Cognition 176 (2018) 255–268

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Original Articles

The development of principled connections and kind representations

Paul Haward[a,*], Laura Wagner[b], Susan Carey[a], Sandeep Prasada[c]

[a] Harvard University, United States
[b] Ohio State University, United States
[c] Hunter College, City University of New York, United States

ARTICLE INFO

Keywords:
Conceptual representation
Kind representations
Modes of explanation
Generic knowledge
Principled connections

ABSTRACT

Kind representations draw an important distinction between properties that are understood as existing in instances of a kind by virtue of their being the kind of thing they are and properties that are not understood in this manner. For example, the property of barking for the kind dog is understood as being had by dogs by virtue of the fact that they are dogs. These properties are said to have a *principled connection* to the kind. In contrast, the property of wearing a collar is not understood as existing in instances by virtue of their being dogs, despite the fact that a large percentage of dogs wear collars. Such properties are said to have a *statistical connection* to the kind. Two experiments tested two signatures of principled connections in 4–7 year olds and adults: (i) that principled connections license normative expectations (e.g., we judge there to be something *wrong* with a dog that does not bark), and (ii) that principled connections license formal explanations which explain the existence of a property by reference to the kind (e.g., that barks *because it is a dog*). Experiment 1 showed that both the children and adults have normative expectations for properties that have a principled connection to a kind, but not those that have a mere statistical connection to a kind. Experiment 2 showed that both children and adults are more likely to provide a formal explanation when explaining the existence of properties with a principled connection to a kind than properties with statistical connections to their kinds. Both experiments showed no effect of age (over ages 4, 7, and adulthood) on the extent to which participants differentiated principled and statistical connections. We discuss the implications of the results for theories of conceptual representation and for the structure of explanation.

1. Introduction

Human beings are alone in the animal kingdom in developing an extraordinary repertoire of intricate kind representations—representations for kinds of entities like *dogs*, *watches*, *cities*, *triangles* and *atoms*. Kind representations play a central role in human thought. They underlie the meanings of most count and mass nouns in natural language, and as such, they provide an important interface between non-linguistic conceptual structure and combinatorial, hierarchical, unbounded linguistically expressible thought.

Given the centrality of kind representations in common sense thought and language, investigating the characteristics of kind representations and how they are acquired is a central task in theories of conceptual representation and of conceptual development (Cimpian, 2016; Gelman, 2003; Macnamara, 1986; Margolis, 1998; Prasada, 2016; Xu, 2005, 2012). To investigate kind representations, we must distinguish between the specific content of the representation of any one kind of thing (i.e., information specific to dogs, tables, and trees) and the abstract structure of kind representations which underlies kind representations of any and all kinds of things. The specific content of kind representations, unlike the abstract structure of kind representations, varies from kind to kind. So, for example, the kind representation for dogs will include information that characterizes dogs (e.g., that they are animals, that they bark, have fur, have four legs) and distinguishes dogs from other kinds of things, and the kind representation for tables will include information that characterizes tables (e.g., that they are made by humans, are furniture, have tops, are for putting things on) and distinguishes tables from other kinds of things, and so on for each specific kind of thing. Despite these differences, there is evidence for a common abstract structure underlying the representations of kind concepts in general. This structure is what makes the representations *kind* representations and explicating it provides an abstract characterization of how humans think and speak about any and all kinds of things.

* Corresponding author at: Department of Psychology, Harvard University, 33 Kirkland St, Cambridge, MA 02138, United States.
E-mail address: haward@fas.harvard.edu (P. Haward).

https://doi.org/10.1016/j.cognition.2018.02.001
Received 27 February 2017; Received in revised form 9 January 2018; Accepted 3 February 2018

1.1. Characteristics of adult kind representations and evidence that young children's kind representations have at least some of the same characteristics

Most fundamentally, all kind representations have the dual function of providing the means for thinking and talking about entities as instances of kinds, and the means for thinking and talking about kinds, themselves (Prasada, 2016). For example, not only can we form the thought that Fido is an instance of the kind dog, we can entertain thoughts in terms of kinds themselves, such as the thought that dogs evolved from wolves (Carlson, 1980; Gelman, 2003). This latter thought is not the thought that individual dogs evolved from individual wolves, but that dogs as a kind evolved from wolves as a kind. Both of these functions of kind representations are evident early in development. Children's extensions of nouns across multiple individuals, as well as the category-based inductions they make of properties across these individuals, reveal their ability to think of distinct things as being instances of the same kind of thing. Furthermore, by age two and a half, and possibly earlier, children can understand and use generic sentences to think and talk about kinds and the properties that characterize them (e.g., "Birds fly," Gelman, 2003).

