)	0.06 (0.085)	0.26 (0.195)
Behavioral similarity	0.11 (0.148)	0.09 (0.104)	0.06 (0.078)	0.25 (0.125)
Physiological similarity	0.05 (0.084)	0.02 (0.043)	0.01 (0.025)	0.03 (0.050)
General similarity	0.06 (0.092)	0.06 (0.087)	0.06 (0.099)	0.10 (0.086)
Extrinsic	0.38 (0.169)	0.43 (0.174)	0.39 (0.206)	0.27 (0.146)
Similar habitat	0.31 (0.161)	0.35 (0.190)	0.29 (0.165)	0.24 (0.151)
Similar diet	0.09 (0.122)	0.11 (0.095)	0.12 (0.122)	0.04 (0.061)
Causal	0.18 (0.187)	0.22 (0.237)	0.44 (0.257)	0.03 (0.044)
Predatory interaction	0.12 (0.136)	0.17 (0.170)	0.26 (0.188)	0.03 (0.044)
Habitat interaction	0.06 (0.100)	0.04 (0.068)	0.13 (0.159)	0.00 (0.000)
Behavioral interaction	0.01 (0.026)	0.01 (0.023)	0.04 (0.053)	0.00 (0.011)
General interaction	0.00 (0.000)	0.02 (0.047)	0.03 (0.050)	0.00 (0.000)

Note: Standard deviations appear in parentheses.

extrinsic inferences), and those that reflect reasoning based on some causal mechanism, including co-occurrence in space and time or direct contact through behavior or predation (henceforth, causal inferences). The precise makeup of these broad categories can be found in Table 2. If a given response was coded as any of the component inference types, it was scored as a taxonomic, extrinsic, or causal inference. Again, these categories were not mutually exclusive. For example, a response based on as category membership, behavioral similarity and habitat interaction would be counted both as a taxonomic inference and a causal inference; a response coded as behavior interaction and predatory interaction would be counted as a causal inference. These three general categories accounted for over 92% of all responses.

2.3.2. Results and Discussion

2.3.2.1. Relative Frequency of Inferences Results demonstrate that undergraduates used a wide range of knowledge to generate inferences about how a novel substance might be distributed among animal species. Overall, 56% of inferences were taxonomic (e.g., an inference from the heron/duck pair to “other birds because substance F might occur naturally in birds”), and almost 38% were extrinsic (e.g., given that herons and ducks have substance F, “a kingfisher might have substance F in its bloodstream because it feeds off fish like the heron and lives by streams like the duck”). Causal inferences (e.g., a projection from heron/duck to “other birds or fish in the area because they may have gotten the substance from a common pond or body of water”) were least common, but still accounted for 18% of responses. One-way ANOVA confirmed that the differences between all three means were statistically significant ($F(2,60) = 21.66, p < 0.0001, h_p^2 = 0.42$).

These results demonstrate the spontaneous use of knowledge about taxonomic relatedness, spatiotemporal contiguity, and causal transmission to guide inductive reasoning. It is important to stress that participants in this experiment were not simply reporting how the premise species were related, nor were they evaluating the plausibility of a fully formed argument. Rather, they were forming hypotheses about the distribution of a novel substance on the basis of information extracted via comparison of the premise species, and then using that hypothesis to generate plausible inferences. Sometimes these inferences were quite specific (e.g., an inference from garter snake/owl to “insects (ants, crickets) and spiders, mice because garter snakes eat insects, I think mice eat garter snakes and I would think owls eat at least mice if not both so maybe substance C travels the food chain”), and sometimes they were very vague (e.g., from garter snake/owl to “Any insects or animals eaten by either—the substance will pass from their blood to the animals”), but the critical point is that a broad range of knowledge about contextual similarity and interaction among premise categories—as well as taxonomic similarity—guided inference generation. Moreover, our open-ended methodology was sufficiently sensitive for us to detect such inferences even in the absence of

specific factual knowledge (about, e.g., what owls and garter snakes eat). More generally, these results revealed that category membership and intrinsic similarity—while clearly important—do not nearly exhaust the spectrum of knowledge utilized to guide inductive inferences.

2.3.2.2. Effects of Premise Relations Our central question in Study One was the degree to which relations among premise categories guide and constrain patterns of inferences. To address this we manipulated relations among premise categories to look for corresponding differences in inference patterns. We originally conceived of this manipulation as a 2 (taxonomic distance) \times 2 (ecological relatedness) design, but results of posttests suggested that although participants viewed the overall relatedness of the premise pairs in the manner in which we intended, individual variability in salience of relations both within and between our planned item classes was larger than we anticipated. The world, apparently, is not a factorial design. Also, the distribution of predation versus shared habitat relations was uneven in our “ecologically related” cells, and this turned out to be a critical distinction. As such, we deemed it more appropriate to construe the salience of taxonomic, habitat, and predation relations between each pair of premise categories as a continuous variable (ranging from weak to strong) rather than a categorical variable (present/absent) as originally conceived.

We used multiple regression analyses to look at the data in this way. For each item, we calculated the relative frequency of positive responses to each belief question (“do these two animals belong to the same biological category,” “do these two animals live in the same habitat,” and “does one of these animal eat the other”) averaged across all participants. We took these scores as indices of the salience of taxonomic relatedness, shared habitat, and predatory relations between premise species in each item. We also assigned each item scores corresponding to the mean relative frequency of taxonomic, extrinsic, and causal inferences for that item, again averaged across all participants. Using arcsine transforms of these scores, we conducted three item-wise multiple regression analyses using beliefs about premise relations to predict the relative frequencies of each type of inference; standardized regression coefficients for these multiple regressions are presented in [Figure 1](#).

This analysis demonstrated that the primary predictor of taxonomic inferences was the salience of the taxonomic relatedness of a premise pair; the frequency of taxonomic inferences increased with taxonomic salience, but was unrelated to the salience of shared habitat or predatory relations ($R^2 = 0.65$, $p = 0.031$). In contrast, extrinsic inferences were positively related to the salience of shared habitat, but unrelated to the salience of taxonomic relatedness. Extrinsic inferences were also negatively related to the salience of predatory relations between premise pairs ($R^2 = 0.64$, $p = 0.033$). Finally, frequency of causal inferences was positively related

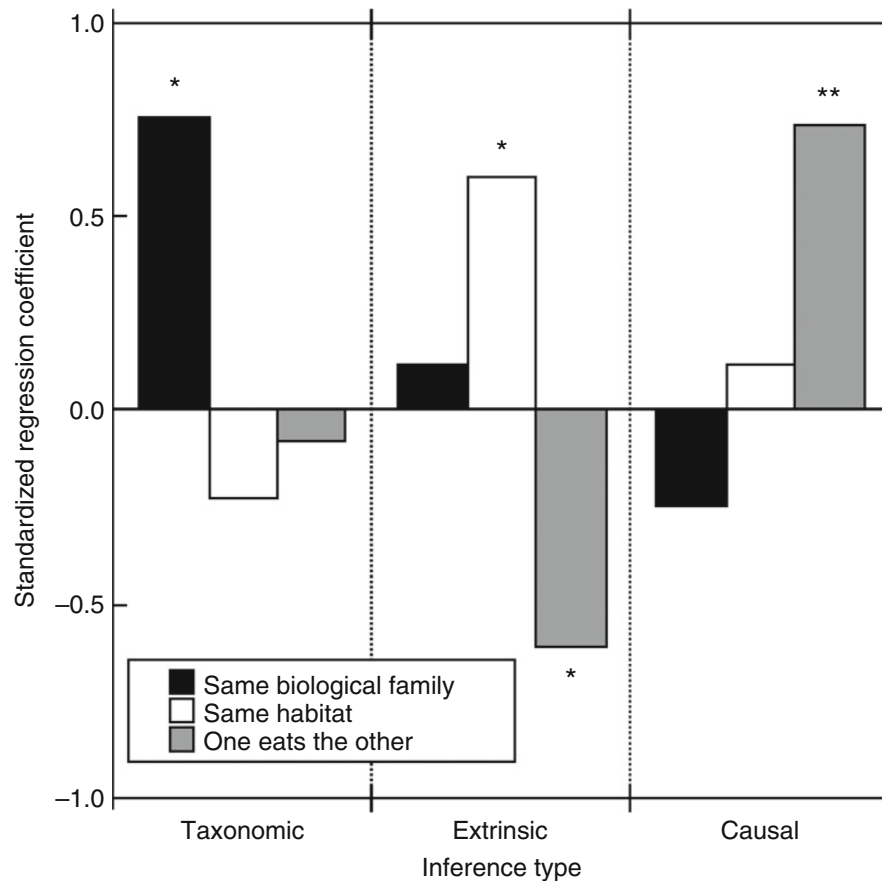


Figure 1 Relations between salience of premise relations and frequency of taxonomic, extrinsic, and causal inferences, Study One. (Note: * $p < 0.05$, ** $p < 0.005$.)

to salience of predatory relations, but unrelated to salience of taxonomic relations or shared habitat ($R^2 = 0.80$, $p = 0.004$).

