THE SAME COGNITIVE MECHANISM UNDERLIES REASONING ABOUT THE FEATURES OF KINDS AND THE TRAITS OF INDIVIDUALS

BY

LIN BIAN

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Arts in Psychology
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2014

Urbana, Illinois

Advisor:

Professor Andrei Cimpian

ABSTRACT

Learning about categories of things in the world (e.g., that lions have manes) and learning about the individuals around us (e.g., that the new neighbor is friendly) are both crucial cognitive tasks whose output enables people to behave adaptively in many novel circumstances. Aside from their importance to our lives, however, these two tasks seem to have little in common: Reasoning about the features of categories appears to rely on a set of computations that is entirely distinct from those involved in reasoning about the traits of individuals. Consistent with this impression, these judgments have been studied in near-complete isolation. Here, however, we propose that inferences about the features of kinds and the traits of individuals are in fact computed by the *same cognitive mechanism*. To test this hypothesis, we investigated whether kind-feature and individual-trait inferences share two distinctive signatures: namely, whether they are both facilitated by features/behaviors that are (1) threatening or (2) unique. Five experiments provided evidence for this prediction. The results also suggested that these signatures were not shared by other types of inferences. That is, the information that facilitated participants' kind-feature and individual-trait inferences did not similarly facilitate quantified inferences concerning whether *all*/*some* members of a kind display a feature and whether an individual *always*/*sometimes* performs a behavior. By suggesting that inferences about the features of kinds and the traits of individuals share a cognitive source, these studies open possibilities for dialogue between the independent research traditions that have investigated these important aspects of human thought.

ACKNOWLEDGMENTS

I owe the success of this project to the many people who have contributed their insights and help over the past three years. First I would like to express the deepest appreciation to my advisor, Dr. Andrei Cimpian, who gave me persistent guidance and encouragement during the entire process. Working with him is enjoyable and fruitful, through which I am consistently growing to become an accomplished graduate student, an independent thinker, and above all, a real researcher in the beloved field of developmental psychology. I would also like to express my gratitude to my co-advisor, Dr. Renee Baillargeon. Her advice on both research as well as on my career have been priceless.

I am also grateful to my committee, Dr. Cynthia Fisher, who is tremendously generous with her knowledge and provides insightful feedback on my thesis. I have greatly benefited from her fresh perspectives on the subject.

Special thanks are given to the members of the Cognitive Development Lab, without whom my project would not have been nearly successful. I always cherish the special opportunity to be group mates with them.

And finally, thanks to my parents, my boyfriend and numerous friends who have been my faithful company all along this voyage. Their streaming love and support have accumulated to be a lasting source of my courage, confidence and the impetus to achieve more and bigger progress.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Humans routinely use limited first-hand evidence to construct broader knowledge about their natural and social worlds. In fact, the ability to make such inferential leaps is a cornerstone of our species' success in predicting and controlling its environment. Two types of these specific-to-general inferences seem to play particularly prominent roles in our mental lives: First, we draw inferences about the features of *kinds of things* in the world on the basis of evidence about specific members of those kinds. After learning, for instance, that one's cat likes to sleep in cardboard boxes, one might conclude that cats, as a kind, like to do the same. Second, we draw inferences about the dispositions of the *individuals* around us by observing their behavior in particular situations. For instance, after seeing a colleague donate money to a charity, one might conclude that she is generous. Our goal in the present research was to investigate the mechanisms that underlie these inferences. Specifically, we ask: What is the relationship between the mechanism that enables people to learn about the features of kinds and the mechanism that enables them to learn about the dispositions of individuals?

Several considerations suggest that the two mechanisms have little, if anything, in common. Intuitively, for example, kind-feature and individual-trait inferences appear to be entirely different sorts of judgments. The former involves generalizing across individuals to infer a feature of a broad category, while the latter involves making an inference from an altogether different sort of entity (namely, a situation or time slice) to an altogether different sort of target (namely, a disposition of a single individual).[1] Consistent with these surface differences, kind-feature and individual-trait inferences have been the focus of largely separate research traditions,

---

[1] We use the terms *trait* and *disposition* interchangeably to refer to a stable quality or characteristic of an individual that makes it likely for that individual to behave in certain ways.

with minimal overlap in terms of the mechanisms proposed and investigated. Inductive inferences about the features of kinds have been a longstanding focus of cognitive psychologists, whose theorizing has focused on factors such as the prototypicality (vs. atypicality) and diversity (vs. homogeneity) of the evidence, or the featural overlap between the evidence and the kind about which one is making inferences (e.g., Carey, 1985; Gelman, 1988; Markman, 1989; Osherson, Smith, Wilkie, López, & Shafir, 1990; Rips, 1975; Sloman, 1993). Inferences about the features of kinds have also been hypothesized to stem from special-purpose cognitive modules that have evolved specifically because of their role in enabling fast, efficient judgments about kinds in domains important to survival (e.g., Atran, 1998; Pinker, 1994, 1997; Setoh, Wu, Baillargeon, & R. Gelman, 2013).

In contrast, inferences about the traits of individuals have been a major concern for social psychologists, whose proposals have tended to focus on an entirely different set of issues. For example, prominent work in this area has highlighted the relationship of these inferences to patterns of covariation in the evidence (e.g., has the colleague donated to other charities?; Heider, 1958; Kelley, 1967, 1973), their automaticity (e.g., Gilbert & Malone, 1995; Trope & Gaunt, 2000; Uleman, 1999), and the errors they introduce into our judgments (e.g., Gilbert & Malone, 1995; Jones, 1979; Jones & Harris, 1967; Ross, 1977). Some have suggested that these inferences also originate from an innate module, except a different one than the module putatively responsible for kind inferences: a biologically adapted mechanism that emerged over the course of our species' evolutionary history because it facilitated reasoning about other individuals' psychological states and characteristics, including their dispositions (e.g., Hamlin, Wynn, & Bloom, 2007; Kuhlmeier, Wynn, & Bloom, 2003; A. Leslie, 1994; Luo & Baillargeon, 2005; Pinker, 1994, 1997).

Thus, the existing literature seems to point to fundamentally different cognitive sources for people's ability to reason about the features of kinds and traits of individuals. Here, however, we propose that kind-feature and individual-trait inferences have much more in common than previous research would lead one to expect. More precisely, our main hypothesis is that these two sorts of judgments may in fact stem from the *same cognitive mechanism* that can output inferences at different levels (about individuals vs. kinds), depending on factors such as the content and scope of the evidence or the reasoner's goals in the moment. In other words, we argue that the differences in current psychological approaches to studying kind-feature and individual-trait inferences do not actually correspond to a difference in the mechanism underlying these inferences—rather, these discrepant approaches may mask a common source. To motivate this Shared Inferential Mechanism (SIM) Hypothesis, we provide a brief, selective overview of previous research on inferences about the features of kinds and the dispositions of individuals. Specifically, we summarize arguments and evidence pertaining to three dimensions of overlap between these two types of inferences: (1) their weak relationship to statistical evidence; (2) their resistance to counterexamples; and (3) their linguistic expression. We then describe five experiments we conducted to provide a formal test of the SIM Hypothesis.

Weak Dependence on Statistical Evidence

Inferences about the features of kinds and the traits of individuals are similar by being only weakly dependent on the *frequency* of the relevant features or behaviors. People sometimes make inferences about kinds even when the features in question are very infrequent, whereas other times they do not make such inferences even when there is considerable statistical evidence to support them (e.g., Abelson & Kanouse, 1966; Cimpian, Brandone, & Gelman, 2010; Cimpian, Gelman, & Brandone, 2010; see also Leslie, 2008). For example, people agree that mosquitoes, as a kind, carry malaria, even though very few mosquitoes are in fact carriers of this disease. In contrast, people don't think that bees, as a kind, are sterile, even though the vast majority of them actually are. Similarly inconsistent uses of statistical evidence are found with respect to inferences about the traits of individuals. Often, a single event or behavior is sufficient for observers to draw conclusions about the enduring dispositions of the actors involved (e.g., Jones & Harris, 1967; Ross, Amabile, & Steinmetz, 1977; Trope & Gaunt, 2000; for reviews, see Gawronski, 2004; Gilbert & Malone, 1995; Trope & Gaunt, 2007). Other times, however, even multiple behaviors consistent with a dispositional inference may not be sufficient for that inference to be formulated (e.g., Gidron, Koehler, & Tversky, 1993; Reeder & Brewer, 1979; Rothbart & Park, 1986; Tausch, Kenworthy, & Hewstone, 2007). For example, although we might conclude that a person is *superstitious* based on a single behavior, much more evidence is typically needed to formulate an inference about a person's *punctuality*—showing up on time to a single meeting is far from sufficient, and several meetings may not be enough either (e.g., Rothbart & Park, 1986).

Note that this loose relationship with the underlying frequencies (whereby a certain level of statistical evidence suffices for some kind-feature or individual-trait inferences but not others) is rather unusual. For many other types of inferences, their acceptability depends almost entirely on whether a certain frequency threshold is met. To illustrate, it would be legitimate to infer that *some*, *most*, *every*, *all*, etc., members of a kind display a feature if and only if a certain number of them do, regardless of what that feature is (e.g., Cimpian et al., 2010). The same point applies to inferences over situations about whether a person *sometimes*, *occasionally*, *always*, etc., does something—all that matters is how often this person performs the relevant action. Thus, the inconsistencies in the evidentiary threshold for inductive inferences about the features of kinds and the dispositions of individuals are a distinctive feature of these inferences,[2] perhaps suggestive of a common source for the two.

Resistance to Counterexamples

Related to the point about weak dependence on statistical evidence, both kind-feature and individual-trait inferences are—once formulated—resistant to contradictory evidence. To begin, inferences about the features of kinds are typically not falsified by counterexamples. Encountering a cat that doesn't meow or a dog with only three legs, for instance, does not cause people to reconsider whether cats meow and dogs have four legs. This robustness to counterexamples has been documented even when the inferences concern novel kinds: Four-year-olds, for example, hold on to such inferences (e.g., that *pagons* are friendly) when they are presented with inconsistent evidence (e.g., a pagon that isn't friendly; Chambers, Graham & Turner, 2008). Likewise, inferences about the traits of individuals are often resistant to counterevidence—that is, they are often maintained even in the face of behaviors that are inconsistent with the relevant traits. If we believe someone to be manipulative, for instance, behaviors that are inconsistent with this inferred disposition (e.g., an honest act) seldom change this belief (e.g., Garlick, 1993; Gidron, Koehler, & Tversky, 1993; Hayden & Michel, 1976).

Linguistic Expression

Semantic analyses of the typical linguistic expression of kind-feature and individual-trait inferences have suggested that there are deep structural similarities between them (e.g., Dahl, 1975; Krifka et al., 1995; Lawler, 1973; Leslie, 2013). In fact, semanticists often classify statements expressing these two inferences as part of the same linguistic phenomenon,

---

[2] This feature is not unique, however. Inferences quantified with *many* and *often* (e.g., Cova & Egré, under review; Keenan & Stavi, 1986) may be similar in this respect. For example, if someone drinks water once a day, it does not seem legitimate to infer that they drink water often. In contrast, if someone goes hiking once a day, it seems quite legitimate to infer that they go hiking often.

genericity. According to a standard treatment of genericity, the meaning of statements conveying kind-feature and individual-trait inferences can be formalized using the same (unpronounced) generic quantifier **GEN** (where **GEN** stands for *generic*; e.g., Krifka et al., 1995; Leslie, 2013).[3]

In the case of statements expressing inferences about the features of kinds, **GEN** quantifies over the members of the relevant kinds. For example, the logical form of a sentence such as *Cats meow* might be represented as,

GEN[*x*] [*x* is a cat] [*x* meows],

where [*x* is a cat] specifies the domain over which generic quantification takes place (i.e., individual members of the cat kind) and [*x* meows] specifies the property being attributed to members of the set that is determined by the combination of the generic quantifier **GEN** and the domain restrictor [*x* is a cat]. A similar analysis is applied to statements expressing inferences about the traits of individuals, except in this case **GEN** quantifies over situations or occasions. For example, a sentence such as *Mary is manipulative* might be represented as,

GEN[*s*] [*s* is a situation involving Mary] [Mary is being manipulative in *s*].

