Cogn Process
DOI 10.1007/s10339-008-0220-4

RESEARCH REPORT

How to reason without words: inference as categorization

Ronaldo Vigo · Colin Allen

Received: 2 February 2008 / Revised: 27 May 2008 / Accepted: 20 June 2008

Abstract The idea that reasoning is a singular accomplishment of the human species has an ancient pedigree. Yet this idea remains as controversial as it is ancient. Those who would deny reasoning to nonhuman animals typically hold a language-based conception of inference which places it beyond the reach of languageless creatures. Others reject such an anthropocentric conception of reasoning on the basis of similar performance by humans and animals in some reasoning tasks, such as transitive inference. Here, building on the modal similarity theory of Vigo [J Exp Theor Artif Intell, 2008 (in press)], we offer an account in which reasoning depends on a core suite of subsymbolic processes for similarity assessment, discrimination, and categorization. We argue that premise-based inference operates through these subsymbolic processes, even in humans. Given the robust discrimination and categorization abilities of some species of nonhuman animals, we believe that they should also be regarded as capable of simple forms of inference. Finally, we explain how this account of reasoning applies to the kinds of transitive inferences that many nonhuman animals display.

Keywords Animal cognition · Categorization · Learning · Reasoning · Similarity

R. Vigo · C. Allen (✉)
Department of History and Philosophy of Science,
Cognitive Science Program, Indiana University,
1011 E Third St, Bloomington, IN 47405, USA
e-mail: colallen@indiana.edu

R. Vigo
e-mail: rvigo@indiana.edu

Introduction

The Stoic philosopher Chrysippus of Soli told of a hunting dog pursuing prey who, on coming to a crossroads, sniffed two of the roads leading away and immediately set off down the third without sniffing. Sorabji (1993) reports that Chrysippus did not take this to show that the dog really reasons, but only that it "virtually" goes through a syllogism: "The animal went either this way, or that way, or the other way. But not this way, or that way. So that way" (Sorabji 1993, p. 26). Medieval logicians referred to the binary choice version of this disjunctive syllogism as *modus tollendo ponens* (MTP), and it has occasionally been called the "rule of dogs".

Whether or not the Ancients and Medievals believed that dogs mentally rehearse syllogisms, they certainly believed that the behavior of Chrysippus's dog conformed to a valid pattern of reasoning that we, as humans, can explicitly employ and understand. Of course, the fact that some nonhuman animal behavior conforms to a rule of inference is not sufficient evidence to accept the hypothesis that these animals are genuinely capable of reasoning. Indeed, recent experiments by Watson et al. (2001) that are designed to test dogs' abilities to use MTP have revealed possibly significant differences between dogs and children. In a searching task, dogs were slower to search a third location (of three) for a hidden reward when the two others had already been eliminated by unsuccessful search, whereas 4 to 6-year-old human children were faster to search the uneliminated third location. Watson and his colleagues conclude that "the observed contrast in response timing between children and dogs would seem most parsimoniously viewed as indicating that dogs rely on associative guidance and children rely, to some degree, on logical guidance when searching for objects that have recently disappeared" (Watson et al. 2001, p. 225). (Note, however, that the evidence remains open to other interpretations: human children would also slow down after some number of unsuccessful searches.)

Watson and colleagues are by no means alone in framing issues in terms of associative learning versus reasoning (see, for example, several of the contributions to Hurley and Nudds 2006, and the review by Watanabe and Huber 2006). McGonigle and Chalmers, who argued that monkeys are rational (McGonigle and Chalmers 1992), stated in a recent paper that "There is a widespread view that the sorts of animal learning mechanisms most frequently studied in the laboratory are inductively too weak and unproductive to generate the kinds of behaviors expressed in higher order forms of human cognition and linguistic adaptation (Chomsky 1980; Fodor and Pylyshyn 1988; Piaget 1971)" (McGonigle and Chalmers 2002). Their mention of linguistic adaptation serves to remind us that the widespread reluctance to attribute genuine reasoning to animals goes hand-in-hand with a specific conception of inference as a process that depends essentially on linguistically represented premises and conclusions. This conception makes it hard to credit animals with reasoning capacities. As Bermúdez (2003, p. 140) puts it, "there is no hope of applying an inference-based conception of rationality at the nonlinguistic level."