Thus, as predicted, we observed strong effects of relations among premise categories on the nature of inferences generated from those categories. Taxonomic inferences varied with the salience of taxonomic relations among premise species, whereas extrinsic and causal inferences varied with the salience of ecological relations among premise species. Indeed, the correspondence between premise relations and nontaxonomic inferences was quite specific; salient shared habitat led to increased extrinsic reasoning, whereas salient predatory relations led to increased causal reasoning and decreased extrinsic reasoning. This last finding hints that participants preferred to draw inferences based on a salient causal mechanism when possible (e.g., projecting from owl/garter snake to “eagle, hawk, water snake, and garden snake—I think that these animals could have this substance because they are predators and if the animal is eaten with the substance, it could get into the other animal’s bloodstream”), and fell back

on extrinsic similarity when no such mechanism was readily apparent (e.g., projecting from chickadee/salamander to “animals that share their neighborhood: chickadees and salamanders don’t eat exactly the same items; if their shared dirt, air and water have the same substance, several animals could be affected”).

2.3.2.3. *Privileged Taxonomic Inferences?* Somewhat to our surprise, our results did not provide support for the privileged status of taxonomic relations in inference generation. If we compare the frequency of taxonomic inferences to that of “nontaxonomic” inferences (i.e., inferences based on extrinsic similarities or causal interaction), the two do not differ (0.56 vs. 0.53; $t(30) = 0.35$, $p = 0.726$). Moreover, there was no evidence that taxonomic relations among premise categories exerted undue influence on which inferences were generated. Taxonomic salience predicted taxonomic inferences but did not suppress extrinsic or causal inferences; on the contrary, the most influential premise relations were predation relations, which simultaneously promoted causal inferences and inhibited extrinsic inferences.

2.4. Summary: Effects of Premise Relations on Inference Generation

We want to emphasize four conclusions from Study One. First, our participants found the task perfectly natural, and spontaneously generated inferences based on contextual similarity and causal relations as well as inferences based on intrinsic similarity and category membership. For example, upon learning that a water snake and a frog both have a certain substance in their bloodstream, participants used what they know about relations between water snakes and frogs to guide further inferences about the substance.

Second, there was a tight linkage between the perceived relations between premise categories and the inferences those categories engendered. Taxonomic relations supported taxonomic inferences; if the belief that snakes and frogs are (somewhat) related taxonomically was most salient, a participant was likely to generate an inference based on taxonomic similarity (e.g., “salamander, bullfrog, turtle: they are all reptiles”). In contrast, shared habitat promoted extrinsic inferences, and predation relations supported causal inferences. If a participant noticed that water snakes and frogs share a habitat, or that water snakes like to eat frogs, they were more likely to generate inferences based on extrinsic similarity or causal interaction (e.g., “bullfrog, fish (freshwater), plants: they live in the same environment and could’ve contracted from the same thing or the snake could’ve just eaten the infected frog,” or “other freshwater life, other snakes, animals that eat small water animals: again it seem to begun on the frog and it could move up the

food chain’’). Thus, salient relations between premise categories served as a foundation for generating inferences.

Third, we found that when premises were causally related via predation, not only were causal inferences promoted, but extrinsic inferences were also suppressed, suggesting that inferences with some casual component were preferred over inferences based on similarity. And finally, we found little evidence for the primacy of taxonomic relations in guiding inductive inferences. Instead, it appears that when people generate inferences in an open-ended task, they consider a wide range of relations as bases for constructing inductive hypotheses.



3. EFFECTS OF PROPERTY ON INFERENCE GENERATION

In the previous section, we presented evidence that salient relations among premise categories strongly influence the kinds of open-ended inductive inferences that university students spontaneously generate. The open-ended inference generation task proved to be sensitive to sophisticated reasoning patterns our participants employed even in the absence of specific biological knowledge. In this section, we focus on another source of information that can provide relevant constraints on generation of inductive inferences: the nature of the property being projected.

3.1. Property Effects in Argument Evaluation

The nature of the property being generalized can have a profound impact on the strength of inductive arguments. For example, if you learn that frogs have gene X, you might infer that toads would also have the gene on the basis of close biological affinity among them and/or common membership in the same taxonomic class. In contrast, if you learn that frogs have disease X, you might reasonably expect fish to also have the disease because both live in aquatic environments, or raccoons might even have the disease from eating sick frogs. More generally, the nature of the property influences the likelihood that you would generate one type of inference over another. Because what we know about genes includes the knowledge that members of the same biological family are genetically similar, this property may increase the salience of taxonomic knowledge about frogs, thereby increasing the likelihood of a taxonomic inference. Likewise, what we know about disease includes knowledge about contagion via contact or eating contaminated food. As such, this property may increase the salience of knowledge about frogs’ habitat and role in the food chain, thereby increasing the likelihood of an inference based on extrinsic similarity or interaction. This

systematic link between the kind of property being projected and the likelihood of different inferences has been termed inductive selectivity.

An early demonstration that property could affect perceived argument strength came from Heit and Rubinstein (1994) who asked participants to estimate the probability that a pair of animal species would share a property. The pairs were anatomically similar (e.g., *mouse* and *bat*) or behaviorally similar (e.g., *sparrow* and *bat*). Likewise, properties were either anatomical (e.g., “having a liver with two chambers that act as one”) or behavioral (e.g., “traveling in back and forth, or zig-zag, trajectory”). They found that probability judgments were determined by the match between the relation shared by the animals in a pair and the projected property. Arguments in which the premise and conclusion were anatomically similar were judged stronger for an anatomical property than for a behavioral property, whereas arguments with a behaviorally similar premise and conclusion were judged stronger for behavioral than anatomical properties. The authors concluded that property influenced the perceived similarity between premises and conclusions, which in turn determined the strength of the inference.

Ross and Murphy (1999) showed that the nature of the property influenced the use of taxonomic or script categories to guide inferences about food. Participants were presented with triads consisting of a target food (e.g., cereal) and two alternatives, one taxonomic (noodles, another member of the *bread & grains* category), and the other one script (milk, another *breakfast food*). They were taught that the target food (cereal) had a biochemical property (a novel enzyme) or a situational property (eaten at a particular ceremony in an unfamiliar culture), and asked to project that property to one of the alternatives. Participants preferred taxonomically related conclusions when projecting a biochemical property, but conclusions related via a common script when projecting situational properties (see also Nguyen & Murphy, 2003; Vitkin, Coley & Feigin, 2005).

Inductive selectivity is present early on. By kindergarten and perhaps earlier, children are able to selectively utilize taxonomic similarity among species to guide inferences about novel physiological properties, and ecological relations to guide inferences about disease (Coley, Vitkin, Seaton, & Yopchick, 2005; Coley et al., 2007; Vitkin Vasilyeva, & Coley, 2007). This suggests that from relatively early in development, children are sensitive to different relations among living things and their selective inductive potential. Likewise, Kalish and Gelman (1992) showed that when reasoning about artifacts, preschoolers projected unfamiliar dispositional properties (e.g., will get fractured if put in really cold water) on the basis of material kind (e.g., glass), but projected unfamiliar functional properties (e.g., used for partitioning) on the basis of object kind (e.g., scissors).

Shafto and Coley (2003) showed that when projecting a novel disease, commercial fishermen utilized knowledge of predator–prey relations among local marine creatures to guide casual inferences; in contrast, when reasoning

about a completely blank property (e.g., has property X), fishermen showed no sign of causal reasoning, and instead utilized taxonomic similarity to guide inferences. Along these same lines, Tenenbaum and colleagues (e.g., Griffiths & Tenenbaum, 2005; Shafto, Kemp, Bonawitz, Coley, & Tenenbaum, 2008; Tenenbaum, Griffiths, & Kemp, 2006) propose a family of models of statistical inference over structured knowledge representations. These models posit separable knowledge structures based on intuitive theories that are called on to guide inductive reasoning in different contexts. In support of this idea, Shafto et al. (2008) explicitly taught biologically naïve undergraduates about food web and taxonomic relations among species. They found that a causal model of the food web predicted inferences about novel diseases but not genes, whereas a tree-structure of taxonomic relations among species predicted inferences about novel genes but not diseases. This is consistent with the view that property serves to indicate which knowledge structure is relevant for evaluating the strength of a given argument.

In sum, there is ample evidence that the property being projected influences the way in which inductive arguments are evaluated. Typically, these effects are interpreted in terms of the property biasing the computation of premise–conclusion relatedness. However, we know little about how the property being projected might influence the way in which inductive inferences are generated.

3.2. Study Two: Investigating Effects of Property on Inference Generation

To examine this question, we utilized the same open-ended inductive inference task that we used in Study One. Our primary goal was to look at the effects of property on patterns of inference generation. To this end, we asked participants to reason about a novel gene, disease, or substance. *Gene* was chosen as a taxonomically biasing property; participants are likely to believe that genes are distributed along taxonomic lines (Shafto, Coley, & Baldwin, 2007). If property affects what relations between premises are noticed and/or constrains what inductive hypotheses are generated, we expect that thinking about the distribution of a novel gene should render taxonomic relations particularly salient and thereby facilitate inferences based on taxonomic similarity. Likewise, *disease* was chosen as an ecologically biasing property; participants are likely to believe that disease can be transmitted along ecological lines via contact and/or contamination (e.g., Shafto, Coley, & Baldwin, 2007; Shafto, Coley, & Vitkin, 2007; Shafto et al., 2008). Thus, reasoning about a disease should highlight shared habitat or predation relations, and thereby facilitate reasoning based on extrinsic similarity or interaction. We included *substance* in order to replicate results of Study One. Moreover, since we observed all three types of inferences of interest among subjects reasoning

about substance in Study One, it represented a relatively ambiguous and unbiased property.