Here, the generic quantification is defined over the variable *s* (i.e., situations); the domain over which **GEN** operates is defined by the restrictor [*s* is a situation involving Mary]; and the property being quantified is defined by the final element, [Mary is being manipulative in *s*]. As may already be apparent, the logical form of this sentence is analogous to that of *Cats meow*, provided above. This analysis thus suggests that the meaning of statements that convey judgments about the features of kinds and the dispositions of individuals can be spelled out in terms of the same unvoiced, implicit quantifier. This idea is consistent with the present claim that these inferences are the products of the same cognitive mechanism but stands in contrast to their

[3] To clarify, we are not claiming that inferences about the features of kinds and the traits of individuals are expressed in language *exclusively* via statements that can be formalized with **GEN**. These statements are, however, a very common linguistic vehicle for such inferences, and thus their semantics are relevant to the present argument.

surface differences, as well as to their prior treatment in the psychological literature.

The Present Research: Testing the Shared Inferential Mechanism Hypothesis

The research reviewed above suggests that inferences about the features of kinds and the traits of individuals have several features in common. However, to our knowledge, no empirical studies have explored whether these inferences stem from the same cognitive mechanism. Not surprisingly, then, almost no experiments have directly compared the two types of inferences in a systematic way (for an exception, see Lawson & Kalish, 2006). Thus, our goal here was to provide a first test of the SIM Hypothesis.

The logic of this test was as follows: Prior work in cognitive psychology and linguistics has suggested that inferences about the features of kinds exhibit a peculiar characteristic—namely, they are sensitive to select aspects of the semantic content of the properties about which the inferences are being made. Specifically, properties that are (1) strongly negatively valenced (dangerous, threatening, etc.) or (2) unique/distinctive are inferred to characterize kinds based on less evidence than properties that are otherwise similar but are neither negative nor distinctive (e.g., Cimpian et al., 2010; Leslie, 2008). For example, people generalize a dangerous-sounding property such as *having poisonous red feathers* to a novel kind more readily than a property such as *having red feathers*, even when given the same amount of evidence (e.g., that 30% of the kind's members display the feature). Similarly, a property such as *having red feathers* is more readily generalized to a novel kind when it is said to be present only among members of that kind (i.e., when it is unique/distinctive) than when members of other kinds are said to display it as well. To return to our argument, if inferences about the properties of kinds and the traits of individuals are computed via the same mechanism (as we propose), then content should matter in exactly the same way for inferences about the dispositional traits of individuals. That is, events involving dangerous or distinctive behaviors should lead to inferences about enduring dispositions particularly readily compared to events involving other behaviors. In sum, the present experiments test the hypothesis that kind-feature and individual-trait inferences are generated via the same mechanism by varying the content of the information over which these inferences are computed while equating all other variables (e.g., the strength of the evidence). The prediction of the SIM Hypothesis is that the two inferences would be affected by these content manipulations in a similar manner.

The prior evidence on whether, and how, inferences about individual dispositions are influenced by content is mixed—particularly with respect to the valence of the content. For instance, some studies have suggested a negativity bias, wherein people require less evidence to make inferences about negatively-valenced dispositions than positively-valenced ones (a result that is broadly consistent with our hypothesis; e.g., Aloise, 1993; Gidron et al., 1993; Rothbart & Park, 1986). Several linguists have made similar suggestions. For example, Carlson (1977) pointed out that less evidence seems needed to endorse the truth of a sentence such as *John beats small children* than of a sentence such as *John repairs cars*. Other studies, however, have found the exact opposite: a positivity bias (e.g., Heyman & Giles, 2004; Lockhart, Chang, & Story, 2002; Rholes & Ruble, 1984). For instance, Heyman and Giles (2004) found that both children and adults were more likely to make dispositional inferences based on positively- than negatively-valenced evidence (e.g., John getting a good grade vs. a bad grade; see Skowronski & Carlston, 1987, and Martijn, Spears, van der Pligt, & Jakobs, 1992, for additional discussion of these apparently contradictory findings).

With respect to distinctiveness, the prior evidence seems more consistently in line with our prediction that dispositional inferences would go through more easily for behaviors or events that are unique (e.g., Higgins & Bryant, 1982; McArthur, 1972; Orvis, Cunningham, & Kelley, 1975). For example, when told that Mary loved a certain movie but nobody else did (which is what Kelley [1967, 1973] called *low consensus* information), people are more likely to infer a dispositional trait (e.g., Mary loves movies) than if told that everyone else loved that movie as well (*high consensus* information). Although these results are consistent with our prediction, in none of these studies were participants also asked to make inferences about *kinds* based on analogous evidence. The multiple methodological differences between these studies and those investigating inferences about the features of kinds (e.g., Cimpian et al., 2010) make it difficult to draw conclusions about the SIM Hypothesis from this evidence. When also considering the contradictory findings with respect to valence, it is clear that the prior literature provides no more than encouraging hints regarding the prediction that judgments about the features of kinds and the traits of individuals are similarly affected by property content.

We conducted five experiments to test the SIM Hypothesis. Across these experiments, the likelihood of making kind-feature and individual-trait inferences was assessed using a procedure that was closely matched across the two types of inferences, which afforded strong comparisons between them. Experiments 1 and 3 investigated the contribution of danger and distinctiveness, respectively, to the likelihood of drawing inferences about the features of kinds and the traits of individuals. The results of these experiments supported our prediction: Danger and distinctiveness had an almost identical facilitative effect on the two inferences, as would be expected if they were the output of the same cognitive mechanism. Experiments 2 and 4 explored the specificity of this effect: Would danger and distinctiveness connotations also have an impact on the likelihood of making quantified[4] inferences over individuals (some/all *X*s do *Y*) and situations (*X* sometimes/always does *Y*)? The results of Experiments 2 and 4 found no evidence for an effect of property content on these quantified inferences, suggesting that inferences about the features of kinds and traits of individuals share a relatively unusual signature. The final experiment explored an unexpected difference between kind-feature and individual-trait inferences that emerged in Experiments 1 and 3—namely, the fact that inferences about the traits of individuals were made more frequently overall than inferences about the features of kinds, given equivalent levels of evidence. The results of Experiment 5 suggested that this difference is more likely to be due to ambiguities in the linguistic stimuli used to express these inferences rather than to differences in their cognitive source.

[4] For the rest of the paper, the term *quantified* will be used to refer to overt, explicit quantification (that is, quantification explicitly marked in language, as with *some* or *all*). Thus, generic quantification (e.g., Krifka et al., 1995) is excluded from the scope of this term as used from this point on.

CHAPTER 2

EXPERIMENT 1: THE EFFECT OF VALENCE ON KIND-FEATURE AND INDIVIDUAL-TRAIT INFERENCES

As a first test of the SIM Hypothesis, in this experiment we explored the influence of valence on inferences about the features of kinds and the traits of individuals. Participants were randomly assigned to one of two conditions, depending on the type of inference assessed. In the Kind-Features condition, participants were provided with evidence about the prevalence of a feature among the members of an unfamiliar kind (e.g., zorbs) and were then asked whether the corresponding kind-feature inference is valid. Half of the stimulus features were negatively-valenced, referring specifically to acts that are dangerous to humans (e.g., chopping people's heads off), and half were neutral in valence (e.g., climbing trees). In the Individual-Traits condition, participants were provided with evidence about the prevalence of a behavior in the life of an unfamiliar individual (e.g., Zorb) and were then asked whether the corresponding individual-trait inference is valid. As in the other condition, half of the stimulus behaviors were dangerous acts (e.g., chopping people's heads off), whereas the other half were neutral (e.g., climbing trees). In the context of this experiment, the prediction of the Shared Inferential Mechanism Hypothesis is that both kind-feature and individual-trait inferences would be made more readily based on evidence with connotations of danger.

Method

Participants. Participants (*N* = 198; 86 men, 112 women) were recruited from Amazon's Mechanical Turk service. They received $0.75 for participation. All participants reported being native English speakers on a prescreening question. An additional 36 participants were tested but excluded from the final sample because their IP addresses were from outside the United States (*n* = 7) or because they failed an attention check at the end of the study (see below; *n* = 29).

Materials, procedure, and design. The instructions for participants were identical in the Kind-Features and Individual-Traits conditions:

> In this study, we will tell you about some animals that live on a planet in a remote galaxy. For each question, you will be given some evidence and then you will be asked to judge if a certain conclusion follows from that evidence.

The two conditions diverged, however, in terms of the information provided on each of the 24 trials that followed the instruction screen (see Table 1 for an example). In the Kind-Features condition, participants were told about a kind (e.g., zorbs) and were then provided with evidence about the prevalence of a feature among its members (e.g., 30% of zorbs climb trees). Finally, the subjects were asked whether this evidence licenses the conclusion that the feature applies to the kind as a whole (see Table 1). Specifically, participants were asked whether the corresponding generic statement (e.g., "zorbs climb trees") is true or false.

On each trial of the Individual-Traits condition, participants were told about an individual (e.g., Zorb) and were then provided with evidence about the prevalence of a behavior in this individual's life (e.g., Zorb has climbed trees in 30% of the situations where he had the opportunity). The qualifier "where he had the opportunity" was added to avoid ambiguity; without this qualifier, the meaning of the percentage is ambiguous because it is unclear what is included in the denominator (e.g., are we counting situations where Zorb is sleeping? how about situations where there are no trees around?). Finally, participants were asked whether this evidence licenses an inference about an enduring disposition (e.g., "Zorb climbs trees"; see Table 1).

Because sentences such as "zorbs climb trees" and "Zorb climbs trees" are understood to express claims about the features of kinds and the dispositions of individuals, respectively (e.g., Cimpian et al., 2010; Krifka et al., 1995; Leslie, 2008), asking subjects about the truth value of these statements was an easy, intuitive means of assessing their endorsement of the corresponding conclusions/inferences.

We used four items: two that had strong danger connotations (chopping people's heads off and ripping out people's guts), and two that did not (climbing trees and digging holes in the ground). Each item was presented at six levels of evidence (1%, 5%, 10%, 30%, 50%, and 70%; 4 items × 6 levels of evidence = 24 trials). A different novel name was used on each trial (e.g., blins/Blin, ludinos/Ludino). Moreover, each name was assigned to a "dangerous" item for half of the subjects and to a "plain" item for the other half; this counterbalancing was designed to avoid confounding property content with particular sets of names. The order of the trials was randomized individually for each participant.

At the end of the task, we asked a question designed to check whether participants had paid attention: "Please think back to the task you just completed. Did the questions ask about single individuals or entire species?" Twenty-nine participants failed this attention check and were thus excluded from the sample. Participants then completed a demographics questionnaire and were debriefed about the goals of the study.

To summarize, the design of the study was as follows: 2 (inference type: Kind-Features vs. Individual-Traits; between subjects) × 2 (property content: dangerous vs. plain; within subject) × 6 (level of evidence: 1%, 5%, 10%, 30%, 50%, and 70%; within subject). The main dependent measure was the proportion of inferences endorsed by subjects within each cell of this design. These data were submitted to analyses of variance (ANOVAs), with Bonferroni-adjusted follow-up tests where needed.