Our goal in this paper is to show that there is hope. We describe an alternative framework that is capable of providing a unified approach to reasoning and the subsymbolic perceptual processes underlying similarity assessment, discrimination, and categorization. The framework is provided by the modal similarity theory (MST) of Vigo (2008), which we describe in "Modal similarity theory". MST introduces the concept of modal similarity and shows how one may construe the propositional connectives as expressing degrees of modal similarity that can be investigated empirically in humans and nonhumans. This approach allows us to recast the notion of inference in subsymbolic, nonlinguistic terms. We do not deny, of course, that inferences can be carried out symbolically or using linguistic vehicles. Nor do we deny that reasoning can be greatly facilitated by interaction with formal or linguistic notations. But we believe that it is not necessary to characterize inferences in such terms. It is our view that the strong connection between reasoning and language that is usually taken for granted is, in fact, a specific cultural product that is partly the result of the way inferences are taught in logic and mathematics courses. That is, reasoning is typically defined and represented by logicians and mathematicians in terms of linguistic constructs such as sentences and propositions. But, as we shall argue below, this need not be the case.

What is inference?

For the purposes of this paper we start from generic conceptions of 'inference' and 'reasoning', such as are found in standard dictionary definitions. Thus, for example, an inference may be defined as "a conclusion reached on the basis of evidence and reasoning [or] the process of reaching such a conclusion", and 'reasoning' may be defined as "the power of the mind to think, understand, and form judgments by a process of logic" (Dictionary 2008). Neither of these definitions mentions sentences, words, or other linguistic vehicles. One might argue, however, that by referring to a "process of logic" in the definition of 'reasoning', the dictionary defers to experts in logic for full specification of what is to count as reasoning and inference.

Logicians' standard definitions of inference make it seem quite implausible to attribute the capacity to nonhuman animals. Logic is concerned with the relationships between premises and conclusions of arguments. The premises of an argument are often explicitly defined as a set of sentences, and the conclusion is defined as another sentence. Semantic conceptions of inference are grounded in semantic relationships between premises and conclusion: namely, those of (truth-preserving) validity and (truth-enabling) inductive strength. In turn, syntactic conceptions of inference are grounded in the notion of a derivation of a conclusion from a set of premises by the sequential application of rules to structured formulae or strings of symbols. Both conceptions are language-driven.

On the one hand, the semantic conception of inference is language-driven because it relies on sentences as bearers of truth values or probability assignments. When the basic sentences are atomic in nature (i.e., not decomposable), as is the case in the sentential calculus, their inner structure is not considered when determining the deductive validity or inductive strength of the inference. In deductive logic, whether a conclusion follows necessarily from a set of premises is encoded in the meaning of the sentential connectives considered as propositional transformations or propositional functions. This functional or relational meaning of the connectives is based on truth values. In other words, under this type of language-oriented sense of inference, the process of reasoning validly is construed as a sequence of transformations that preserves truth among sets of sentences.

On the other hand, the syntactic conception of inference treats reasoning as a process of rule satisfaction. Reasoning is characterized as the application of symbolic rules to structured strings of symbols. The logical connectives are syncategorematic symbols, whose sole purpose is to serve as the structural "glue" between sentential parts. This viewpoint is consistent with the view known as formalism in the philosophy of mathematics. Formalists hold that all inference amounts to nothing more than the application of rules of symbolic manipulation to strings of symbols. This view of reasoning has been challenged with questions such as "what is the nature of the rules?" and "why are certain rules chosen over an infinitude of others?" Even if these challenges can be met, the formalist's assumption that reasoning necessitates symbolic or linguistic representation of premises and conclusions deserves scrutiny.