Study One revealed a surprising sensitivity to the specific nature of the ecological relation among premise species. Therefore, in Study Two we were more careful about the specific nature of the ecological relation between related pairs; we manipulated whether the pair was related via shared habitat only (e.g., koala–kangaroo), or via predation as well as habitat (e.g., lion–zebra). We also tripled the number of items used. Based on results of Study One, we expected taxonomic relations to promote taxonomic inferences and predation relations to promote causal inferences. We also expected extrinsic inferences to increase with the salience of shared habitat and decrease with the salience of predation.

We were particularly interested in potential interactions between property and premise relations. One possibility is that premise relations and property may have relatively independent effects on inferences: each may serve to render particular classes of inferences more or less likely. Participants' naïve theories about how different properties are distributed may increase the overall likelihood of property-congruent types of inferences. For example, reasoning about disease may render causal inferences likely regardless of whether premises themselves are related via predation. Another possibility is that property may influence the way that premise relations influence inferences. For example, predation relations between premise categories may increase the likelihood of causal inferences about *disease*, but not about *gene*. These two possibilities are not mutually exclusive. In Study Two, we examine both of them.

3.2.1. Method

3.2.1.1. Research Design and Procedure Seventy-two Northeastern University undergraduates were recruited from introductory psychology classes and participated for course credit. For this study we developed 36 pairs of animal species. As in the previous study, pairs were either taxonomically near or far. However, we were more precise in manipulating the ecological relations among species. Specifically, pairs were either ecologically related via predation (one species typically preyed on the other), ecologically related via shared habitat (both species are typically found in the same habitat but do not prey on each other), or ecologically unrelated (species are typically found in nonoverlapping habitats³). As in Study One, ecological relatedness was orthogonal to taxonomic distance. A complete list of stimuli is presented in [Table 4](#).

³ We also manipulated whether species were exotic or local—18 pairs (6 from Experiment 1 and 12 new pairs) were local species (native to Massachusetts) and another 18 pairs (all new) represented exotic species (not typically occurring naturally in New England). However, we collapsed across that variable for purposes of this chapter.

Table 4 Stimulus Pairs, Study Two.

Ecological relation	Taxonomic distance	
	Taxonomically near	Taxonomically far
Predation	Lion/Zebra*	Herring/Penguin*
	Harp seal/Polar bear*	Anteater/Leaf-cutter ant*
	Hammerhead shark/Sardine*	Lemming/Snowy owl*
	Green frog/Water snake	Salmon/Black bear
	Fox/Rabbit	Brown bat/Mosquito
	Fly/Spider	Hawk/Field mouse
Shared habitat	Tarantula/Scorpion*	Dolphin/Seahorse*
	Koala bear/Kangaroo*	Elephant/Crocodile*
	Toucan/Parrot*	Macaw/Jaguar*
	Porcupine/Moose	Spotted turtle/Beaver
	Heron/Duck	Lobster/Tuna
	Bee/Butterfly	Owl/Deer
Unrelated	Gorilla/Caribou*	Clownfish/Tiger*
	Emu/Flamingo*	Giraffe/Puffin*
	Emerald tree boa/Komodo dragon*	Gecko/Peacock*
	Hummingbird/Canada goose	Bullfrog/Chipmunk
	Newt/Box turtle	Raccoon/Pelican
	Humpback whale/Squirrel	Chickadee/Salamander

Note: * denotes exotic items.

In addition to manipulating relations among premise species, we also manipulated the property that participants were asked to reason about. As discussed above, participants reasoned about novel substances, genes, or diseases. In sum, each participant reasoned about a single property and generated inferences from 18 local or exotic pairs of animals. Taxonomic distance and ecological relation were manipulated within-subjects.

The instructions and procedure were the same as in Study One, except that in the disease condition, the animal pairs were “discovered to have a certain disease” rather than “a specific, naturally occurring substance in their bloodstreams” and in the gene condition they were “discovered to have a certain gene.” Following the inference task, participants completed a belief posttest as in Study One.

3.2.1.2. Coding The coding procedure was identical to that used in Study One. Consensus (defined as agreement between N-1 coders) was reached on over 93% of codes. Disagreements were resolved by discussion. As in

Study One, the most common inferences were those based on category membership ($M = 31\%$), similar habitat ($M = 29\%$), and food chain interaction ($M = 15\%$). Means for all inference types in all three property conditions are given in Table 3. We again collapsed the initial coding categories into those that reflect taxonomic, extrinsic, and causal inferences. Together, these coding categories accounted for 91% of all responses.

3.2.2. Results and Discussion

3.2.2.1. Relative Frequency of Inferences Again, undergraduates used a wide range of knowledge to generate inferences about how novel properties might be distributed among animal species. Across all conditions, 58% of inferences were taxonomic (e.g., an inference from hummingbird/Canada goose to “Quail, all geese types, because it seems likely that since hummingbirds and Canada geese have it all bird relatives might have substance P in their bloodstream”). Likewise, 35% of inferences were extrinsic (e.g., one participant projected a substance from green frog/water snake to “eels, tadpoles, alligators, crocodiles: they all for the most part dwell in water”). Finally, 23% of inferences were causal (e.g., projecting “disease E” from lobster/tuna to “clams, seaweed, or algae, because lobster could eat the clams and that could suggest the whole water species has it or it could’ve been something the tuna ate, i.e., seaweed/algae that infected it and then once the lobster ate the infected tuna, it too became ‘E positive’”). These differences were significant via one-way ANOVA ($F(2,138) = 50.12$, $p < 0.0001$, $h_p^2 = 0.42$). These results again demonstrate the spontaneous use of knowledge about taxonomic relatedness, spatiotemporal contiguity, and causal transmission to guide inductive reasoning. Moreover, the absolute means were remarkably similar across the two studies, providing some confidence that the distribution of inferences observed in Study One were not due solely to idiosyncrasies of premise pairs.

3.2.2.2. Property Effects A central question in Study Two was the degree to which properties guide and constrain the generation of inductive inferences. To address this question, we first collapsed across all items in order to look at the overall impact of property on the frequencies of our three types of inferences. In general, we expected substance to show the same overall pattern as in Study One. In contrast, we thought disease might bias participants toward extrinsic and/or causal inferences and away from taxonomic inferences, and conversely, gene might bias participants away from extrinsic or causal inferences and toward taxonomic inferences. To test these hypotheses, we conducted a 3 (Property) \times 3 (Inference Type) ANOVA. As expected, the distribution of inferences differed markedly depending on the kind of property being projected (Inference Type \times Property interaction ($F(4,138) = 17.32$, $p < 0.0001$, $h_p^2 = 0.33$)).

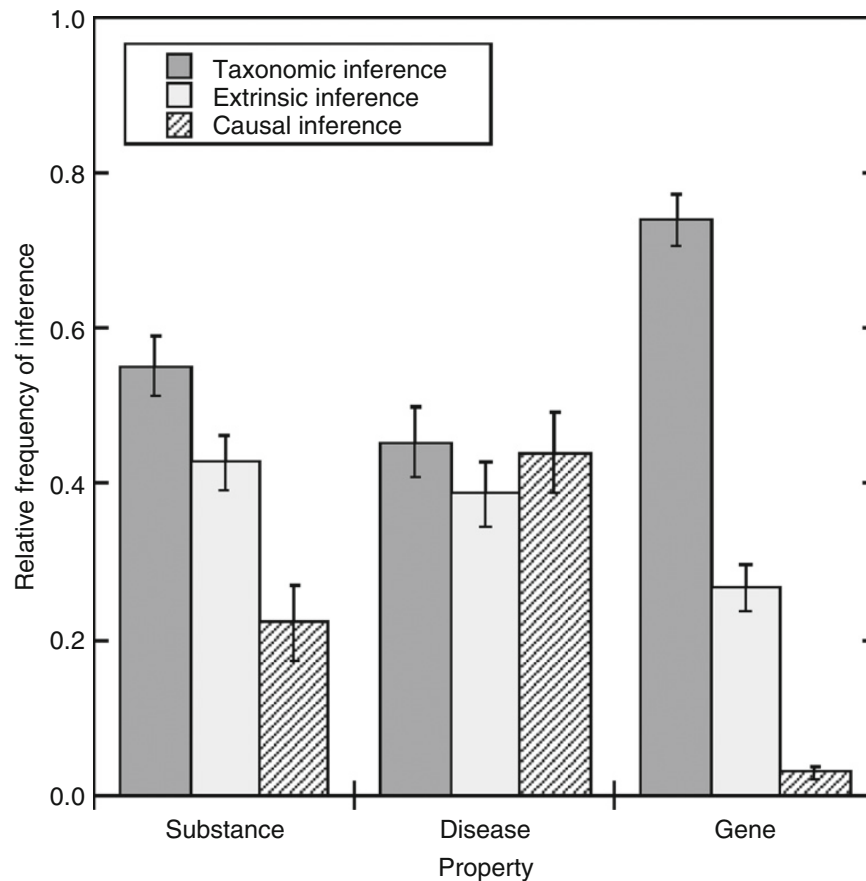


Figure 2 Mean relative frequency of taxonomic, extrinsic, and causal inferences in each property condition, Study Two. Error bars represent one standard error of the mean.