Results and discussion

The main prediction of the SIM Hypothesis was that both kind-feature and individual-trait inferences would be endorsed more often when the evidence carries connotations of danger and threat. In the context of a 2 (inference) × 2 (property content) × 6 (evidence level) ANOVA, this broad prediction translates into two finer-grained predictions: First, we expect a significant main effect of property content (with more inferences for the dangerous items) and/or a significant interaction between property content and evidence level (with more inferences for the dangerous items at the lower evidence levels but perhaps not at the higher ones, where we may see ceiling effects). Second, we expect that neither of the above effects will be moderated by the type of inference assessed. This absence of significant interactions between (1) inference type (kind-feature vs. individual-trait) and content, and (2) inference type, content, and level of evidence would suggest that the facilitative effect of danger-related content is not different for inferences about the features of kinds and the traits of individuals, in line with the SIM Hypothesis.

The results of the ANOVA provided support for these predictions. First, we found an overall inferential advantage for the dangerous items, $F(1, 196) = 7.16$, $p = .008$, $\eta_p^2 = .035$, which was qualified by the expected interaction with evidence level, $F(5, 980) = 2.85$, $p = .015$, $\eta_p^2 = .014$. As illustrated in Figure 1, the inferential advantage for dangerous items was most apparent at the lower evidence levels (1% level: $p = .004$; 5% level: $p = .039$); in contrast, higher levels of evidence licensed almost uniform endorsement of the inferences, with no differences by property content.

Second, we explored whether the effect of property content was moderated by the type of inference participants were asked to make (kind-feature vs. individual-trait). As predicted, there was no evidence of such moderation: Neither the inference × property content interaction, F(1,

196) = 0.03, $p = .852$, nor the inference × property content × evidence level interaction, $F(5, 980) = 0.93$, $p = .460$, was significant. These results are consistent with the claim that property content has an analogous effect on kind-feature and individual-trait inferences (see also Figure 1).

Finally, the ANOVA revealed several other main effects and interactions. The main effect of evidence level, $F(5, 980) = 84.66$, $p < .001$, $\eta_p^2 = .302$, suggested that participants were more likely to endorse inferences for which they had more evidence (see Figure 1). Interestingly, we also found a main effect of inference type, $F(1, 196) = 39.24$, $p < .001$, $\eta_p^2 = .167$, and an interaction between inference type and evidence level, $F(5, 980) = 11.12$, $p < .001$, $\eta_p^2 = .054$. Participants endorsed more individual-trait than kind-feature inferences overall (for a similar result, see Lawson & Kalish, 2006), and particularly at the lower levels of evidence. We will explore several explanations for these differences (e.g., possible ambiguities in the stimulus statements, difficulties in truly equating the evidence for the two types of inferences) in Experiment 5 and the General Discussion.

In sum, the results of Experiment 1 suggest that both inferences about the features of kinds and inferences about the traits of individuals are similarly facilitated by information with strong negative valence. These findings are consistent with the claim that these two types of inferences are the outputs of a single underlying mechanism.

CHAPTER 3

EXPERIMENT 2: NO EFFECT OF VALENCE ON QUANTIFIED INFERENCES OVER INDIVIDUALS AND SITUAIONS

The effect of valence on kind-feature and individual-trait inferences (Experiment 1) suggests a common source for these inferences. However, the claim of a special shared origin for these inferences would be strengthened if we also found that the effect of valence was *not* broadly shared by other sorts of inferences. Providing such evidence was our goal in the present experiment. Specifically, we tested whether the valence effects observed in Experiment 1 would also be obtained for two other types of specific-to-general inferences that operate over individuals and situations: namely, existentially-quantified (some/sometimes) and universally-quantified (all/always) inferences. The prediction of the SIM Hypothesis is that danger connotations should have no effect on the likelihood of making these other, quantified, inferences. Rather, these inferences should be wholly dependent on the *amount* of evidence provided, not its content. For example, the only determinant of whether it is legitimate to conclude that *all* zorbs display a certain feature should be *how many* of them display this feature, not whether the feature is dangerous or plain.

Method

Participants. Participants (*N* = 221; 67 men, 109 women; demographic information for the remaining 45 subjects was not recorded) were recruited from two sources: a university undergraduate subject pool and Amazon's Mechanical Turk service. Participants received partial course credit or $0.75, respectively, for their participation. An additional 27 subjects were tested but excluded from the final sample either because their IP addresses were from outside the US (*n* = 6) or because they failed the attention check at the end of the study (*n* = 21).

Materials, procedure, and design. The method of this experiment was identical to that of Experiment 1, with one key exception: Participants were asked to make *quantified* inferences over individuals and situations. Specifically, participants were asked to judge statements expressing either existential generalizations (e.g., "*Some* zorbs climb trees," "Zorb *sometimes* climbs trees") or universal generalizations (e.g., "*All* zorbs climb trees," "Zorb *always* climbs trees").

Thus, the design of this study was as follows: 2 (quantification type: existential [some/sometimes] vs. universal [all/always]; between subjects) × 2 (inference type: over individuals [some/all] vs. over situations [sometimes/always]; between subjects) × 2 (property content: dangerous vs. plain; within subject) × 6 (level of evidence: 1%, 5%, 10%, 30%, 50%, and 70%; within subject).

Results and Discussion

The main prediction of the SIM Hypothesis was that the quantified inferences tested here should not be influenced by property content. Thus, we expect the main effect of property content and its interaction with evidence level—both of which were significant in Experiment 1—to be non-significant here. This prediction was confirmed: As shown in Figure 2, quantified inferences about the danger-related items were endorsed to the same extent as quantified inferences about the plain items both in the aggregate, $F(1, 217) = 0.18$, $p = .670$, and when the different levels of evidence are taken into account, $F(5, 1085) = 1.50$, $p = .187$.

The only significant effect involving property content was the four-way interaction between all factors in the ANOVA, $F(5, 1085) = 2.67$, $p = .021$, $\eta_p^2 = .012$. Importantly, however, an inspection of the relevant means and follow-up comparisons uncovered no evidence of an advantage for the dangerous items. The four-way interaction seemed to be driven by an unusual pattern of differences in participants' existentially-quantified inferences across situations (*X* sometimes does *Y*): At the 1% evidence level, participants made stronger inferences from the plain evidence than from the danger-related evidence (p = .005), whereas at the 10% evidence level they made stronger inferences from the danger-related evidence (p = .003). Analogous differences were not found for the existentially-quantified inferences over individuals (some *X*s do *Y*), or for either type of universally-quantified inferences; these asymmetries in turn led to the significant four-way interaction. While this interaction is not easy to interpret, it also provides little support for the claim that negatively-valenced information facilitates quantified inferences.

The other significant effects uncovered by the ANOVA were as follows: Existentially-quantified inferences were endorsed more often than universally-quantified ones, $F(1, 217)$ = 1353.17, $p < .001$, $\eta_p^2 = .862$, particularly at the lower prevalence levels, $F(5, 1085) = 12.84$, $p < .001$, $\eta_p^2 = .056$. And, as expected, the likelihood of endorsing these two types of quantified inferences was higher at the higher levels of evidence, $F(5, 1085) = 17.69$, $p < .001$, $\eta_p^2 = .075$.

Together, these results suggest that quantified (existential and universal) inferences are not facilitated by information that has connotations of threat and danger. As a result, the fact that both inferences about the features of kinds and inferences about the traits of individuals *are* facilitated by this information makes it all the more plausible that they stem from the operation of the same separate cognitive mechanism.

CHAPTER 4

EXPERIMENT 3: THE EFFECT OF DISTINCTIVENESS ON KIND-FEATURE AND INDIVIDUAL-TRAIT INFERENCES

In this experiment, we explored the influence of distinctiveness on inferences about the features of kinds and the traits of individuals. The SIM Hypothesis predicts that both types of inferences would be facilitated in a similar manner when the properties/behaviors that constitute the evidence are distinctive or unique.

Method

Participants. Participants (*N* = 194; 69 men, 123 women; 1 participant did not report gender and another reported “male and female”) were recruited from two sources: a university undergraduate subject pool and Amazon’s Mechanical Turk service. An additional 16 subjects were tested but excluded from the final sample either because their IP addresses were from outside the US (*n* = 2) or because they failed the attention check (*n* = 14).

Materials, procedure, and design. The method of this experiment was identical to that of Experiment 1, except for the items. In this experiment, we manipulated the distinctiveness of the items by including additional information that suggested the relevant properties were either distinctive or commonplace (see Table 3 for an example). Each of the four items used in this study (climbing trees, digging holes in the ground, making nests out of grass, and washing oneself with water) was said to be distinctive for approximately half of the participants and commonplace/plain for the other half. Thus, across subjects, the actual properties about which participants made their judgments were equated between the distinctive and plain items.

In summary, the design of this study was as follows: 2 (inference type: Kind-Features vs. Individual-Traits; between subjects) × 2 (property content: distinctive vs. plain; within subject) ×

6 (level of evidence: 1%, 5%, 10%, 30%, 50%, and 70%; within subject).

Results and Discussion

As before, the SIM Hypothesis makes two main predictions. First, the distinctive items should license more inferences than the plain items, especially at the lower levels of evidence. Second, this facilitative effect should not be moderated by whether participants are asked to make inferences about the features of kinds or the traits of individuals.

In line with the first prediction, the ANOVA uncovered a significant interaction between property content and level of evidence, $F(5, 960) = 2.57$, $p = .025$, $\eta_p^2 = .013$. As illustrated in Figure 3, the effect of distinctiveness was present when the evidence was weak (5% level: $p = .089$; 10% level: $p = .023$) and absent when the evidence was more abundant. The overall main effect of property content did not reach significance in this study, $F(1, 192) = 2.10$, $p = .149$, $\eta_p^2 = .011$.[5]

In line with the second prediction, there was no evidence that the type of inference participants were making moderated the facilitative effect of distinctiveness. Neither the two-way interaction between inference type and property content, $F(1, 192) = 0.23$, $p = .629$, nor the three-way interaction between these two factors and evidence level, $F(5, 960) = 0.32$, $p = .904$, was significant.

The remaining effects uncovered by the ANOVA were analogous to those found in Experiment 1. First, there was a significant main effect of evidence level, $F(5, 960) = 112.82$, $p < .001$, $\eta_p^2 = .370$, indicating that participants were more likely to make inferences when more

[5] Overall, these effects are smaller in magnitude than those in Cimpian et al.'s (2010) experiments, which investigated the impact of danger and distinctiveness on inferences about kinds. There are multiple methodological differences that could explain the differences in the magnitude of the effects (e.g., the content of the stimuli, the testing circumstances [in-lab vs. online]). To illustrate, the stimulus sentences used here do not include the words "dangerous" or "distinctive," whereas Cimpian et al.'s did (e.g., "Lorches have purple feathers" [plain] vs. "Lorches have *distinctive* purple feathers" [distinctive; emphasis added]). It is possible that subtle differences of this sort have a non-trivial impact on participants' judgments.

evidence was provided. Second, inferences about the traits of individuals were overall more frequent than inferences about the features of kinds, $F(1, 192) = 19.36$, $p < .001$, $\eta_p^2 = .092$, but particularly at the lower levels of evidence, $F(5, 960) = 6.63$, $p < .001$, $\eta_p^2 = .033$ (see also Figure 3). Additional discussion of this difference will be provided in Experiment 5 and the General Discussion.

In sum, the results of the present study suggest that both kind-feature and individual-trait inferences are made more easily when the properties under consideration are unique or distinctive. These results provide converging support for the proposal that inferences about the features of kinds and the dispositional traits of individuals are generated by the same cognitive process.

CHAPTER 5

EXPERIMENT 4: NO EFFECT OF DISTINCTIVENESS ON QUANTIFIED INFERENCES OVER INDIVIDUALS AND SITUATIONS

Is the facilitative effect of distinctiveness specific to kind-feature and individual-trait inferences? If so, that would reinforce the argument that these inferences are both the outputs of a single, separate cognitive mechanism. Thus, in the present experiment we tested the prediction that the distinctiveness of a property has no effect on the likelihood of drawing quantified (existential and universal) inferences.