Both conceptions are language-driven, and their status as orthodoxy leads even those who are quite impressed with the cognitive capacities of nonhuman animals to attribute, at most, a "proto-reasoning" capacity to languageless creatures (e.g., Bermúdez 2003). The underlying conception of reasoning, and the semantic and syntactic approaches to inference, are products of attempts by philosophers and scientists, dating back to Aristotle, to render precise the logical discourse which is the backbone of mathematics and science. Unfortunately, from our perspective, the real, live, reasoning human agent has been ignored in this tradition. That is, reasoning as a living, everyday process underlying decision making took a back seat to interpretations enforced upon it by the needs of the project of rationally reconstructing the methods of science and mathematics.

Not until the emergence of psychology as a legitimate field of science was reasoning construed as not being necessarily language or symbol oriented and dependent. For example, some of the early Gestalt psychologists explained perceptual judgments as the product of processes that subconsciously mirror deductively valid and inductively strong inferences while acting on physical stimuli in accordance with basic principles (Köhler 1929). But rather than thinking of the general capacity for inference as "ratiomorphic" (Brunswick 1955), we prefer to think of the culturally acquired skill of manipulating formal symbols under the neologism "percipiomorphic" to emphasize the more fundamental role of the perceptual processes.

In "Inference without language" we argue that inference is not necessarily language-dependent, or even symbol-dependent. In subsequent sections we argue instead that inference is a subsymbolic process grounded in the cognitive meaning of the logical connectives. The phrase "cognitive meaning" denotes the cognitive capacities and processes that explain agents' use of the logical connectives.

What is the cognitive meaning of the logical connectives? Our answer to this question will be that modal similarity judgments explain how the logical connectives are used. The intuition to be developed here is that the connectives are relations that specify which of a set of alternatives are possible members of a category. Take, for instance, material implication. The formula $p \supset q$ encodes the elimination of one possible state (p and not-q) and the inclusion of the remaining three of the four possible states that the two entities or features (not necessarily propositions) represented by p and q may be in. The cognitive process of modal similarity assessment involves the comparison of the positive and negative states of (in this example) the conditional category to the template provided by the biconditional modal category. In the most general sense, the states of interest concern the presence or absence of discernable features. Thus MST, which we present more fully in the section titled "Modal similarity theory", assumes a far more general domain of application for the logical connectives than propositions, by generalizing the objects of inference to features of stimuli.

Inference without language

So, how plausible is it that inference can be understood as a nonlinguistic phenomenon? Language provides a number of cognitive benefits beyond its communicative role. We can use the phonological and visual properties of spoken and written or gestured language as aids to memory of concepts and events. Large amounts of information can be packaged in an efficient symbolic wrapping that by its very structure helps us overcome some of the limitations of storing and retrieving information in the brain. Thus it appears that linguistic representations can optimize human usage of available short-term and long-term memory resources. Moreover, the importance of language associations for semantic memory and rote learning has long been studied (Bower and Clark 1969).

Enhancement of learning and memory is not all that language appears to facilitate for human cognition. It may also organize the material over which inferential processes operate, by structuring high level concepts. Under this interpretation of language, its symbols provide categories such as subject, predicate, verb, which our cognitive system in turn utilizes recursively for organizing other sorts of categories or concepts. It was this role of language that the great mathematician Joseph Euler perhaps alluded to in his statement: "thought (and therefore reasoning) is not possible without language". We must also acknowledge that human facility with externalized symbol systems (e.g., marks on paper) supports much longer chains of reasoning than is possible without such "cognitive tools" (Clark 1998) and in "From ethological to logical" we discuss results that suggest the notations themselves can become the objects of modal categorization processes.