We analyzed this interaction in two ways. First, we looked at the relative distribution of inferences in each property condition, as depicted in [Figure 2](#). Second, to examine the degree to which reasoning about disease and gene changed the frequency of different inferences relative to the neutral substance condition, we compared the frequency of each type of inference for people reasoning about disease and gene to the frequency of the inference among people reasoning about substance. Difference scores are presented in [Figure 3](#).

3.2.2.2.1. Substance The distribution of inferences about substance replicated Study One; taxonomic inferences were marginally more frequent than extrinsic inferences, and both were more frequent than causal inferences, which were nevertheless fairly prevalent ($F(2,46) = 13.08$, $p < 0.0001$, $h_p^2 = 0.36$). However, taxonomic inferences were not more

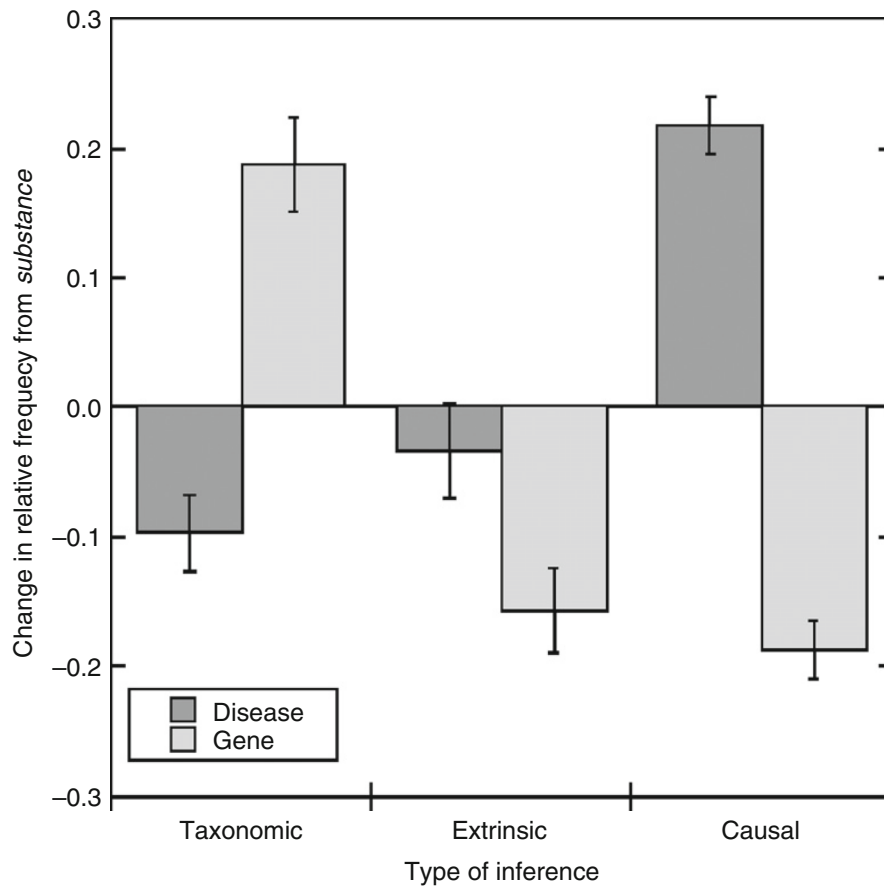


Figure 3 Mean change in relative frequency of inferences from *substance* condition, Study Two. Error bars represent one standard error of the mean.

frequent than nontaxonomic (extrinsic and causal inferences taken together) (0.58 vs. 0.55, $t(23) = 0.40$, $p = 0.691$), again replicating Study One.

3.2.2.2.2. Disease As expected, reasoning about disease led to a marked 22% increase in causal inferences relative to the substance condition ($t(46) = 3.06$, $p = 0.004$, Cohen's $d = 0.89$) and a 10% decrease in taxonomic inferences ($t(46) = 1.65$, $p < 0.11$, Cohen's $d = 0.48$; see Figure 3).⁴ The small decrease in extrinsic inferences did not approach significance. As a result, when compared to each other, the relative frequency of the three inference types in the disease condition did not differ ($F(2,46) < 1$; see Figure 2). However, the relative frequency of nontaxonomic inferences considered together was significantly higher than

⁴ Although the t -test by subjects only approached statistical significance, the t -test by items was highly significant ($t(35) = 3.37$, $p < 0.002$) and a sign test revealed that for 28/36 items, taxonomic similarity-based inferences were equally or less common for *disease* than for *substance* ($p = 0.007$). Together, these results plus the moderate size of the effect give us confidence in the reliability of the difference.

frequency of taxonomic inferences (0.67 vs. 0.47, $t(23) = 2.41$, $p = 0.024$, Cohen's $d = 0.84$). This pattern of inferences suggests that to some degree, reasoning about disease selectively supported inferences based on causal relations at a cost to inferences based on taxonomic similarity. However, it is clear that all three types of inferences were deemed relevant for reasoning about disease. In support of this view, responses in the disease condition were more likely to receive multiple codes than in the other two conditions ($F(2,69) = 6.17$, $p = 0.003$, $h_p^2 = 0.15$). This suggests that individual inferences about disease were more complex than inferences about substance or gene. In sum, although inferences generated about disease were more causal and less taxonomic than inferences generated about substance, they were by no means exclusively causal or extrinsic, and thus represented an increase in multidimensional reasoning.

3.2.2.2.3. Gene In contrast, *gene* was a strongly taxonomically biasing property, as we expected it to be. Compared to the *substance* condition, taxonomic inferences increased dramatically for people reasoning about genes, whereas extrinsic and causal inferences both decreased markedly ($t(46) > 3.44$, $p \leq 0.001$, Cohen's $d \geq 1.00$; see [Figure 3](#)). Consequently, as seen in [Figure 2](#), taxonomic similarity inferences were much more frequent than extrinsic similarity inferences which in turn were much more frequent than causal inferences ($F(2,46) = 141.35$, $p < 0.0001$, $h_p^2 = 0.86$). Not surprisingly, taxonomic inferences were also more frequent than nontaxonomic inferences (0.75 vs. 0.25, $t(23) = 9.03$, $p < 0.0001$, Cohen's $d = 3.11$). Thus, in both an absolute sense and a relative sense, reasoning about genes greatly increased the likelihood of generating taxonomic inferences. Interestingly, a close look at [Table 3](#) reveals that the increase in taxonomic reasoning about genes was not due to an increase in category-based inferences, which actually decreased in frequency (albeit not reliably). Rather, the increase in taxonomic reasoning stemmed from an increase in inferences based on perceptual similarity ($t(46) = 2.88$, $p = 0.006$, Cohen's $d = 0.83$) and behavioral similarity ($t(46) > 4.61$, $p < 0.0001$, Cohen's $d = 1.33$) relative to *substance*. This suggests that rather than simply falling back on category membership, participants may have attempted to connect the hypothetical gene with specific perceptual or behavioral attributes of premise species, and then base projections on those specific attributes. For example, one participant projected a gene from humpback whale/squirrel to "opossum, mole, gray mouse, dolphin: they are all gray in color" and from raccoon/pelican to "squirrel, seagull, pigeon: these are animals that rummage through things." This raises the interesting possibility that when people think about genes, they give more weight to their potential to give rise to certain observable characteristics than to their general association with a taxonomic class. In other words, people in this task seemed to be projecting "gray color genes" or "rummaging genes" rather than "mammal genes" or "bird genes."

In sum, property had large effects on the relative frequency with which different inferences were generated. Inferences about substance mirrored those of Study One. Inferences about disease were more complex and multidimensional than for other properties; relative to *substance*, inferences about disease were more likely to be causal, and less likely to be taxonomic, although all three types of inferences were seen as equally appropriate. In contrast, inferences about genes were strongly biased toward taxonomic similarity. We next examine effects of premise relations on inference generation, and in particular, the degree to which premise relations and property interact in constraining inference generation.

3.2.2.3. Effects of Premise Relations One motivation for conducting Study Two was to be more careful in our manipulations of ecological relations among premise species. As such, we strove to choose pairs that were related via predation and shared habitat, pairs that were related via shared habitat only, and unrelated pairs. Results of posttests suggested that although participants viewed the relatedness of the premise pairs in the manner in which we intended, individual variability in salience of relations both within and between our planned item classes was again larger than we anticipated. Therefore, as in Study One, we decided to trust our participants' beliefs about premise relatedness rather than our *a priori* expectations, and to construe the salience of taxonomic, habitat, and predation relations between each premise pair as continuous variables (ranging from weak to strong, based on participants' ratings) rather than as categorical variables (present/absent) as originally conceived. We present multiple regression analyses comparable to those in Study One—using salience of shared habitat, predation, and taxonomic relations to predict item-wise frequency of each type of inference—rather than ANOVA.

First, we averaged across property conditions to get an overall picture of how premise relations predicted inferences. Based on the results of Study One, we expected different inferences to be sensitive to different premise relations; of interest was whether Study Two replicated the specific relations between premises and inferences we observed in Study One. Results of this analysis are presented in [Figure 4](#).