Method

Participants. Participants ($N = 213$; 88 men, 118 women, 1 not reporting gender; demographic information for the remaining 6 subjects was not recorded due to a programming error) were recruited from two sources: a university undergraduate subject pool and Amazon's Mechanical Turk service. An additional 20 subjects were excluded because they failed the attention check ($n = 20$).

Materials, procedure, and design. The method of this study was identical to that of Experiment 2 (which also looked at quantified inferences), except that the items were taken from Experiment 3. Thus, this experiment had the following design: 2 (quantification type: existential [some/sometimes] vs. universal [all/always]; between subjects) × 2 (inference type: over individuals [some/all] vs. over situations [sometimes/always]; between subjects) × 2 (property content: distinctive vs. plain; within subject) × 6 (level of evidence: 1%, 5%, 10%, 30%, 50%, and 70%; within subject).

Results and Discussion

Consistent with the SIM Hypothesis, the ANOVA revealed no evidence that property content had any influence on participants' quantified inferences (see Figure 4). In fact, none of the interactions involving property content even approached significance, nor did its main effect, Fs < 0.54, ps > .465.

The only significant results were the more obvious ones involving level of evidence and quantification type: Participants made more existential than universal inferences, $F(1, 209) = 1523.69$, $p < .001$, $\eta_p^2 = .879$, particularly at the lower prevalence levels, $F(5, 1045) = 12.60$, $p < .001$, $\eta_p^2 = .057$. They also made more inferences from stronger evidence, $F(5, 1045) = 29.24$, $p < .001$, $\eta_p^2 = .123$. An unexpected additional result was that participants made significantly more existentially-quantified inferences over individuals (e.g., "Some zorbs climb trees") than over situations (e.g., "Zorb sometimes climbs trees") at the lowest two evidence levels (1% and 5%), ps < .038. Because similar differences were absent at the higher levels of evidence, as well as for the universally-quantified inferences, the ANOVA revealed a significant three-way interaction between quantification type, inference type, and evidence level, $F(5, 1045) = 2.84$, $p = .015$, $\eta_p^2 = .013$.[6] For our purposes here, however, the important thing to keep in mind is that the content of the properties was not involved in any of these effects.

The results of this experiment suggest that existentially- and universally-quantified inferences are not influenced by the distinctiveness of the properties being considered. Together with the results of Experiment 2 (which suggested that danger/threat connotations are irrelevant to these inferences as well), these findings provide support for the conclusion that sensitivity to content may be specific to kind-feature and individual-trait inferences. In turn, this conclusion reinforces the Shared Inferential Mechanism Hypothesis.

[6] This three-way interaction subsumed two lower-order interactions: that between generalization type (over individuals vs. over situations) and level of evidence, $F(5, 1045) = 4.82$, $p < .001$, $\eta_p^2 = .023$, and that between generalization type and quantification type, $F(1, 209) = 5.27$, $p = .023$, $\eta_p^2 = .025$.

CHAPTER 6

EXPERIMENT 5: WHY WERE INDIVIDUAL-TRAIT INFERENCES ENDORSED MORE OFTEN THAN KIND-FEATURE INFERENCES?

We are proposing that inferences about the features of kinds and the traits of individuals have a common cognitive source. Although the evidence so far is consistent with this possibility, we also found that people seem overall more willing to make inferences about the traits of individuals than about the features of kinds, even when given what appears to be the same evidence (see Experiments 1 and 3). Does this difference speak against the SIM Hypothesis? In the present experiment, we explored a plausible reason to think it does not—that is, we explored an explanation for this difference that is compatible with the existence of a single underlying mechanism. This explanation was prompted by some of our participants' comments in an open-ended question during the debriefing portion of the studies. When asked to describe their approach to the task, a number of participants in the Individual-Traits conditions of Experiments 1 and 3 provided answers suggesting that they interpreted the stimulus sentences (e.g., "Zorb climbs trees") as expressing not a *dispositional inference* (roughly, Zorb is generally disposed to climb trees), as we had intended, but rather an *ability* (roughly, Zorb is able to climb trees). For instance, one person said, "Even if the animal only did something in 1 percent of presented opportunities, that animal *can do that behavior*"; another stated, "I answered all the questions *true* because no matter the percentage the animal mentioned *could perform the act stated*" (our italics).

If the ability interpretation of our stimulus sentences was common, it could explain why participants were endorsing these statements so freely: Intuitively, endorsement of ability statements of this sort requires minimal evidence. Even if John has run a four-minute mile only once, for example, it still logically follows that he *is capable of* running a four-minute mile. This is a deductive inference—no inductive leap is involved, and no further evidence is needed. Importantly, this problem does not extend to the stimulus statements we used to assess participants' inferences about the features of kinds (e.g., "zorbs climb trees"). In this case, even if the evidence (e.g., 1% of zorbs climb trees) was interpreted as signaling an ability (e.g., these 1% can climb trees), a participant would still have to perform an inductive inference in order to conclude that the kind as a whole possesses this ability. The trivial, deductive option is not available here, which suggests that endorsement of the conclusions in the Kind-Features condition is unlikely to have been artificially inflated.

Our goal in Experiment 5 was to investigate these issues more formally. Specifically, we assessed (1) the extent to which people adopt the ability interpretation of sentences such as those used in our Individual-Traits conditions, (2) the extent to which adopting this interpretation leads to higher endorsement of the truth of these statements, and (3) whether taking these ability interpretations into account would be sufficient to explain the endorsement differences between kind-feature and individual-trait statements (such as those observed in Experiments 1 and 3). If the ability interpretation is found to be common among participants, to promote high levels of endorsement, and to explain away the apparent differences between the individual-trait and kind-feature inferences, then we would have good reason to attribute the unexpected differences found in previous studies to ambiguities in our stimulus sentences.

Method

Participants. Participants (N = 132; 63 men, 69 women) were recruited from the Amazon's Mechanical Turk service. An additional 16 subjects were excluded either because their IP addresses were outside the US (n = 3) or because they failed the attention check (n = 13).

Materials, procedure, and design. The method of this study was similar to that of Experiment 3, including a Kind-Features and an Individual-Traits condition, each with 24 trials. There were, however, two major differences from the method of Experiment 3. First, we no longer manipulated distinctiveness—that is, we removed mention of whether the properties were displayed by few versus many other animals. The distinctiveness manipulation was omitted because in this study we are simply interested in people's interpretation of the conclusion statements (e.g., "Zorb climbs trees").

Second, we assessed the appeal of the ability interpretation for the statements in the Individual-Traits condition (where such an interpretation would be problematic). Specifically, we asked participants in the Individual-Traits condition to evaluate several paraphrases of the stimulus sentences. Participants were told, for example, "Imagine you need to paraphrase the following statement: *Kazz climbs trees*. How good are the following paraphrases?" The paraphrase of most interest here was the one that described an ability (e.g., "Kazz is able to climb trees"); the other two paraphrases presented to subjects will be described below. Participants marked their evaluations of the paraphrases on a scale from 1 ("not good at all") to 9 ("very good"). These paraphrase questions were asked at the very end of the task, after participants completed the 24 trials of the main task. To minimize interference from these prior judgments, the stimulus sentences were paired with novel names that had not occurred in the main task (e.g., Kazz instead of Zorb). For comparison purposes, we also elicited evaluations of two other paraphrases, both of which conveyed quantified inferences: a broad-scope paraphrase (e.g., "Kazz generally climbs trees") and a narrow-scope paraphrase (e.g., "Kazz sometimes climbs trees"). The meaning of these quantified paraphrases provides only a rough approximation of a dispositional generalization (e.g., Krifka et al., 1995), which—as explained in the introduction—

is much less tied to specific frequencies than quantified generalizations are. Nevertheless, these quantified paraphrases can provide a useful point of comparison for the ability paraphrases. The presentation order of the three paraphrases was randomized individually for each participant.

Results and discussion

The first goal of this study was to determine the extent to which participants interpret the stimulus statements we used to probe individual-trait inferences as expressing an *ability* to perform a certain behavior. Participants' evaluation of the ability paraphrases suggested that these paraphrases were indeed deemed quite good, $M = 7.39$ on a 1 ("not good at all") to 9 ("very good") scale ($SD = 1.90$). This rating was significantly above the midpoint of the 1–9 scale, $t(64) = 10.17$, $p < .001$. For a comparison with the "generally" and "sometimes" paraphrases, participants' ratings were entered into a one-way repeated-measures ANOVA. This analysis revealed that participants differentiated between the three paraphrases, $F(2, 128) = 21.36$, $p < .001$, $\eta_p^2 = .250$, with the ability paraphrase being judged more acceptable than both the "generally" ($M = 5.12$, $SD = 2.33$) and the "sometimes" ($M = 6.00$, $SD = 2.08$) paraphrases, *p*s < .001. Together, these results suggest that the ability interpretation of Individual-Traits sentences in Experiments 1 and 3 (e.g., "Zorb climbs trees") may have been relatively prevalent and may thus have been a strong competitor for the dispositional inference interpretation, which is the one we had intended for these sentences.

The second goal of this study was to test the link between ability interpretations of the stimulus sentences and higher endorsement of these sentences in the main task: Were the participants who favored the ability interpretation of the stimulus sentences also more likely to endorse these statements as true no matter what the evidence? Such a correlation seems likely, especially given that that the ability interpretation is only a small step removed from even the lowest level of evidence. Indeed, there was a significant positive correlation between participants' rating of the ability paraphrases and their average endorsement of the stimulus statements in the main task, $r(63) = .30$, $p = .017$. This relationship was present even when we partialled out participants' ratings of the "generally" and "sometimes" paraphrases, $rs(62) = .29$ and .32, $ps = .022$ and .010, respectively. These results suggest that, as predicted, adopting an ability interpretation of the stimulus sentences may make it easier to accept these statements on the basis of the evidence provided in the main task.

The third goal of this study was to assess whether the differences observed in Experiments 1 and 3 between kind-feature and individual-trait responses might be explained by the availability of ability interpretations. If these interpretations are sufficient to account for the observed differences, then the Individual-Traits participants who did *not* favor the ability interpretations should behave similarly to the Kind-Features participants. To investigate this prediction, we first divided our Individual-Traits condition into two subsamples based on whether participants gave the ability paraphrase the highest rating among the three paraphrases ($n = 47$) or not ($n = 18$). We then submitted these data to a 3 (group: Individual-Traits [ability interpretation favored] vs. Individual-Traits [ability interpretation disfavored] vs. Kind-Features) × 6 (level of evidence: 1%, 5%, 10%, 30%, 50%, and 70%) ANOVA.

Consistent with our prediction, the results of this ANOVA suggested that differences between the Kind-Features and Individual-Traits conditions are reduced to non-significance when the confounding effect of ability interpretations is taken into account. To illustrate, pairwise comparisons that followed up on the main effect of group, $F(2, 129) = 11.27, p < .001$, $\eta_p^2 = .149$, revealed no significant difference between the Kind-Features participants ($M = .54$ *true* responses, $SD = .37$) and the Individual-Traits participants who *did not* favor the ability interpretation (*M* = .68, *SD* = .25), $p = .365$. However, as predicted, the Kind-Features participants were significantly less likely to endorse the stimulus statements than the Individual-Traits participants who *did* favor the ability interpretation (*M* = .83, *SD* = .26), $p < .001$.

To summarize, the results of this experiment established three main conclusions. First, the stimulus sentences used to probe inferences about the dispositions of individuals (e.g., "Zorb climbs trees") also admit a relatively salient ability interpretation (e.g., Zorb is able to climb trees). Second, this interpretation leads to agreement with these sentences based on a trivial amount of evidence, without the need for any sort of inductive leap. Third, the availability of these ability interpretations seems to be a key reason for the differences observed in Experiments 1 and 3 between the overall acceptability of kind-feature and individual-trait inferences. In the present experiment, the difference between these two conditions was no longer significant when considering just the participants who disfavored the ability interpretation for the Individual-Traits stimulus sentences. However, some caution may be warranted with respect to this last result: Although the condition difference was not significant, the averages still point to lower endorsement of the stimulus sentences in the Kind-Features condition (.54, as opposed to .68 for the Individual-Traits participants who disfavored the ability interpretations). In addition, this comparison may have been underpowered, as there were only 18 participants in our sample who did not prefer the ability interpretation. Nevertheless, the totality of the evidence from this study suggests that stimulus ambiguity is one possible explanation for the differences previously observed between inferences about the features of kinds and the traits of individuals. By helping to explain this unexpected difference, this evidence also reinforces the proposal that these inferences may be the products of the same cognitive mechanism.