We contend, however, that language is no more necessary for logical reasoning than possession of a word for pain is necessary to feel pain. Although we cannot provide conclusive proof for this claim, there are several suggestive lines of evidence which motivate our attempt to develop an alternative framework. For example, case studies of aphasic patients provide evidence in support of the view that language is not essential to reasoning. Aphasia due to trauma to the regions of the the brain responsible for speech and language comprehension and generation may leave patients devoid of their fundamental ability to communicate even though other intelligence and reasoning abilities are intact (Siegal et al. 2001) leading many aphasia researchers to argue that the inferential capacity of humans can be exercised at a nonlinguistic level. Further evidence that human reasoning does not necessarily depend on language comes from experiments indicating that humans possess an uncanny ability to solve problems without making a single conscious, linguistically vehicled inference. Of course, some may argue that the fact that the process is subconscious does not make it language-independent. There is, however, evidence that whenever a solution to one of these problems emerges at a subconscious level, it is regions of the brain other than those responsible for language that exhibit any notable activity (Pernet et al. 2004). In addition, there are tasks on which humans excel rationally, such as geometric reasoning, but that are, prima facie at least, devoid of language-based inference (Battista and Clements 1991).

These arguments cast doubt on language as a necessary condition for reasoning. Given this doubt, we believe it is worth investigating the idea that other capacities are perhaps more fundamental than language for reasoning. One of these more fundamental capacities is categorization. Indeed, without the ability to form categories one would not be able to learn a language in the first place. That is, one needs categorical thought in order to make sense of both the syntax and the semantics of any language. For without the ability to group together and differentiate the various types of linguistic objects and their meanings one would only have within one's grasp a mental alphabet soup. Thus, language depends on categories but not vice versa.

This point regarding the centrality of categorization and conceptual abilities sets the stage for our eventual goal which is to show that reasoning depends on the very same modal and categorization abilities used in prelinguistic categorization tasks. Using Vigo's MST, we will lay out the following argument, which constitutes the "master argument" for this paper.

1. Logical connectives (conditional, biconditional, etc.) define modal similarity categories.
2. Inference is reducible to conditional categorization.
3. Hence, inference is modal similarity categorization.
4. Modal similarity categorization is a prelinguistic process.
5. Hence, inference is a prelinguistic process.

Although we cannot provide definitive proof for the premises of this argument, we believe that our approach provides a cognitively grounded way to understand the inferential capacities of a wide range of cognitive agents. MST explains how inferential capacities are composed of modal and conceptual capacities. Conceptual capacities are in turn composed of capacities for attention and categorization, which themselves are derived from primitive similarity and discrimination abilities. Building in the other direction, we surmise that the human capacity for reflective reasoning is built out of inferential capacities combined with the capacity for language. We do not argue for reflective reasoning in animals. Nor do we argue that nonhuman animals have the same set of reasoning capabilities as humans, but rather a subset of these, due perhaps to cognitive limitations such as lower storage capacity for certain types of memory, or different perceptual abilities.

Before introducing MST, we first describe the basic results for transitive inference in animals, which anchor our argument that nonhuman animals are capable of some form of inferential reasoning. Controversy about transitive inference has generated a wealth of articles examining the phenomenon, including various attempts to explain it in purely associationistic terms. (For introductions to this literature, see Zentall 2001; Allen 2006; Watanabe and Huber 2006.) MST provides a rigorous, empirically tractable conceptual underpinning that we believe is sufficient to reorient the millennia-old debate about animal reasoning.

Transitive inference

Animals living in social hierarchies observe and participate in dominance interactions whose ramifications extend beyond the immediate participants. A capacity for transitive inference would be advantageous in such circumstances, and various field and laboratory studies seem to indicate that such a capacity is present in a wide range of taxonomic groups. Although there is disagreement about the mechanisms underlying this capacity, mammals including primates (McGonigle and Chalmers 1992) and rodents (Dusek and Eichenbaum 1997), birds including pigeons (Zentall 2001), jays (Bond et al. 2003), chickens (Hogue et al. 1996), and Siamese fighting fish (Grosenick et al. 2007) all seem capable of responding to novel pairings as one would predict using TI. Such capacities may contribute to the fitness of the individuals living in dominance hierarchies (Seyfarth and Cheney 2002; Allen 2006; Grosenick et al. 2007).