Two things are notable in [Figure 4](#). First, the way in which premise relations facilitated inferences was identical to what we observed in Study One. Second, unlike in Study One, the salience of shared habitat rendered taxonomic inferences less likely. Specifically, the frequency of taxonomic inferences increased with taxonomic salience, decreased with the salience of shared habitat, but was unrelated to the salience of predation relations ($R^2 = 0.68, p < 0.0001$). In contrast, extrinsic inferences were positively related to the salience of shared habitat, but unrelated to taxonomic salience ($R^2 = 0.31, p = 0.007$). The negative relation between the salience of predatory relations and the frequency of inferences based on

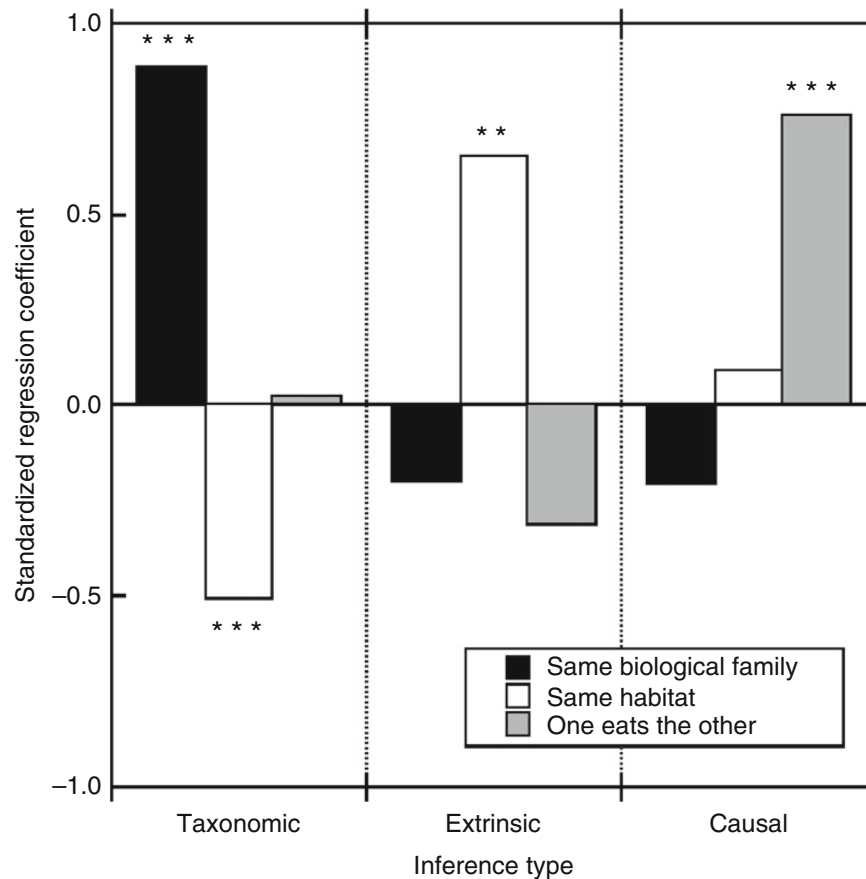


Figure 4 Relations between salience of premise relations and frequency of taxonomic, extrinsic, and causal inferences averaged across property conditions, Study Two. (Note: ** $p < 0.005$, *** $p < 0.0005$.)

extrinsic similarity observed in Study One was marginally significant overall; as we shall see, this particular relation varied by property. Finally, causal inferences were positively related to salience of predatory relations, but unrelated to salience of taxonomic relations or shared habitat ($R^2 = 0.69$, $p < 0.0001$) (as discussed below, this relation also varied somewhat with property). In sum, as in Study One, we observed a tight coupling between the salience of relations among premise categories and inferences drawn from those categories. Taxonomic relations promoted taxonomic inferences, shared habitat promoted extrinsic inferences, and predation relations promoted causal inferences.

It is notable that, unlike in Study One, the salience of shared habitat strongly inhibited taxonomic inferences, suggesting that in the presence of a salient alternative relation, the appeal of taxonomic inferences faded. One possible explanation is that taxonomic inferences serve as a default, and when people notice a salient habitat relation they may tend to believe that this is

what was being specifically “communicated” to them by this premise pair (according to the relevance theory of [Medin et al., 2003](#)) rendering them less likely to make a default taxonomic inference. Alternatively, the presence of salient habitat relations may have led participants to develop alternative contextual hypotheses that reduced the strength of taxonomic hypotheses, consistent with findings of [McDonald et al. \(1996\)](#). In either case, since we do not see consistent reciprocal effects of taxonomic relations on other types of inferences, we can speculate that people may have an internal “relevance ranking” of different relations, with contextual relations ranked fairly high.

3.2.2.4. Does Property Influence how Premise Relations Generate Inferences? So far, results show clear effects of property and of premise relations on generation of inductive inferences. However, we were particularly interested in whether these effects were independent of each other, or whether the way premise relations led participants to generate inferences varied by property. To examine this question, we performed separate multiple regressions on item-wise salience and inference scores for each property condition. Standardized regression coefficients are presented in [Figure 5](#); below we discuss results for each type of inference in turn.

3.2.2.4.1. Taxonomic inferences As seen in [Figure 5](#), taxonomic inferences increased with salience of taxonomic relations between premise categories, decreased with the salience of shared habitat, and were unaffected by the salience of predation relations in all three property conditions (Substance: $R^2 = 0.48$, $p < 0.0001$; Disease: $R^2 = 0.55$, $p < 0.0001$; Gene: $R^2 = 0.32$, $p = 0.005$). This suggests that the property being projected had little influence on the way in which premise relations licensed taxonomic inferences. Although the absolute level of taxonomic inferences varied from 74% for *gene* to 45% for *disease*, in all cases, salient taxonomic relations among premises facilitated the generation of taxonomic inferences, whereas salience of shared habitat inhibited them. Thus, property and premise relations exerted independent effects on taxonomic inferences.

3.2.2.4.2. Extrinsic inferences Extrinsic inferences were more weakly predicted by premise relations, and the nature of the relationship varied by property. As depicted in [Figure 5](#), for those reasoning about substance, frequency of extrinsic inferences increased with the salience of shared habitat, but was unrelated to the salience of taxonomic and predation relations ($R^2 = 0.31$, $p = 0.007$) whereas for gene, extrinsic inferences increased with the salience of shared habitat, and decreased with the salience of both taxonomic and predation relations ($R^2 = 0.35$, $p = 0.003$). This pattern suggests that—unlike taxonomic inferences—property changed the way premise relations promoted extrinsic inferences. While any detailed explanation of this pattern of results would be speculation, results clearly

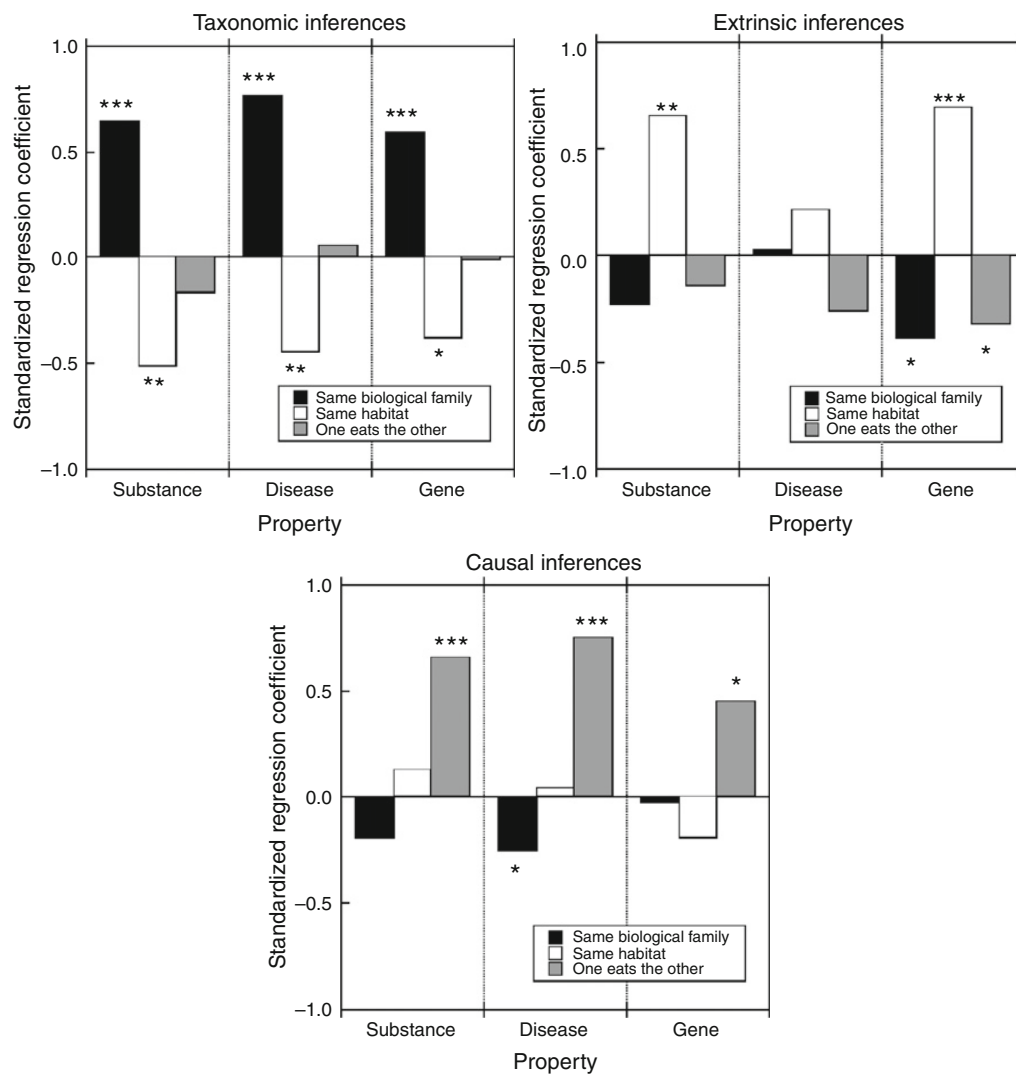


Figure 5 Relations between salience of premise relations and frequency of taxonomic, extrinsic, and causal inferences in each property condition, Study Two. (Note: * $p < 0.05$, ** $p < 0.005$, *** $p < 0.0005$.)

demonstrate the interplay of background knowledge about distribution of properties on the one hand, and salient relations among premise categories on the other.