CHAPTER 7

GENERAL DISCUSSION

Human learning is powerful in large part because we are able to formulate broad inferences based on limited first-hand evidence. These inferences are incredibly diverse—they take a variety of forms and serve a variety of uses, from learning about the phonology of one's language (e.g., Kuhl, 2004) to developing a general "working model" of attachment relationships (e.g., Bowlby, 1982; Johnson, Chen, & Dweck, 2007) to learning about the edible resources in one's environment (e.g., Wertz & Wynn, 2014). This diversity of forms and functions is likely to be accompanied by a diversity of underlying neurocognitive mechanisms. However, functional differences among inferences may not always correspond to differences in underlying mechanisms. In fact, in this paper we propose that two types of inferences that are different both in their form and in their use—and that might at first appear to be the products of different cognitive systems—may actually share an origin. The first type consists of inductive inferences about the features of categories, which are a crucial component of learning about the natural world (e.g., Keil, 1989; Markman, 1989; Murphy, 2002; Smith & Medin, 1981). The second type consists of inductive inferences about the dispositions of the individuals with whom we interact, which are a crucial component of navigating our complex social world (e.g., Heider, 1958; Gilbert, 1998; Jones, 1979; Ross, 1977). Despite the multiple differences between these two types of inferences (e.g., in their function, their scope, and their prior treatment in the psychological literature), we propose that they are in fact computed by the same cognitive process.

Of course, demonstrating beyond doubt that two judgments have the same cognitive origin is not a trivial task—just as it is not trivial to determine whether, say, two anonymous paintings were made by the same artist. In such circumstances, one reasonable strategy is to look for certain "signatures" associated with the suspected originator. If these trademark features are present in both instances, then one might be able to claim a shared origin. One's confidence in this claim may be increased further if the relevant signatures were also unusual. For example, it would not help much if the presumed artist's signature brushstroke was in fact shared by most painters of that period. This is, in broad outline, the strategy we adopted here with respect to inferences about the features of kinds and the traits of individuals. Prior research has revealed two signatures of kind-feature inferences: namely, that they are facilitated by (1) danger-relevant and (2) distinctive properties (Cimpian et al., 2010; Leslie, 2008). Here, we tested whether individual-trait inferences exhibit these signatures as well. To provide further evidence for the claim of a common source or mechanism, we also sought to show that these signatures are not shared by other types of inferences.

Summary of Results

The results of our five experiments supported the SIM Hypothesis. Experiments 1 and 3 suggested that inferences about the dispositional traits of individuals are indeed facilitated by danger-related and distinctive behaviors, just as inferences about the features of kinds are. Of note, the procedure for testing these two types of inferences was—in all relevant respects—identical (see Tables 1 and 3); this feature of the design enables us to draw strong conclusions from these comparisons. Moreover, Experiments 2 and 4 established that the two signatures of kind-feature and individual-trait inferences are unusual, at least insofar as they are not shared by existentially-quantified (some/sometimes) and universally-quantified (all/always) inferences over situations and individuals. This result strengthens the claim that kind-feature and individual-trait inferences have a common origin, just as one might be more confident that two paintings are the work of the same artist if they shared a highly unusual brushstroke pattern.

Finally, Experiment 5 followed up on an unexpected difference between kind-feature and individual-trait inferences, wherein the latter were more frequently endorsed than the former. In Experiment 5, we suggested that this overall difference may have been an artifact of the stimulus sentences used to probe dispositional inferences. These sentences admitted another salient interpretation that did not require participants to go beyond the evidence. Specifically, these sentences could have been interpreted as conveying that, if an individual performs a behavior, then that individual has the ability to perform that behavior. Because this ability interpretation made it acceptable to endorse the stimulus sentences on the basis of minimal evidence, it also gave rise to an apparent advantage for the individual-trait inferences. However, when we adjusted for this ability interpretation in Experiment 5, this unexpected advantage became smaller and fell to non-significance.

There are other potential reasons to doubt that the observed endorsement advantage for individual-trait over kind-feature inferences contradicts the SIM Hypothesis. For instance, one reason is that, although we were careful to match the percentages of individuals and situations described in the evidence (e.g., 5% of zorbs vs. 5% of situations involving Zorb), this is only a rough means of equating the actual strength of evidence. Recall that the individuals in the Individual-Traits condition were said to perform the relevant behaviors in a certain percentage of the situations where they had the opportunity to do so. Even within a relatively short time span, however, any individual has countless opportunities to perform any behavior imaginable. The vast majority of these behaviors are never actually performed; as a result, the base rate for any one of them is correspondingly low. In this context, even the lowest levels of evidence in our task may have constituted relatively strong evidence. For example, if an individual took advantage of even 1% of the opportunities for a certain behavior (e.g., digging holes), that would still translate into quite a few instances of that behavior—perhaps sufficient to license a dispositional inference. The same may not be true when the evidence concerns the members of a kind. In this case, the base rates are more uncertain: Any particular feature could in principle be widespread among the members of any particular kind. Thus, low levels of evidence (e.g., 1%) seem less impressive than in the individual case. To summarize, potential differences in assumed base rates may have led to differences in the perceived strength of the evidence for individual-trait and kind-feature inferences. In turn, it is possible that the differences we observed in people's willingness to make kind-feature and individual-trait inferences are not due to differences in the underlying computational mechanisms but rather to differences in the strength of the evidence we provided as input to a single, shared mechanism.

Overall, the present experiments provide promising evidence for the SIM Hypothesis. In the long run, however, it will be important to augment this evidence by testing other predictions of this hypothesis. For example, future studies could explore other possible characteristics of the shared inferential mechanism hypothesized here, beyond its sensitivity to threatening and distinctive stimuli. As we speculate in the next section, this mechanism may also be influenced by the *explanations* generated for the evidence at hand and may emerge *early in development*. If these are indeed characteristics of the shared inferential mechanism, then inferences about the features of kinds and the traits of individuals should exhibit them to the same extent. Further investigation of whether the mechanism responsible for these inferences is separate from the mechanisms responsible for other inductive inferences, including quantified inferences beyond those explored here, would also be worthwhile.

How Does the Shared Inferential Mechanism Work?

We have argued that people make inferences about the traits of individuals and the features of categories via the same cognitive mechanism. What are the characteristics of this mechanism? While a full answer to this question is beyond the scope of this paper, we nevertheless provide a few speculative suggestions. The first two characteristics discussed below map directly onto the present experiments, while the last two go beyond them.

First, it is likely that this mechanism privileges *survival-relevant* information, such that evidence pertaining to potential threats in the environment licenses stronger inferences about the relevant targets (individuals or kinds; see Experiment 1). These inferences have adaptive value insofar as they steer the organism away from future circumstances in which the relevant threats would reemerge (that is, away from future encounters with the individual or the members of the kind inferred to be a threat). Threat-relevant information is privileged in other cognitive–perceptual systems as well, presumably because of its adaptive value (for reviews, see Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001; Vaish, Grossmann, & Woodward, 2008). For example, even young children show faster orienting responses and enhanced memory for threatening stimuli (e.g., Baltazar, Shutts, & Kinzler, 2012; Hamlin, Wynn, & Bloom, 2010; Kinzler & Shutts, 2008; LoBue & DeLoache, 2010).

Second, the shared mechanism we hypothesize here is likely to attend to the *diagnosticity*, or informational value, of the evidence (see Experiment 3). A feature is diagnostic of a kind to the extent that it distinguishes that kind from others; analogously, a behavior is diagnostic of an individual to the extent that it distinguishes that individual from others. The role of diagnosticity has been independently recognized in the literatures on category and trait inferences, although it has not always been termed as such. The research on categories, for example, has identified distinctiveness as an important reason why people formulate and endorse generic inferences such as *lions have manes* or *cardinals are red* (e.g., Cimpian et al., 2010; Leslie, 2007, 2008). Although the features involved in these inferences are present only in a minority of kind members (e.g., only adult male lions have manes), they are nevertheless distinctive of these kinds (i.e., members of few other kinds have them). In turn, the distinctiveness (or diagnosticity) of these features boosts the acceptability of corresponding kind inferences relative to potential competitors (e.g., that cardinals are brown). The role of diagnosticity has been recognized in the literature on trait inferences as well. For example, in their groundbreaking treatise Jones and Davis (1965) argued it was a central determinant of whether a trait inference (or, in their terminology, a correspondent inference) is formulated: The likelihood of making a correspondent inference depends on the informational value provided by an action, which is trivially low if this action is unremarkable—that is, if this action is what anyone else would have done in the same context. For instance, to the extent that greeting one's coworkers with a smile is simply what everyone does, then this behavior cannot serve as a basis for stronger inferences about a person's general traits (e.g., friendliness). This argument about the role of diagnosticity or distinctiveness influenced subsequent work on this topic, much of which takes up the same conclusion in various forms (e.g., Ajzen & Fishbein, 1975; Kelley, 1967, 1973; Reeder & Brewer, 1979; Skowronski & Carlston, 1987).[7] In sum, it seems likely that the cognitive mechanism responsible for inferences about the features of kinds and the traits of individuals is sensitive to the diagnosticity of the evidence available.[8]

---

[7] Diagnosticity might explain the occasional positivity bias documented in inferences about the traits of individuals (as mentioned in the introduction). In certain domains (such as achievement), positive behaviors may be particularly diagnostic. Even the best artist, for example, does not create a masterpiece with every single work. Thus, in such a domain, a positive outcome (e.g., a masterpiece) may be more diagnostic than a negative one (e.g., a flop), and therefore more likely to lead to a trait inference (e.g., Skowronski & Carlston, 1987). The same result would probably hold with respect to inferences about the features of kinds, although this prediction has not yet been tested (as far as we know).

[8] Many threatening features/behaviors are also distinctive. Might it be sufficient to posit that the hypothesized inferential mechanism is sensitive to diagnosticity, with threat being just a special case of diagnosticity? We suspect

Third, inferences about the features of kinds and the traits of individuals may be intimately linked with people's *explanations* for the evidence at hand. A fundamental aspect of human psychology—acknowledged across various subdisciplines of psychological science—is the drive to make sense of the world via explanations (e.g., Cimpian & Salomon, in press; Gopnik, 1998; Keil, 2006; Lombrozo, 2006; Murphy & Medin, 1985; Ross, 1977; Weiner, 1985). The explanation generated for a certain observation imposes strong constraints on the breadth of the inferences that can be drawn from it. For example, if a feature of an unfamiliar animal (e.g., patches of hairless skin) is explained as an accident (e.g., the animal was in a fight), then it is unlikely that this feature will be inferred to characterize the animal's kind (e.g., Cimpian & Markman, 2008; Gelman, 1988). Similarly, if an instance of behavior (e.g., a nervous laughter) is explained in terms of the particular circumstances in which the behavior occurred (e.g., a first date), then broader inferences about personal dispositions are unlikely. As with diagnosticity, the key role of explanation has been independently recognized by cognitive and social psychologists. A currently dominant view of categorization has in fact been termed the *theory-based* view so as to highlight the crucial role it assigns to causal-explanatory understandings in people's reasoning about categories (e.g., Carey, 1985; Gelman, 2003; Murphy & Medin, 1985). Similarly, social psychologists have conceptualized the process of reasoning about individuals' traits as an instance of *causal attribution* since their first systematic explorations of this subject matter (e.g., Heider, 1958; Jones & Davis, 1965). In sum, it seems plausible to claim that the mechanism by which inferences about the features of kinds and the traits of individuals are formulated is influenced by the explanations people generate to make

---

that diagnosticity alone is not sufficient: The evidence in the broader literature suggests that the effects of negative or threatening stimuli emerge even when diagnosticity is taken into account (see Baumeister et al., 2001; Hitchcock & Knobe, 2009). Extrapolating from this evidence, we believe that these two dimensions are likely to exert (at least partly) independent effects on people's kind-feature and individual-trait inferences.

sense of the world.