The experimental basis for the claim that nonhuman animals engage in transitive inference originates from an experiment originally conducted by Piaget to study the logical development of children (Inhelder and Piaget 1958). This experiment involves training subjects on adjacent pairs of stimuli drawn from a strictly transitive ordering, and testing them on a novel pair drawn from the same set of stimuli. The most simple-minded version of the experiment uses only two pairs of (arbitrarily labeled) stimuli: (*a* +, *b*−) and (*b* +, *c*−). Here the plus sign stands for the presence of a reward (positive reinforcement) if that stimulus is selected by the subject, and the minus sign stands for the absence of a reward. The letters *a* through *c* are merely our labels for the stimuli, not the actual stimuli themselves, which may be arbitrary shapes, odors, etc. The order of presentation of the stimuli is typically counterbalanced across trials in experimental tests of transitive inference. Hence, the rewarded stimulus is never uniformly on the left or the right, for instance. When trained to assess these pairs of stimuli via operant conditioning, pigeons and rats are more likely to select stimulus *a* when presented with the novel pair (*a*, *c*). However, this particular result admits of a very simple associative explanation. During training, *a* was always rewarded and *c* never rewarded. Hence the preference for *a* over *c* can be explained entirely in terms of the past reinforcement history for the individual elements; the animal is simply picking the one that has been rewarded in the past.

This result leads to a slightly more sophisticated experiment that has become the industrial standard for laboratory investigations of transitive inference. In the 5-element procedure, due to Bryant and Trabasso (1971), the subjects are trained with four pairs of stimuli: (*a*+, *b*−), (*b*+, *c*−), (*c*+, *d*−), and (*d*+, *e*−). Once they have reached a certain criterion level of correct performance on these pairs, the subjects are then tested with the novel pair (*b*, *d*). Many kinds of animal (e.g., rats, pigeons, monkeys) tested in this way reliably select *b*. In the training set, *b* is rewarded exactly as frequently as *d* (on average, 50% of the time—i.e., always when paired with *c* and *e* respectively, and never when paired with *a* and *c*). Consequently there is no explanation of the preference for *b* over *d* simply in terms of the past history of direct reinforcement of choosing each of these individual elements. Although accounts of indirect reinforcement have been attempted (e.g., Fersen et al. 1991; Zentall 2001), many scientists believe a capacity for transitive inference is well established for at least some species of nonhuman animals (see Allen 2006; McGonigle and Chalmers 2002; Treichler 2007).

What exactly allows animals to induce new relationships among familiar stimuli? We hope to fill a theoretical gap, which leads some psychologists to look beyond traditional associationist concepts such as respondents and operants toward new notions such as "emergents" (Rumbaugh et al. 1996). A more concrete suggestion comes from McGonigle and Chalmers (2002) who invoke the idea of "private codes" to anchor "relational primitives" which order stimuli into the sequences required for transitive inference—e.g., bigness as an anchor for size relations. Similarly, Treichler et al. (2003) regard the standard associationist tool kit to be insufficient to account for facts surrounding the ability of monkeys to link three 5-element transitively ordered lists into a single 15-element list, and they suggest the need for internal organizational processes operating over hierarchically structured memory representations. We believe our approach is compatible with these, but we would argue that the formal approach offered by MST holds out the prospect of more specific models and predictions than the alternatives.

Modal similarity theory

Model similarity theory was developed by Vigo (2008) to show the relationship between the propositional connectives, similarity assessment, and categorization. The goal was to formulate a similarity measure that works at the level of single features, but is more distributive and relational than current representations of similarity. The key idea here is that modal similarity is a higher-order relationship between structures that are defined over the presence and absence of a first order feature φ in pairs of items that make up the modal categories. A modal category is a relational category consisting of pairs of items (*p* and *q*) where φ is either present in both $(p_{\mathsf{P}}(\varphi), q_{\mathsf{P}}(\varphi))$, absent in both $(p_{\mathsf{A}}(\varphi), q_{\mathsf{A}}(\varphi))$ or present in one and absent in the other (see Box 1). The importance of modal categories lies in the fact that the logical connectives (i.e., the building blocks of formal logic) define the possible modal categories in the sense that each function from a set of truth values to a truth value specifies completely and uniquely the general form of a particular modal category. The central claim of MST is that the propositional connectives as used by human agents in language express degrees of similarity between the modal states (presence or absence of features) of pairs of stimuli. Of course, for some theoretical purposes (for example, the project of rationally reconstructing the methods of science and mathematics that we mentioned in "What is inference?") one may choose to ignore everything but the truth functional aspect of the connectives. However, our claim is that cognitive subjects treat the connectives as defining relational categories and expressing modal similarity. MST thus builds a foundational bridge between similarity-based processes and the rule-based reasoning processes described by symbolic logic.