In contrast, for disease, frequency of extrinsic inferences was unrelated to any premise relations ($R^2 = 0.08$, $p = 0.422$). However, it is important to point out that even though extrinsic inferences about disease were not predicted by premise relations, their frequency was nevertheless relatively high. Thus, disease appears to independently promote extrinsic inferences. Such a pattern could be due to participants relying on a general theory—or overhypothesis (Goodman, 1955)—stating that diseases are distributed via

spatial or contextual relations, which would make extrinsic inferences appealing regardless of premise relations.

3.2.2.4.3. Causal inferences As seen in Figure 5, the way in which premise relations predicted causal inferences also varied by property, but less so. In all property conditions, generation of causal inferences increased with the salience of predation relations, and was unrelated to salience of shared habitat. Additionally, for participants reasoning about disease (but not substance or gene), causal inferences decreased with the salience of taxonomic relations (Substance: $R^2 = 0.52$, $p < 0.0001$; Disease: $R^2 = 0.68$, $p < 0.0001$; Gene: $R^2 = 0.19$, $p = 0.075$). In sum, causal reasoning was consistently promoted by salience of predation relations between premise categories, but unrelated to salience of shared habitat. This suggests that contextual similarity was necessary but not sufficient to promote causal inferences, which were rendered particularly tempting when participants were reminded of predator–prey interactions among premise species. This reminding may have provided a salient causal mechanism to explain a shared property. Even for those reasoning about genes, despite the relative dearth of causal inferences (3%), such inferences were still positively predicted by the salience of predation relations among premise species. Although effects of property on the kinds of knowledge recruited to guide causal inferences were not dramatic, they confirm that the nature of the property can influence the way premise relations are used to guide inference generation.

3.3. Summary: Effects of Property on Inference Generation

Results of Study Two show that property influenced inference generation at two levels. First, naïve theories about the nature of the properties affected the relative frequency with which participants generated taxonomic, extrinsic, and causal inferences. Reasoning about substance replicated Study One, whereas reasoning about genes strongly biased participants toward taxonomic inferences, and reasoning about disease promoted causal reasoning, but also resulted in a more complex and multidimensional inference pattern. Second, property influenced both the degree to which relations are recruited to guide inferences and the quality of the effects of premise relations on inferences, creating a property-specific facilitation/inhibition profile. In addition, for extrinsic and causal inferences, the effects of premise relations varied by property, whereas for taxonomic inferences, they did not. Finally, Study Two also replicated the overall distribution of inferences, and the effects of premise relations on inference generation, from Study One. Salient taxonomic relations increased taxonomic inferences, salient habitat relations increased extrinsic inferences, and salient predation

relations increased causal inferences. The one departure from Study One was the finding that salience of shared habitat consistently inhibited taxonomic inferences.

4. INFERENCE GENERATION: CONCLUSIONS AND IMPLICATIONS

In two experiments utilizing a novel open-ended induction task we have demonstrated that salient relations among premise categories, and the nature of the property being projected, both guide and constrain the ways in which people generate inductive inferences about novel properties of animals. In this section, we summarize our main findings about the process of inference generation and discuss possible implications for the broader study of inductive reasoning.

4.1. What Have We Learned About Inference Generation?

In contrast to traditional methods used in the study of inductive inference, which require participants to evaluate the strength of inductive arguments, participants in our open-ended induction task generated their own inferences from the premise categories and properties we supplied. This approach encouraged them to generate a variety of inferences. Not surprisingly, taxonomic inferences—based on common category membership or shared intrinsic features—were generated most frequently (e.g., an inference from lemming/snowy owl to “other species of owl and similar species of lemming because of biological similarities between similar animals” or from tiger/clownfish to “a zebra because clownfish and tigers both have stripes. A zebra also has stripes”). Extrinsic inferences—based on shared situational or contextual features—were also quite common (e.g., an inference from lobster/tuna to “crabs, catfish, salmon, oysters, shrimp, because they all live in similar environmental conditions”). Perhaps, most striking was the finding that 20% of inferences generated by participants were based on causal relatedness or interaction (e.g., projecting a substance from salmon/black bear to “other bears and fish, because the bear might get substance A in their bloodstream by eating salmon, which also has substance A. So any other animal that eats salmon would probably have it also” or projecting a disease from ant/anteater to “birds because the disease may come from the ants themselves. By eating them the anteater got the disease, as would birds”). Clearly, a broad range of knowledge is used in the process of generating inductive inferences.

Moreover, the type of knowledge used to generate inferences varied systematically with the specifics of each inductive problem. Salient relations

among premise categories had a pronounced effect on the nature of inferences generated from those categories. Participants often explicitly referred to relations among premise species to explain their inferences. For example, one participant projected a substance from humpback whale/squirrel to “other mammals because whales and squirrels are both mammals.” Another projected a substance from owl/deer to “rabbit because all are found in woods.” A third projected a substance from elephant/crocodile to “rhino, hippo, alligator because all have tough, thick skins. Maybe substance E has to do with producing leathery skin.” Even more telling was the fact that many participants found themselves at a loss to generate an inference from an unrelated premise pair. The response of one participant, when confronted with the bullfrog/chipmunk pair, was typical: “No clue. I can’t think of a relationship between the two.” Indeed, the links between premise relations and inferences were quite specific. The salience of taxonomic relatedness consistently predicted taxonomic inferences, the salience of shared habitat consistently predicted extrinsic inferences, and the salience of predation relations consistently predicted causal inferences. Premise relations also had inhibitory effects. Most strikingly, salience of shared habitat reliably (in Study Two, at least) inhibited taxonomic inferences.

In addition to premise relations, property also had a large effect on the inferences participants generated. One way in which property influenced inference generation was to invoke naïve theories about how kinds of properties are likely to be distributed or transmitted. *Substance* served as a more or less neutral property; taxonomic and nontaxonomic inferences about substances were equally frequent (although more specifically, taxonomic inferences were more frequent than extrinsic inferences, which were more frequent than causal inferences). Compared to *substance*, participants reasoning about novel genes were biased in the direction of taxonomic inferences, whereas those reasoning about novel diseases were biased in the direction of causal inferences.

Even more strikingly, property influenced what relations among premise categories were seen as relevant. To illustrate, in the *gene* condition participants responding to the lion/zebra item tended to generate taxonomic inferences like “tiger, gazelle, horse, because they all have 4 legs, with similar features,” and “tigers and giraffes, because tigers and lions are similar animals, and zebras and giraffes are similar animals.” In contrast, in the *disease* condition participants tended to generate causal inferences from the same pair, like “hyenas, and lion prey, because lion could have gotten the disease from eating the zebra and spread it to any other animal it came in contact with,” and “Tigers/scavengers that eat zebras because zebras may carry the disease.” Thus, not only did different properties engender different inferences from the very same premise pair, but they also rendered different relations among the premise categories salient. Reasoning about genes rendered taxonomic knowledge salient because of what we believe about

genes and how they work; therefore, what seemed most relevant about lions and zebras is that both are quadrupedal mammals. In contrast, reasoning about disease rendered knowledge of spatiotemporal interactions salient because of what we believe about diseases and how they work; therefore, what seemed most salient about lions and zebras is the fact that lions eat zebras.

One final and striking finding was the frequency with which people generated vague inferences (e.g., one participant projected a substance from leaf-cutter ant/anteater to “an animal that is a predator to an anteater. I can’t think of any ‘cause I’m not an animal expert. Anteater eats ants, and they both have this substance. So I assume whatever eats an anteater will have it too or receive it by eating it.”) Inferences like this were quite common and reinforce the idea that people can generate sophisticated and subtle inferences based on framework theories, often despite the lack of specific knowledge. Indeed, this pattern of response is strongly reminiscent of the idea of overhypotheses. [Goodman \(1955\)](#) suggested that people possess abstract beliefs describing the scope of properties, and that these beliefs could constrain possible hypotheses about how properties could be projected. When one of our participants projected a novel gene from leaf-cutter ant/anteater to “other animals in the same family as the anteater and leaf cutter ant, because related animals have similar genes,” they unwittingly exemplified this idea perfectly.

In sum, our results suggest that people generate inductive inferences by extracting salient relations from premise categories in light of what they understand about the property being projected, and then drawing inferences consistent with those relations. This process emphasizes the degree to which categorical induction is both flexible and knowledge-driven. We next consider the broader implications of these findings.