Finally, borrowing an idea from Leslie (2007, 2008) and Gelman (2010; Hollander, Gelman, & Star, 2002), we might also speculate that this shared mechanism is the *default* means available to our species for drawing general conclusions from limited evidence. By *default*, we mean two things here. First, this mechanism may be developmentally primitive—the first-emerging and most basic means of going beyond the evidence to derive expectations about new exemplars or situations. Consistent with this possibility, inferences about both the features of kinds and the traits of individuals seem to be available to infants even before their first birthdays (e.g., Baldwin, Markman, & Melartin, 1993; Graham, Kilbreath, & Welder, 2004; Hamlin et al., 2007; Luo & Baillargeon, 2005; Song, Baillargeon, & Fisher, 2005; Woodward, 1998, 1999). There is also evidence that the ability to reason with these inferences *precedes* the ability to reason with quantified (e.g., existential, universal) inferences, which further strengthens claims of developmental primacy (Hollander et al., 2002; Mannheim, Gelman, Escalante, Huayhua, & Puma, 2011; Tardif, Gelman, Fu, & Zhu, 2012). It should be noted, however, that the latter studies pertain exclusively to inferences/generalizations over individuals (to either kinds or quantified sets). In future work, it will be important to test whether this conclusion also holds with respect to inductive inferences over situations.

The shared inferential mechanism we hypothesize here may be default in a second sense as well: It may be the mechanism that people call upon spontaneously and without conscious awareness in order to routinely extract general conclusions from particular samples of evidence. The idea that dispositional inferences are computed ubiquitously, spontaneously, and outside of conscious awareness is commonly accepted in social psychology (for a review, see Uleman, Saribay, & Gonzalez, 2008). Recent evidence suggests this conclusion extends to inferences about the features of kinds as well. Sutherland, Cimpian, Leslie, and Gelman (under review) found that participants made kind inferences in a task in which such inferences detracted from correct performance (which speaks to their ubiquity and spontaneity), and even when participants were under a cognitive load. In fact, kind-feature and individual-trait inferences may be default in this second sense of "ubiquitously running in the background" in part because they require few cognitive resources to be computed (e.g., Hampton, 2012; Leslie & Gelman, 2012; Leslie, Khemlani, & Glucksberg, 2011; Todorov & Uleman, 2003).

To summarize, we have argued that inferences about the features of kinds and the traits of individuals are computed via a shared mechanism that (1) is particularly sensitive to threat-relevant information, (2) takes into account the diagnosticity of the evidence available, (3) is bound up with our intuitive explanations for how the evidence came about, and (4) may be the default generalization mechanism that humans possess.

Conclusion

The present experiments provide support for the SIM Hypothesis, according to which inferences about the features of kinds and the traits of individuals are computed via the same cognitive mechanism. Looking to the future, our argument opens up opportunities for cross-fertilization between two literatures that have long and rich traditions but that have seldom been in contact. If inferences about the features of kinds and the traits of individuals truly stem from the same source, as we have argued, then combining the insights accumulated independently by cognitive and social psychologists could be used to break tremendous new ground in understanding this important aspect of human reasoning. We hope that the present work will embolden researchers in both areas to investigate kind-feature and individual-trait inferences jointly, in closely matched experimental designs, further exploring the many predictions that follow from the Shared Inferential Mechanism Hypothesis.

REFERENCES

Abelson, R. P., & Kanouse, D. E. (1966). Subjective acceptance of verbal generalizations. In S. Feldman (Ed.), *Cognitive consistency: Motivational antecedents and behavioral consequents* (pp. 171–197). New York: Academic Press.

Atran, S. (1998). Folk biology and the anthropology of science: Cognitive universals and cultural particulars. *Behavioral and Brain Sciences, 21*, 547-609.

Baltazar, N. C., Shutts, K., & Kinzler, K. D. (2012). Children show heightened memory for threatening social actions. *Journal of Experimental Social Psychology, 112*, 102-110.

Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: MIT Press.

Ajzen, I., & Fishbein, M. (1975). A Bayesian analysis of attribution processes. *Psychological Bulletin, 82*(2), 261-277.

Aloise, P. A. (1993). Trait confirmation and disconfirmation: The development of attribution biases. *Journal of Experimental Child Psychology*, *55*(2), 177-193.

Baldwin, D. A., Markman, E. M., & Melartin, R. L. (1993). Infants' ability to draw inferences about nonobvious object properties: Evidence from exploratory play. *Child Development*, *64*(3), 711-728.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology, 5*(4), 323-370.

Bowlby, J. (1982). *Attachment and loss: Vol. 1. Attachment*. New York: Basic Books.

Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: MIT Press.

Carlson, G. N. (1977). *Reference to kinds in English*. Doctoral dissertation, University of Massachusetts, Amherst.

Chambers, C. G., Graham, S. A., & Turner, J. N. (2008). When hearsay trumps evidence: How generic language guides preschoolers' inferences about unfamiliar things. *Language and Cognitive Processes*, *23*(5), 749-766.

Cimpian, A., Brandone, A. C., & Gelman, S. A. (2010). Generic statements require little evidence for acceptance but have powerful implications. *Cognitive Science, 34*(8), 1452-1482.

Cimpian, A., Gelman, S. A., & Brandone, A. C. (2010). Theory-based considerations influence the interpretation of generic sentences. *Language and Cognitive Processes*, *25*(2), 261-276.

Cimpian, A., & Markman, E. M. (2008). Preschool children's use of cues to generic meaning. *Cognition*, *107*(1), 19-53.

Cimpian, A., & Salomon, E. (in press). The inherence heuristic: An intuitive means of making sense of the world, and a potential precursor to psychological essentialism. *Behavioral and Brain Sciences*.

Cova, F., & Egré, P. (under review). *Moral asymmetries and the semantics of* many.

Dahl, Ö. (1975). On generics. In E. L. Keenan (Ed.), *Formal semantics of natural language: Papers from a colloquium sponsored by the King's College Research Centre, Cambridge* (pp. 99-111). Cambridge, England: Cambridge University Press.

Dewar, K., & Xu, F. (2009). Do early nouns refer to kinds or distinct shapes? Evidence from 10-month-old infants. *Psychological Science*, *20*(2), 252-257.

Garlick, R. (1993). Trait relationships in situations of behavioral inconsistency. *Journal of Social Behavior and Personality*, *8*(2), 203-220.

Gawronski, B. (2004). Theory-based bias correction in dispositional inference: The fundamental attribution error is dead, long live the correspondence bias. *European Review of Social Psychology,* 15, 183-217.

Gelman, S. A. (1988). The development of induction within natural kind and artifact categories. *Cognitive Psychology, 20*, 65-95.

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

Gelman, S. A. (2004). Learning words for kinds: Generic noun phrases in acquisition. In D. G. Hall & S. R. Waxman (Eds.), *Weaving a lexicon* (pp. 445-484). Cambridge, MA: MIT Press.

Gelman, S. A. (2010). Generics as a window onto young children's concepts. In F. J. Pelletier (Ed.), *Kinds, things, and stuff: Mass terms and generics* (pp. 100-120). New York: Oxford University Press.

Gidron, D., Koehler, D. J., & Tversky, A. (1993). Implicit quantification of personality traits. *Personality and Social Psychology Bulletin*, *19*, 594-604.

Gilbert, D. T. (1998). Ordinary personology. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The Handbook of Social Psychology, Vol. 2* (pp. 89-150). New York: McGraw Hill.

Gilbert, D. T., & Malone, P. S. (1995). The correspondence bias. *Psychological Bulletin*, *117*(1), 21.

Gopnik, A. (1998) Explanation as orgasm. *Minds and Machines* 8:101–18.

Graham, S. A., Kilbreath, C. S., & Welder, A. N. (2004). Thirteen-month-olds rely on shared labels and shape similarity for inductive inferences. *Child Development*, *75*(2), 409-427.

Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature*, *450*(7169), 557-559.

Hamlin, J. K., Wynn, K., & Bloom, P. (2010). Three-month-olds show a negativity bias in their social evaluations. *Developmental Science, 13*(6), 923-929.

Hampton, J. A. (2012). Thinking intuitively: The rich (and at times illogical) world of concepts. *Current Directions in Psychological Science, 21*(6), 398-402.

Hayden, T., & Mischel, W. (1976). Maintaining trait consistency in the resolution of behavioral inconsistency: The wolf in sheep's clothing? *Journal of Personality*, *44*(1), 109-132.

Heider, F. (1958).*The psychology of interpersonal relations*. New York: John Wiley & Sons.

Heyman, G. D., & Giles, J. W. (2004). Valence effects in reasoning about evaluative traits. *Merrill-Palmer Quarterly*, *50*(1), 86-109.

Higgins, E. T., & Bryant, S. L. (1982). Consensus information and fundamental attribution error: The role of development and in-group versus out-group knowledge. *Journal of Personality and Social Psychology*, *43*(5), 889-900.

Hitchcock, C., & Knobe, J. (2009). Cause and norm. *The Journal of Philosophy, 106*, 587-612.

Hollander, M. A., Gelman, S. A., & Star, J. (2002). Children's interpretation of generic noun phrases. *Developmental Psychology*, *38*(6), 883-894.

Johnson, S. C., Dweck, C. S., & Chen, F. S. (2007). Evidence for infants' internal working models of attachment. *Psychological Science*, *18*(6), 501-502.

Jones, E. E. (1979). The rocky road from acts to dispositions. *American Psychologist*, *34*(2), 107-117.

Jones, E. E., & Davis, K. E. (1965). A theory of correspondent inferences: From acts to dispositions. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 2, pp. 219-266). New York: Academic Press.

Jones, E. E., & Harris, V. A. (1967). The attribution of attitudes. *Journal of Experimental Social Psychology*, *3*(1), 1-24.

Keates, J., & Graham, S. A. (2008). Category markers or attributes: Why do labels guide infants' inductive inferences? *Psychological Science, 19*, 1287-1293.

Keenan, E. L., & Stavi, J. (1986). A semantic characterization of natural language determiners. *Linguistics & Philosophy, 9*, 253-326.

Kelley, H. H. (1967). Attribution theory in social psychology. In D. Levine (Ed.), *Nebraska Symposium on Motivation* (Vol. 15, pp. 192-240). Lincoln, NE: University of Nebraska Press.

Kelley, H. H. (1973). The processes of causal attribution. *American Psychologist*, *28*(2), 107-128.

Keil, F.C. (1989). *Concepts, kinds and cognitive development*. Cambridge, MA: MIT Press.

Keil, F. C. (2006). Explanation and understanding. *Annual Review of Psychology*, *57*, 227-254.

Kinzler, K. D., & Shutts, K. (2008). Memory for "mean" over "nice": The influence of threat on children's face memory. *Cognition, 107*, 775-783.

Krifka, M., Pelletier, F. J., Carlson, G. N., ter Meulen, A., Link, G., & Chierchia, G. (1995). Genericity: An introduction. In G. N. Carlson, & F. J. Pelletier (Eds.), *The generic book* (pp. 1-124). Chicago, IL: The University of Chicago Press.

Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *Nature Reviews Neuroscience*, *5*(11), 831-843.

Kuhlmeier, V., Wynn, K., & Bloom, P. (2003). Attribution of dispositional states by 12-month-olds. *Psychological Science*, *14*(5), 402-408.