Importantly, kind representations allow us to characterize kinds in ways that do not reduce to noting what is true of all, most, or many members of the kind. By at least three years of age, generic statements like *dogs have four legs* and *watches tell time* are understood as attributing properties to kinds and are interpreted distinctly from statements about quantified sets of individuals (e.g., *some/all/most dogs wear collars*) (Brandone, Gelman, & Hedglen, 2014; Hollander, Gelman, & Star, 2002; Leslie & Gelman, 2012). This capacity to predicate properties of kinds is one characteristic of the human endowment for representing kinds. A consequence of the profound difference between kind representations and quantified representations of sets of individuals is an asymmetry in the statistical inferences that *follow* from *learning* a generic generalization and the statistical evidence that supports *judging the truth* of a generic generalization. For example, when participants are introduced to a property of a novel kind using a generic statement (e.g., "Lorches have purple feathers."), they expect the property to apply to nearly *all* members of the kind (Cimpian, Brandone, & Gelman, 2010). Conversely, if they simply learn of particular Lorches that some of them have purple feathers (e.g., 30% or 50% of the Lorches they encounter have purple feathers), they subsequently judge the same generic ("Lorches have purple feathers.") as true, and they do this for a wide variety of percentages (even in conditions where as few as 10% of Lorches have purple feathers). This striking asymmetry in how generic statements are interpreted is a reliable part of the way 4-to 7-year-olds, as well as adults, interpret generics and incorporate property information into their kind representations (Brandone et al., 2014). Thus, these expectations likely reflect the abstract structure of kind representations.

In addition to the distinction detailed above between kind representations and representations of prevalent features of sets of individuals, further aspects of the abstract structure of kind representations distinguish kind representations within specific domains (e.g., the animal kind dog) from representations of quantified sets of individuals (e.g., all existent dogs, all of the dogs who ever played Lassie on the TV series, Lassie). For instance, natural kind representations, including representations of animals, plants, and substances, are structured by the assumptions of psychological essentialism (Gelman, 2003)—the assumption that causally deep, perhaps unknown, features of the members of the kind explain how new members come into existence and explain why those members have their kind relevant properties. These schemata, too, are abstract, early developing, and are made available when we think about things from the perspective of a natural kind.

1.2. A further abstract characteristic of adult kind representations: Principled connections between kinds and properties

Recent work has confirmed Aristotle's observation (Charlton, 1970) that kind representations draw an important distinction between properties that are understood as existing in instances of a kind *by virtue of* their being the kind of thing they are and properties that are not understood in this manner.[1] For example, the properties of *barking* or *having four legs* for the kind dog are both understood as being had by dogs by virtue of the fact that they are dogs: these are among the properties that are understood as making the kind what it is. In contrast, the property of *wearing a collar* is not understood as existing in instances by virtue of their being dogs, despite the fact that a large percentage of dogs wear collars (Prasada & Dillingham, 2006). This distinction is central to kind representations and generalizes to all kinds: for example, for the artifact kind *watch* the property of *telling time* is understood to be true of individual watches because they are the kinds of things they are, whereas the property of having a round face is not understood as being true of individual watches because they are the kinds of things they are, even though we assent to the generic proposition *watches have round faces*.

Properties such as *barking* for *dogs*, or *telling time* for *watches* are said to bear a *principled connection* to the kind (Prasada & Dillingham, 2006, 2009). They are those properties that capture part of what it means to be a member of that kind of thing. And, critically, *all* kinds across *all* domains possess properties that bear a principled connection to the kind. For example, for abstract mathematical kinds, such as triangles, the property of having three sides is understood to have a principled connection to being a triangle, and for social kinds such as architects, the property of designing buildings is understood to have a principled connection to being an architect. Furthermore, principled connections can be distinguished from *statistical connections,* which represent properties that are merely highly statistically correlated with particular kinds. In this context, statistical connections between properties and kinds serve as a control in the search for signatures of principled connections: both support generic generalizations (we assent to "dogs bark" and "dogs wear collars," "watches tell time" and "watches have round faces", and so on, in spite of only the first generic in each pair expressing a principled connection).

Principled connections between properties and kinds have a number of unique conceptual and linguistic consequences. First, properties that bear a principled connection to the kind license *normative expectations* concerning the presence of properties in instances of the kind (Prasada & Dillingham, 2006, 2009). For example, adults judge that there is something *wrong* with a dog that does not bark, but they do not have equivalent expectations for statistical connections: there is nothing wrong with a dog that does not wear a collar. That is, if an instance of a kind lacks a principled property, it is judged to be incomplete or to have something wrong with it, whereas if it lacks a property merely statistically associated with the kind, no such judgment is licensed.

In addition, principled connections license *formal explanations*—references to the kind in order to explain the existence of a property in instances of that kind. For example, we can explain the existence of a property that has a principled connection to a kind in an instance of the kind by simply citing the category: e.g., that thing tells time *because it is a watch.* In contrast, strong statistical relations are not enough; although barns are typically red (a statistical property of barns), explanations that seek to explain a barn's redness by citing the fact that it is a barn were rated as being significantly less natural (Prasada & Dillingham, 2006, 2009; Prasada, Khemlani, Leslie, & Glucksberg, 2013). As with the licensing of *by virtue of* statements and normative expectations, this signature of principled connections generalizes to all domains. For example, we can explain why a person designs buildings by citing the fact that she is an architect.

Though previous research has not explicitly investigated the development of principled connections and these signatures, there are

[1] The distinction in Aristotle regards kinds, not kind representations, but it is kind representations that concern us as psychologists.