4.2. Implications

Taken together, these results show that salient relations derived from comparison of premise categories—in concert with knowledge activated by the property being projected—provide important constraints on the generation of inductive inferences. In some sense, these results should be reassuring in that they reinforce findings that have emerged from the use of argument evaluation. We knew that property influenced how people evaluate arguments (e.g., [Heit & Rubinstein, 1994](#); [Kalish & Gelman, 1992](#); [Ross & Murphy, 1999](#); [Shafto & Coley, 2003](#); [Shafto, Coley, & Baldwin, 2007](#)), and now we know it also influences how they generate inferences. We knew that premise relations had an impact on argument evaluation ([McDonald et al., 1996](#); [Medin et al., 2003](#)), and now we know they also have an impact on inference generation. In other words, the picture of inductive reasoning that emerges from considering inference generation in

addition to argument evaluation seems to be a coherent one. However, we believe that our perspective has also highlighted aspects of inductive reasoning that might otherwise have remained in the shadows.

4.2.1. Salience of Taxonomy in Category-Based Induction

Traditional accounts have emphasized the role of taxonomic similarity in evaluating category-based inductive arguments. In contrast, our results clearly show that when generating inferences, participants spontaneously appealed to extrinsic similarity and causal relatedness as often as taxonomic similarity. In particular, the prevalence of causal reasoning in these experiments is surprising given previous research showing such reasoning is common among experts, but rare among folk biological novices like the undergraduates who participated in these experiments (e.g., Coley, Shafto, et al., 2005; Coley, Vitkin, et al., 2005; Coley et al., 1999). Past research—utilizing argument evaluation—has shown that experts tend to flexibly utilize knowledge of taxonomic, extrinsic, and causal relations, whereas novices are strongly biased toward taxonomic inferences (e.g., López et al., 1997; Shafto & Coley, 2003). As discussed above, forced-choice or argument-evaluation tasks require participants to recognize relations between given premise and conclusion categories. In contrast, our task allowed participants to generate their own inferences, and the way we coded responses gave participants credit for the knowledge underlying their inferences, even if it was vague (e.g., projecting a disease from parrot/toucan to “other birds that live in the tropical climates” or from newt/box turtle to “other creatures that eat newts and box turtles. . .”) or factually incorrect (e.g., projecting a substance from snowy owl/lemming to “all owls because lemmings and snowy owls are both owls” or a disease from penguin/herring to “ostriches and guinea hens, because ostriches and penguins both can’t fly, and I’m not sure what a herring is but I think it might be related to a guinea hen”). Thus, despite the lack of specific factual knowledge about tropical birds, what might eat a box turtle, or what a lemming is, this format enabled participants to nevertheless generate and articulate relatively sophisticated causal and extrinsic inferences. This suggests that the relative paucity of ecological and causal reasoning among folk-biologically naïve participants in previous research may be due in part to the fact that they were being asked to recognize such relations, rather than generate them.

Besides potentially taking taxonomic *inferences* down a peg or two, our results also have implications for the salience of taxonomic *relations*. A number of studies have shown that taxonomic knowledge dominates other conceptual relations in terms of salience, speed of access (e.g., Ross & Murphy, 1999; Vitkin et al., 2005), and use in guiding inductive inferences (e.g., Shafto, Coley, & Baldwin, 2007). In contrast, our results provide little evidence that taxonomic relations between premise categories are privileged in terms of their impact on inference generation. Indeed, if

anything, we observed the opposite; the presence of salient ecological relations among premise was more likely to suppress taxonomic reasoning than vice versa.

We have several thoughts on these findings. First, because they were generating their own inferences, rather than evaluating our best guesses as to what they deem plausible arguments, participants were not constrained by lack of specific knowledge (nor, indeed, by facts or reality). As such, informationally vague yet causally sophisticated inferences—which would not be detectable in an argument evaluation paradigm—were relatively common. Second, because our task did not involve any time pressure or speeded responding, and was in fact deliberately reflective in that participants were asked to explain their inferences as well as generate them—baseline differences in knowledge accessibility (Shafto, Coley, & Baldwin, 2007; Shafto, Coley, & Vitkin, 2007) were probably not a factor. In other words, the results of tasks involving time pressure suggest that taxonomic knowledge might be initially more accessible, but our results suggest that given sufficient time, other knowledge is readily recruited to guide inductive inferences.

Third, inference generation may involve stronger differentiation between specific kinds of relatedness than argument evaluation does. Beyond assessing whether the premises are sufficiently related in a general way consistent with the projected property—as required for argument evaluation—our participants had to generate novel hypotheses and then articulate the relationships between premises and their hypotheses. Internally labeling taxonomic and ecological relations among premises as such might promote discounting of irrelevant relations and focus attention on more relevant relations. If taxonomic relations are highly salient, yet on some occasions they are viewed as irrelevant for projection, such a selective approach would diminish the effect of taxonomic relatedness on inference generation compared to argument evaluation.

Finally, it may also be that, more generally, taxonomic and relational categories have differing cognitive functions. Ross and Murphy (1999) point out that in the domain of food, taxonomic categories—based on intrinsic properties—are useful for categorization and identification, whereas script categories—based on habitual co-occurrence in space and time—are useful for generating solutions to problems like “what should I have for breakfast?” Likewise, in folk biology, relational categories like pond animals, or even noncategorical relations like predator–prey, may be especially useful for generating solutions to problems like “what other species are likely to have this substance/disease” because they embody relations seen as causally relevant for explaining how such properties could come to be shared among species. As such, by focusing on inference generation we may have tapped into precisely the kind of cognitive task that such categories are most useful for.

4.2.2. Challenges for Models of Category-Based Induction

4.2.2.1. What Needs to be Explained? Our findings expand the range of inductive phenomena that any successful theory of inductive reasoning needs to explain. First of all, any successful model has to incorporate a variety of potential bases for inductive inference; at the very least, these must include both taxonomic and extrinsic similarity, and causal relations, but we make no claim about whether this list is exhaustive.⁵ We emphatically reinforce the point (made elsewhere, e.g., Coley, Shafto, et al., 2005; Coley, Vitkin, et al., 2005; Lassaline, 1996; Medin et al., 2003) that similarity alone—no matter how flexibly conceived—cannot be sufficient to explain inductive reasoning.

Second, any theory of inductive reasoning must take into account the fact that the kinds of knowledge used to generate an inference depend critically on the property being projected and on salient relations among premise categories. We think that the kind of models being developed by Tenenbaum and colleagues (e.g., Griffiths & Tenenbaum, 2005; Shafto et al., 2008; Tenenbaum et al., 2006)—which rely on property to indicate which knowledge structure might be most relevant for assessing a given argument—are a step in the right direction. However, our results suggest that not only does property influence the kinds of knowledge upon which participants base their inferences, but it also influences their interpretation of relations among premise categories, and the way in which those relations influence inferences. Any successful theory of inductive reasoning must take into account this interplay between domain knowledge, beliefs about premise relatedness, and beliefs about the property.

Finally, any successful theory must take into account the fact that inferences vary widely in their specificity. This is reminiscent of Keil's (Keil, 2003; Rozenblit & Keil, 2002) proposals about the "illusion of explanatory depth" in the sense that participants probably do not have a detailed understanding about mechanisms of epidemiology or genetics, or about specifics of food webs, but that relatively abstract and cursory framework principles can nevertheless effectively guide inductive reasoning. Likewise, this finding fits with Coley and colleagues' work on hierarchical induction (Coley et al., 1997, 2004). Although they investigated a very different issue—namely the degree to which knowledge of concepts at different levels of abstraction corresponded to the relative strength of inferences to those concepts—they found that the level at which participants expected category members to share novel properties differed from participants' knowledge of actual properties shared by category members. Coley et al. (2004) conclude that "inductive inference is driven by

⁵ It may well be exhaustive in the context of folk biological inductive reasoning, but for other domains such as reasoning about artifacts or about social categories, no doubt other kinds of inferences might be generated.

expectations about conceptual structure that go beyond what is known about particular category members” (p. 249). Our findings on inference generation support that conclusion.

4.2.2.2. Focus on Process By focusing on how people generate inductive inferences—how they use what they know to make sensible guesses about what they do not know—we hope to direct some attention to the little-studied but critical issue of process in inductive reasoning. In most previous work on inductive inference, the target behavior has been evaluation of a complete argument or choice from among a limited set of alternatives. As such, the questions to be explained—and therefore the natural and appropriate goals of empirical and theoretical investigation—have concerned factors that predict argument strength ratings or choices. These have tended to focus on characteristics of arguments (and implicitly or explicitly, the interactions of these characteristics with the knowledge of the reasoner) that render them strong or weak (e.g., McDonald et al., 1996; Medin et al., 2003; Osherson et al., 1990; Sloman, 1993).

There is no reason why studies of argument evaluation cannot in principle examine process issues; indeed, a few have done so. For instance, Shafto, Coley, and Baldwin (2007) showed that time pressure lowers strength ratings for inductive arguments based on extrinsic relations, but has no effect on arguments based on taxonomic relations. Likewise, Feeney et al. (2010) have shown that premise reading times are related to the changes in argument strength brought about by those premises; larger changes in argument strength are associated with longer reading times, and presumable deeper processing. Rather, the focus on process inherent in the inference generation approach is more a difference in emphasis.

When the target behavior is inference generation, rather than argument evaluation, the questions to be explained concern what inferences are generated under different conditions, why they are generated, and how they are generated. These questions naturally focus attention on the characteristics of the process of inductive inference. We have assumed that the process of inference generation involves accessing knowledge about premise categories and the property being projected, making decisions about what knowledge is relevant, and then generating an actual response. Clearly, there are many processes involved in even this cursory description. These include searching semantic memory for relevant knowledge, comparing premise categories for salient relations, accessing explanatory theories about the nature of the property, and potentially searching for relevant conclusion categories once a basis for inference has been determined, to name just a few.