Lawler, J. M. (1973). Studies in English generics. *University of Michigan Papers in Linguistics, 1*(1), 1-184.

Lawson, C. A., & Kalish, C. W. (2006). Inductive inferences across time and identity: Are category members more alike than single individuals? *Journal of Cognition and Development*, *7*(2), 233-252.

Leslie, A. (1994). ToMM, ToBY, and agency: Core architecture and domain specificity. In L. A. Hirschfeld & S. A. Gelman (eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 119-148). New York: Cambridge University Press.

Leslie, S. J. (2007). Generics and the structure of the mind. *Philosophical Perspectives*, *21*(1), 375-403.

Leslie, S. J. (2008). Generics: Cognition and acquisition. *Philosophical Review*, *117*(1), 1-47.

Leslie, S. J. (2013). Generics oversimplified. *Noûs*. doi: 10.1111/nous.12039

Leslie, S. J., Khemlani, S., & Glucksberg, S. (2011). Do all ducks lay eggs? The generic overgeneralization effect. *Journal of Memory and Language*, *65*(1), 15-31.

Leslie, S. J., & Gelman, S. A. (2012). Quantified statements are recalled as generics: Evidence from preschool children and adults. *Cognitive Psychology, 64(3)*, 186-214.

LoBue, V., & DeLoache, J. S. (2010). Superior detection of threat-relevant stimuli in infancy. *Developmental Science*, *13*(1), 221-228.

Lockhart, K. L., Chang, B., & Story, T. (2002). Young children's beliefs about the stability of traits: Protective optimism? *Child Development*, *73*(5), 1408-1430.

Lombrozo, T. (2006). The structure and function of explanations. *Trends in Cognitive Sciences*, *10*(10), 464-470.

Luo, Y., & Baillargeon, R. (2005). Can a self-propelled box have a goal? Psychological reasoning in 5-month-old infants. *Psychological Science*, *16*(8), 601-608.

Mannheim, B., Gelman, S. A., Escalante, C., Huayhua, M., & Puma, R. (2011). A developmental analysis of generic nouns in Southern Peruvian Quechua. *Language Learning and Development*, *7*(1), 1-23.

Markman, E. M. (1989). *Categorization and naming in children: Problems of induction*. MIT Press: Cambridge, MA.

Martijn, C., Spears, R., Van der Pligt, J., & Jakobs, E. (1992). Negativity and positivity effects in person perception and inference: Ability versus morality. *European Journal of Social Psychology*, *22*(5), 453-463.

McArthur, L. A. (1972). The how and what of why: Some determinants and consequences of causal attribution. *Journal of Personality and Social Psychology*, *22*(2), 171-193.

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

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, *92*(3), 289-316.

Orvis, B. R., Cunningham, J. D., & Kelley, H. H. (1975). A closer examination of causal inference: The roles of consensus, distinctiveness, and consistency information. *Journal of Personality and Social Psychology*, *32*(4), 605-616.

Osherson, D. N., Smith, E. E., Wilkie, O., Lopez, A., & Shafir, E. (1990). Category-based induction. *Psychological Review*, *97*(2), 185-200.

Pinker, S. (1994). *The language instinct: How the mind creates language*. New York: William Morrow and Company.

Pinker, S. (1997). *How the mind works*. New York: W. W. Norton & Company.

Reeder, G. D., & Brewer, M. B. (1979). A schematic model of dispositional attribution in interpersonal perception. *Psychological Review*, *86*(1), 61-79.

Rholes, W. S., & Ruble, D. N. (1984). Children's understanding of dispositional characteristics of others. *Child Development, 55*(2), 550-560.

Rips, L. J. (1975). Inductive judgments about natural categories. *Journal of Verbal Learning and Verbal Behavior*, *14*(6), 665-681.

Ross, L. D. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 173-220). New York: Academic Press.

Ross, L. D., Amabile, T. M., & Steinmetz, J. L. (1977). Social roles, social control, and biases in social-perception processes. *Journal of Personality and Social Psychology, 35*(7), 485-494.

Rothbart, M., & Park, B. (1986). On the confirmability and disconfirmability of trait concepts. *Journal of Personality and Social Psychology*, *50*(1), 131-142.

Setoh, P., Wu, D., Baillargeon, R., & Gelman, R. (2013). Young infants have biological expectations about animals. *Proceedings of the National Academy of Sciences, 110*, 15937-15942.

Song, H., Baillargeon, R., & Fisher, C. (2005). Can infants attribute to an agent a disposition to perform a particular action? *Cognition, 98*, B45-B55.

Skowronski, J. J., & Carlston, D. E. (1987). Social judgment and social memory: The role of cue diagnosticity in negativity, positivity, and extremity biases. *Journal of Personality and Social Psychology*, *52*(4), 689-699.

Sloman, S. A. (1993). Feature-based induction. *Cognitive Psychology*, *25*(2), 231-280.

Smith, E. E., & Medin, D. L. (1981). *Categories and concepts*. Cambridge, MA: Harvard University Press.

Sutherland, S. L., Cimpian, A., Leslie, S. J., and Gelman, S. A. (under review). *Memory errors reveal a bias to spontaneously generalize to categories*.

Tardif, T., Gelman, S. A., Fu, X., & Zhu, L. (2012). Acquisition of generic noun phrases in Chinese: Learning about lions without an '-s'. *Journal of Child Language*, *39*(1), 130-161.

Tausch, N., Kenworthy, J. B., & Hewstone, M. (2007). The confirmability and disconfirmability of trait concepts revisited: Does content matter? *Journal of Personality and Social Psychology*, *92*(3), 542-556.

Todorov, A., & Uleman, J. S. (2003). The efficiency of binding spontaneous trait inferences to actors' faces. *Journal of Experimental Social Psychology*, *39*(6), 549-562.

Trope, Y., & Gaunt, R. (2000). Processing alternative explanations of behavior: Correction or integration? *Journal of Personality and Social Psychology*, *79*(3), 344-354.

Trope, Y., & Gaunt, R. (2007). Attribution and person perception. *The SAGE Handbook of Social Psychology: Concise Student Edition*, 176.

Uleman, J. S. (1999). Spontaneous versus intentional inferences in impression formation. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 141-160). New York: Guilford Press.

Uleman, J. S., Saribay, S. A., & Gonzalez, C. M. (2008). Spontaneous inferences, implicit impressions, and implicit theories. *Annual Review of Psychology*, *59*, 329-360.

Vaish, A., Grossmann, T., & Woodward, A. (2008). Not all emotions are created equal: The negativity bias in social–emotional development. *Psychological Bulletin, 134*(3), 383-403.

Weiner, B. (1985). “Spontaneous” causal thinking. *Psychological Bulletin*, *97*(1), 74-84.

Wertz, A. E., & Wynn, K. (2014). Selective social learning of plant edibility in 6-and 18-month-old infants. *Psychological Science*. doi:10.1177/0956797613516145

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor’s reach. *Cognition, 69*, 1-34.

Woodward, A. L. (1999). Infants’ ability to distinguish between purposeful and non-purposeful behaviors. *Infant Behavior and Development*, *22*(2), 145-160.

APPENDIX

Table 1. *Sample item from Experiment 1*

Property Content	Condition	
	Kind-Features	**Individual-Traits**
Dangerous	**Background information:** One animal is called morseths. They live up to 50 years.[a] **Evidence:** *x%* of morseths chop off people's heads. **Conclusion:** Morseths chop off people's heads. Is this conclusion true or false?	**Background information:** One animal is called Morseth. He is 50 years old.[a] **Evidence:** Morseth has chopped off people's heads in *x%* of the situations where he had the opportunity. **Conclusion:** Morseth chops off people's heads. Is this conclusion true or false?
Plain	**Background information:** One animal is called zorbs. They live up to 50 years.[a] **Evidence:** *x%* of zorbs climb trees. **Conclusion:** Zorbs climb trees. Is this conclusion true or false?	**Background information:** One animal is called Zorb. He is 50 years old.[a] **Evidence:** Zorb has climbed trees in *x%* of the situations where he had the opportunity. **Conclusion:** Zorb climbs trees. Is this conclusion true or false?

[a] This second sentence served to disambiguate the referent of "one animal" in the preceding sentence (and of "some animals" in the instructions to participants, which were the same across conditions). "They live up to 50 years" signaled that the item is about a kind, whereas "He is 50 years old" signaled that the item is about a specific individual. These sentences were the same across the 24 trials in a condition.

Table 2. *Sample item from Experiment 2*

Property Content	Condition	
	Quantified Inferences over Individuals	**Quantified Inferences over Situations**
Dangerous	**Background information:** One animal is called morseths. They live up to 50 years.	**Background information:** One animal is called Morseth. He is 50 years old.
	Evidence: *x%* of morseths chop off people's heads.	**Evidence:** Morseth chopped off people's heads in *x%* of the situations where he had the opportunity.
	Conclusion: [Some/All][a] morseths chop off people's heads.	**Conclusion:** Morseth [sometimes/always][a] chops off people's heads.
	Is this conclusion true or false?	Is this conclusion true or false?
Plain	**Background information:** One animal is called zorbs. They live up to 50 years.	**Background information:** One animal is called Zorb. He is 50 years old.
	Evidence: *x%* of zorbs climb trees.	**Evidence:** Zorb climbed trees in *x%* of the situations where he had the opportunity.
	Conclusion: [Some/All][a] zorbs climb trees.	**Conclusion:** Zorb [sometimes/always][a] climbs trees.
	Is this conclusion true or false?	Is this conclusion true or false?

[a] "Some" and "sometimes" express existentially-quantified inferences. "All" and "always" express universally-quantified inferences.

Table 3. *Sample item from Experiment 3*

Property Content	**Condition**	
	Kind-Features	**Individual-Traits**
Distinctive	**Background information:** One animal is called morseths. They live up to 50 years. **Evidence:** (1) *x%* of morseths climb trees. (2) On this planet, very few other animals perform this behavior. This behavior is extraordinary. **Conclusion:** Morseths climb trees. Is this conclusion true or false?	**Background information:** One animal is called Morseth. He is 50 years old. **Evidence:** (1) Morseth has climbed trees in x% of the situations where he had the opportunity. (2) On this planet, very few other animals have ever performed this behavior when they had the opportunity. This behavior is extraordinary. **Conclusion:** Morseth climbs trees. Is this conclusion true or false?
Plain	**Background information:** One animal is called zorbs. They live up to 50 years. **Evidence:** (1) *x%* of zorbs wash themselves with water. (2) On this planet, many other animals also perform this behavior. This behavior is unremarkable. **Conclusion:** Zorbs wash themselves with water. Is this conclusion true or false?	**Background information:** One animal is called Zorb. He is 50 years old. **Evidence:** (1) Zorb has washed himself with water in *x%* of the situations where he had the opportunity. (2) On this planet, many other animals have also performed this behavior when they had the opportunity. This behavior is unremarkable. **Conclusion:** Zorb washes himself with water. Is this conclusion true or false?

Table 4. *Sample item from Experiment 4*

Property Content	Condition	
	Quantified Inferences over Individuals	**Quantified Inferences over Situations**
Distinctive	**Background information:** One animal is called morseths. They live up to 50 years. **Evidence:** (1) *x%* of morseths climb trees. (2) On this planet, very few other animals perform this behavior. This behavior is extraordinary. **Conclusion:** [Some/All][a] morseths climb trees. Is this conclusion true or false?	**Background information:** One animal is called Morseth. He is 50 years old. **Evidence:** (1) Morseth has climbed trees in *x%* of the situations where he had the opportunity. (2) On this planet, very few other animals have ever performed this behavior when they had the opportunity. This behavior is extraordinary. **Conclusion:** Morseth [sometimes/always][a] climbs trees. Is this conclusion true or false?
Plain	**Background information:** One animal is called zorbs. They live up to 50 years. **Evidence:** (1) *x%* of zorbs wash themselves with water. (2) On this planet, many other animals also perform this behavior. This behavior is unremarkable. **Conclusion:** [Some/All][a] zorbs wash themselves with water. Is this conclusion true or false?	**Background information:** One animal is called Zorb. He is 50 years old. **Evidence:** (1) Zorb has washed himself with water in *x%* of the situations where he had the opportunity. (2) On this planet, many other animals have also performed this behavior when they had the opportunity. This behavior is unremarkable. **Conclusion:** Zorb [sometimes/always][a] washes himself with water. Is this conclusion true or false?