At the moment, we lack answers about the role of any of these processes in inference generation. We do not, however, lack questions. For example,

what is the mechanism by which properties constrain inference generation? Do they focus the search at a relatively early point and thereby limit the candidate inferences that are evaluated? Or do they serve mainly to cull an exhaustive list of possible inferences generated via premise comparison down to a few likely candidates? We hope that our initial look at inference generation prepares the ground and invites further work examining the processes that underlie flexible inductive reasoning.

4.3. Conclusions

At the risk of repeating ourselves, we cannot and should not base a psychology of inductive reasoning solely on studies of argument evaluation. In the hope of putting “reasoning” back into the study of category-based induction, we have presented an initial look at how people generate inductive inferences. Our results show that salient relations derived from comparison of premise categories—in concert with knowledge about the property being projected—provide important constraints on the generation of inductive inferences. We have also shown that such inferences vary widely in their specificity, and make contact with a broad range of real-world knowledge. In taking this approach, we hope to draw attention to the process of inductive reasoning as well as the outcome, and to emphasize the knowledge-driven and creative nature of human inductive inference.

ACKNOWLEDGMENTS

This chapter is based upon work supported by the National Science Foundation under Grant No. 0236338. We are indebted to Anna Vitkin and Allison Baker for their important contributions to the research reported here. We thank Brett Hayes and Gregory Murphy for careful and thoughtful comments on previous incarnations of this paper. We are especially grateful to Kaitlyn Amato, Yui Anzai, Nicole Ciampanelli, Lindsey Davis, Konstantin Feigin, Ruiwen Hu, Janelle LaMarche, Brianna Roche, Claire Seaton, Carissa Shafto, Courtney Steller, and Jennelle Yopchick for their Herculean efforts to collect and code the data reported here.

REFERENCES

- Chomsky, N. (1980). *Rules and representations*. Oxford: Basil Blackwell.
- Coley, J. D., Hayes, B., Lawson, C., & Moloney, M. (2004). Knowledge, expectations, and inductive inferences within conceptual hierarchies. *Cognition*, *90*, 217–253.
- Coley, J. D., Medin, D. L., & Atran, S. (1997). Does rank have its privilege? Inductive inferences within folkbiological taxonomies. *Cognition*, *64*, 73–112.
- Coley, J. D., Medin, D. L., Proffitt, J. B., Lynch, E. B., & Atran, S. (1999). Inductive reasoning in folkbiological thought. In D. L. Medin & S. Atran (Eds.), *Folkbiology* (pp. 205–232). Cambridge, MA: MIT Press.

- Coley, J. D., Shafto, P., Stepanova, O., & Barraff, E. (2005). Knowledge and category-based induction. In W. Ahn, R. L. Goldstone, B. C. Love, A. B. Markman, & P. Wolff (Eds.), *Categorization inside and outside the laboratory: Essays in honor of Douglas L. Medin* (pp. 69–85). Washington, DC: American Psychological Association.
- Coley, J. D., Vitkin, A. Z., Seaton, C. E., & Yopchick, J. E. (2005). Effects of experience on relational inferences in children: The case of folk biology. In B. G. Bara, L. Barsalou, & M. Bucciarelli (Eds.), *Proceedings of the 27th annual conference of the Cognitive Science Society* (pp. 471–475). Mahwah, NJ: Lawrence Erlbaum Associates.
- Coley, J. D., Vitkin, A. Z., Vasilyeva, N. Y., & Amato, K. (2007). Experience increases flexible ecological reasoning. In: *Paper presented at the 15th Biennial Conference of the Australasian Human Development Association*. Sydney, New South Wales, Australia.
- Feeney, A., Coley, J. D., & Crisp, A. (2010). The relevance theory of category-based induction: Evidence from garden path arguments. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 36.
- Feeney, A., Shafto, P., & Dunning, D. (2007). Who is susceptible to the conjunction fallacy in category-based induction? *Psychonomic Bulletin & Review*, 14, 884–889.
- Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought*. New York: Oxford University Press.
- Gelman, S. A., & Coley, J. D. (1990). The importance of knowing dodo is a bird: Categories and inferences in 2-year-old children. *Developmental Psychology*, 26, 796–804.
- Goodman, N. (1955). *Fact, fiction, and forecast*. Indianapolis, IN: Bobbs-Merrill.
- Griffiths, T. L., & Tenenbaum, J. B. (2005). Structure and strength in causal induction. *Cognitive Psychology*, 51, 354–384.
- Heibeck, T., & Markman, E. (1987). Word learning in children: An examination of fast mapping. *Child Development*, 58, 1021–1024.
- Heit, E. (2000). Properties of inductive reasoning. *Psychonomic Bulletin & Review*, 7, 569–592.
- Heit, E., & Feeney, A. (2005). Relations between premise similarity and inductive strength. *Psychonomic Bulletin & Review*, 12, 340–344.
- Heit, E., & Rubinstein, J. (1994). Similarity and property effects in inductive reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 411–422.
- Kalish, C. W., & Gelman, S. A. (1992). On wooden pillows: Multiple classifications and children's category-based induction. *Child Development*, 75, 1871–1885.
- Keil, F. C. (1981). Constraints on knowledge and cognitive development. *Psychological Review*, 88, 197–227.
- Keil, F. C. (2003). Folkscience: Coarse interpretations of a complex reality. *Trends in Cognitive Sciences*, 7, 368–373.
- Kemp, C., Perfors, A., & Tenenbaum, J. B. (2007). Learning overhypotheses with hierarchical Bayesian models. *Developmental Science*, 10, 307–321.
- Lassaline, M. E. (1996). Structural alignment in induction and similarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 754–770.
- Lin, E. L., & Murphy, G. L. (2001). Thematic relations in adults' concepts. *Journal of Experimental Psychology: General*, 130, 3–28.
- López, A., Atran, S., Coley, J. D., Medin, D., & Smith, E. E. (1997). The tree of life: Universal and cultural features of folkbiological taxonomies and inductions. *Cognitive Psychology*, 32, 251–295.
- McDonald, J., Samuels, M., & Rispoli, J. (1996). A hypothesis-assessment model of categorical argument strength. *Cognition*, 59, 199–217.
- Medin, D., Coley, J. D., Storms, G., & Hayes, B. (2003). A relevance theory of induction. *Psychonomic Bulletin & Review*, 10, 517–532.

- Muratore, T. M., & Coley, J. D. (2009). The role of knowledge in folk biological induction. In: *Paper presented at the international conference on Biological Understanding and Theory of Mind*. Reims, France.
- Nguyen, S. P., & Murphy, G. L. (2003). An apple is more than just a fruit: Cross-classification in children's concepts. *Child Development, 6*, 1783–1806.
- Osherson, D. N., Smith, E. E., Wilkie, O., Lopez, A., & Shafir, E. (1990). Category-based induction. *Psychological Review, 97*, 185–200.
- Proffitt, J. B., Coley, J. D., & Medin, D. L. (2000). Expertise and category-based induction. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 811–828.
- Rips, L. J. (1975). Inductive judgments about natural categories. *Journal of Verbal Learning and Verbal Behavior, 14*, 665–681.
- Ross, B. H., & Murphy, G. L. (1999). Food for thought: Cross-classification and category organization in a complex real-world domain. *Cognitive Psychology, 38*, 495–553.
- Rozenblit, L. R., & Keil, F. C. (2002). The misunderstood limits of folk science: An illusion of explanatory depth. *Cognitive Science, 26*, 521–562.
- 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: Learning, Memory, and Cognition, 29*, 641–649.
- Shafto, P., Coley, J. D., & Baldwin, D. (2007). Effects of time pressure on context-sensitive property induction. *Psychonomic Bulletin & Review, 14*, 890–894.
- Shafto, P., Coley, J. D., & Vitkin, A. (2007). Availability in category-based induction. In A. Feeney & E. Heit (Eds.), *Inductive reasoning: Experimental, developmental, and computational approaches* (pp. 114–136). Cambridge University Press.
- Shafto, P., Kemp, C., Bonawitz, E. B., Coley, J. D., & Tenenbaum, J. B. (2008). Inductive reasoning about causally transmitted properties. *Cognition, 109*, 175–192.
- Sloman, S. A. (1993). Feature-based induction. *Cognitive Psychology, 25*, 231–280.
- Sloman, S. A. (1994). When explanations compete: The role of explanatory coherence on judgments of likelihood. *Cognition, 52*, 1–21.
- 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.
- Tenenbaum, J. B., Griffiths, T. L., & Kemp, C. (2006). Theory-based Bayesian models of inductive learning and reasoning. *Trends in Cognitive Sciences, 10*, 309–318.
- Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5*, 207–232.
- Vitkin, A., Coley, J. D., & Feigin, K. (2005). Accessibility of taxonomic and script knowledge in the domain of food. *Paper presented at the 46th annual meeting of the Psychonomic Society, Toronto*.
- Vitkin, A. Z., Vasilyeva, N. Y., & Coley, J. D. (2007). Experience and the development of flexible inductive reasoning in biology. *Paper presented at the annual meeting of British Psychological Society 2007 Developmental Section Conference*. Plymouth, UK.