[a] "Some" and "sometimes" express existentially-quantified inferences. "All" and "always" express universally-quantified inferences.

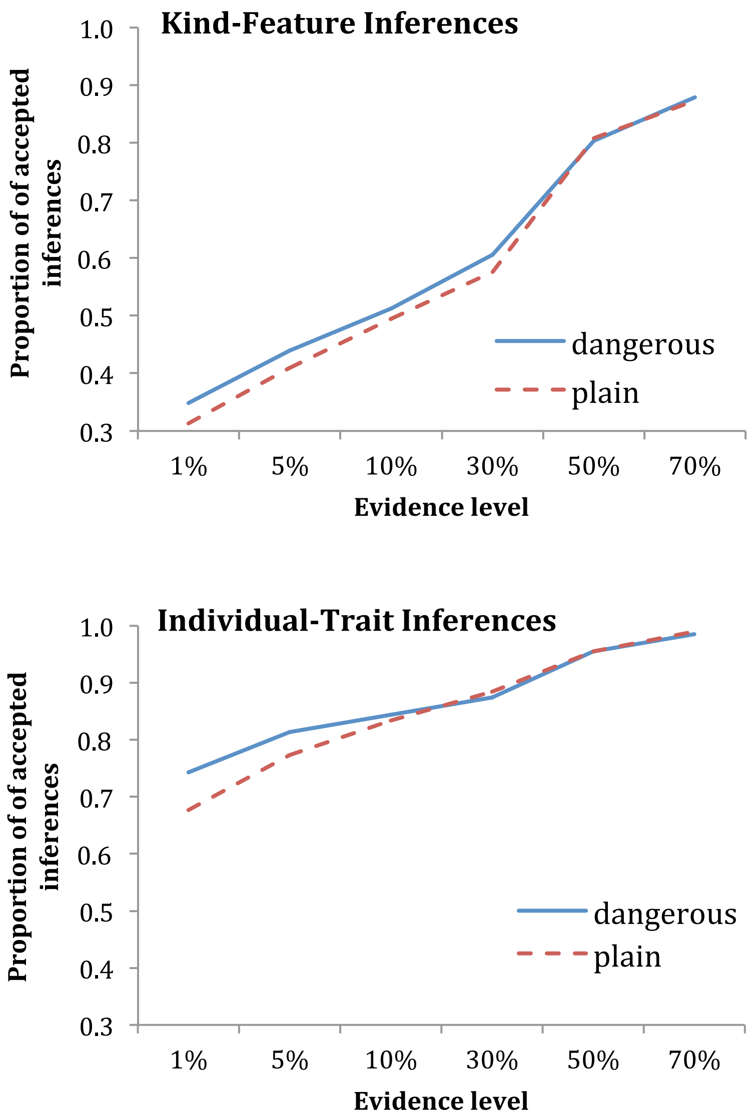


Figure 1. Average proportion of "true" responses in Experiment 1, by inference type, property content, and level of evidence.

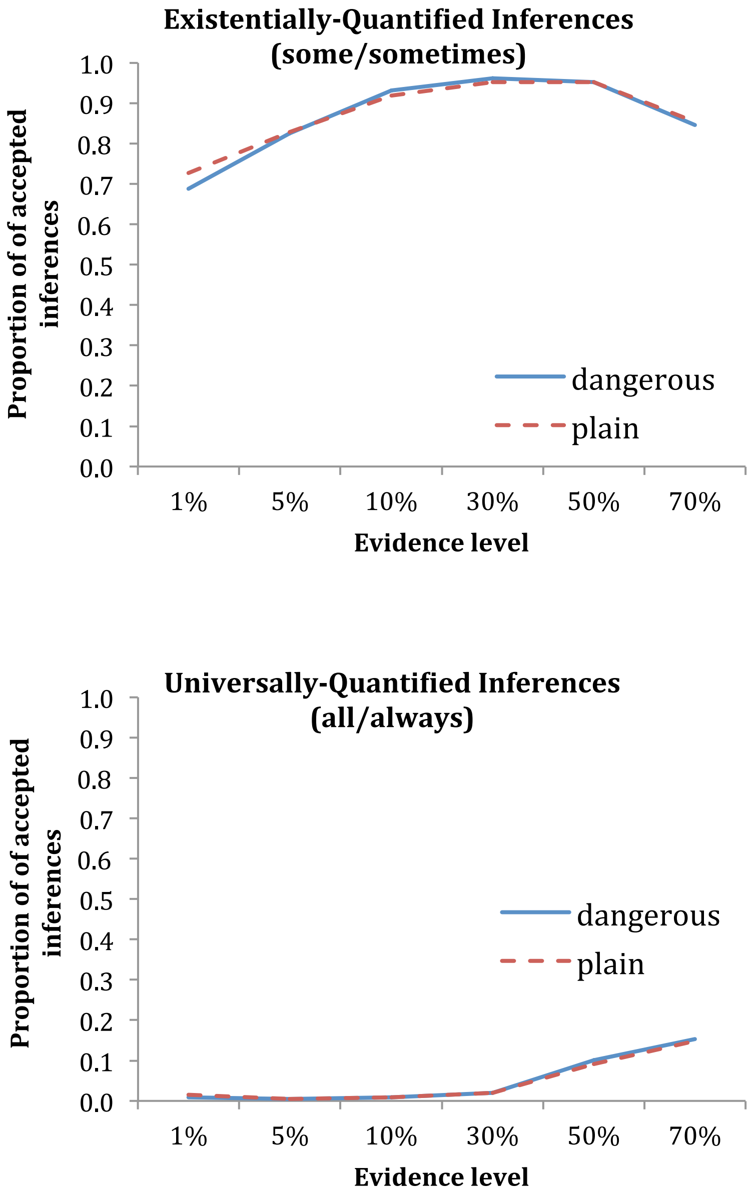
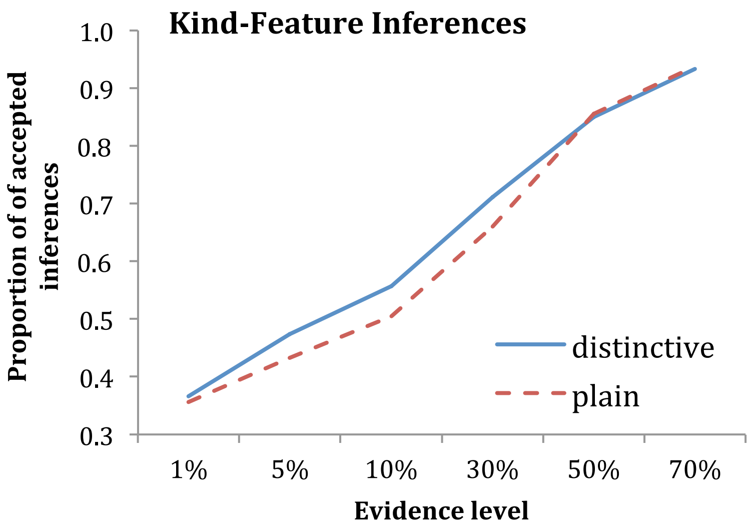


Figure 2. Average proportion of “true” responses in Experiment 2, by quantifier, property content, and level of evidence.



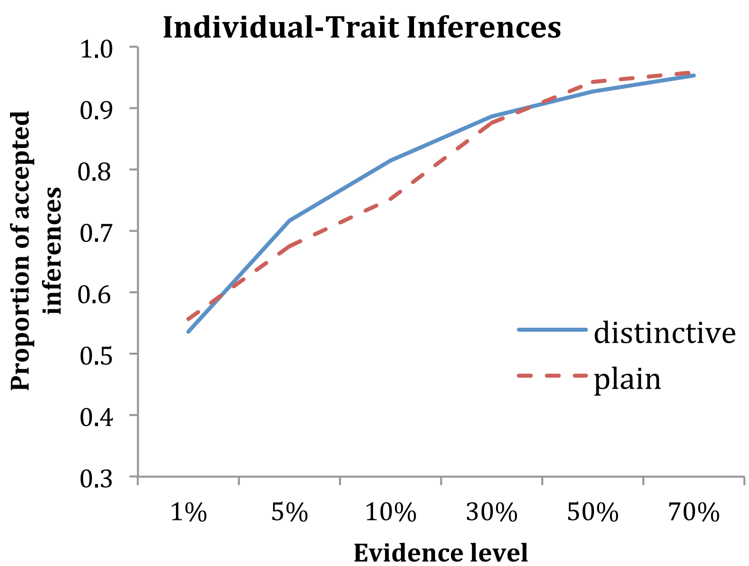
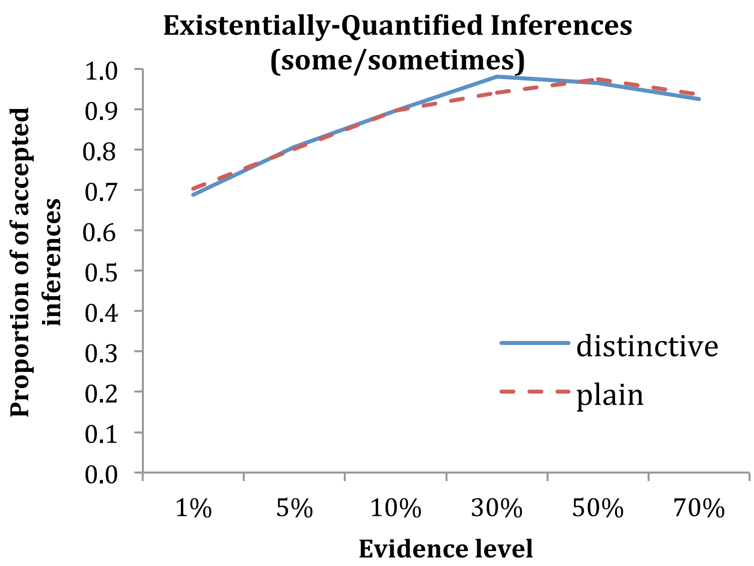


Figure 3. Average proportion of "true" responses in Experiment 3, by inference type, property content, and level of evidence.



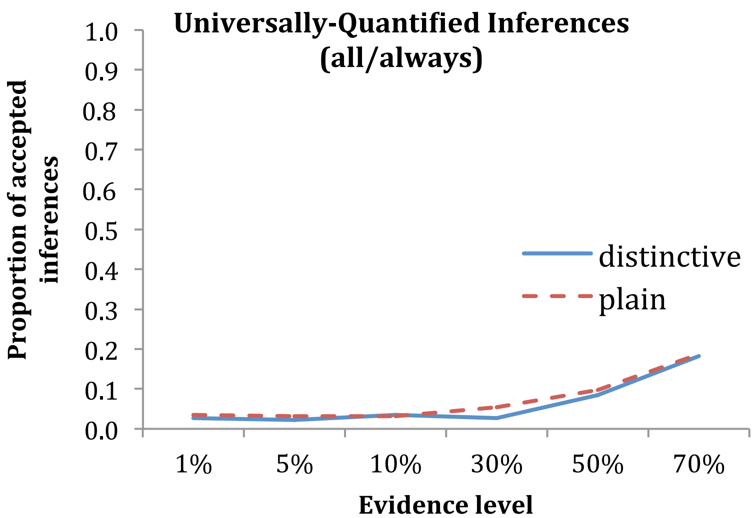


Figure 4. Average proportion of “true” responses in Experiment 4, by quantifier, property content, and level of evidence.