At the center of MST is the modal category **E**, defined by logical equivalence. **E** is particularly significant because

Box 1 MST analytic axioms and the modal similarity measure

The eight axioms for the modal similarity measure are presented here for inspection only. For explanation and justification see Vigo (2008). It should be noted that these analytic axioms are not meant as the basis for an axiomatic theory of representation; instead they are meant to summarize the basic assumptions underlying the measure. Below, $\mathbf{E} = \langle (p_{\mathsf{P}}(\varphi), q_{\mathsf{P}}(\varphi)), (p_{\mathsf{A}}(\varphi), q_{\mathsf{A}}(\varphi)) \rangle$, Ψ is a modal category, φ is a property or feature, and the subscripts $_{\mathsf{P}}$ and $_{\mathsf{A}}$ represent the presence and absence of the feature φ respectively.

1. Existence property I (of primitive similarities and dissimilarities). For any pair of objects p and q there exist two primitive modal similarity measures and two primitive modal dissimilarity measures from which our total measure will be derived; namely, $\mathrm{sim}(p_{\mathsf{P}}(\varphi), q_{\mathsf{P}}(\varphi)), \mathrm{sim}(p_{\mathsf{A}}(\varphi), q_{\mathsf{A}}(\varphi)), \mathrm{dis}(p_{\mathsf{P}}(\varphi), q_{\mathsf{A}}(\varphi))$, and $\mathrm{dis}(p_{\mathsf{A}}(\varphi), q_{\mathsf{P}}(\varphi))$. These two measures are real-valued functions between zero and one that satisfy postulates 3 and 4 below.
2. Compositional property. Total modal similarity is a real-valued function of primitive similarity and dissimilarity: $\mathrm{Sim}_M(\mathbf{E}, \Psi, \varphi) = f(\mathrm{sim}(p_{\mathsf{P}}(\varphi), q_{\mathsf{P}}(\varphi)), \mathrm{sim}(p_{\mathsf{A}}(\varphi), q_{\mathsf{A}}(\varphi)), \mathrm{dis}(p_{\mathsf{P}}(\varphi), q_{\mathsf{A}}(\varphi)), \mathrm{dis}(p_{\mathsf{A}}(\varphi), q_{\mathsf{P}}(\varphi)))$
3. Order property. The primitive similarity and dissimilarity measures are ordered as follows: $\mathrm{sim}(p_{\mathsf{P}}(\varphi), q_{\mathsf{P}}(\varphi)) > \mathrm{sim}(p_{\mathsf{A}}(\varphi), q_{\mathsf{A}}(\varphi))$ and $\mathrm{dis}(p_{\mathsf{A}}(\varphi), q_{\mathsf{P}}(\varphi)) > \mathrm{dis}(p_{\mathsf{P}}(\varphi), q_{\mathsf{A}}(\varphi))$
4. Summation postulate. Partial similarity adds up to 1 and partial dissimilarity adds up to 1: $\mathrm{sim}(p_{\mathsf{P}}(\varphi), q_{\mathsf{P}}(\varphi)) + \mathrm{sim}(p_{\mathsf{A}}(\varphi), q_{\mathsf{A}}(\varphi)) = 1; \mathrm{dis}(p_{\mathsf{P}}(\varphi), q_{\mathsf{A}}(\varphi)) + \mathrm{dis}(p_{\mathsf{A}}(\varphi), q_{\mathsf{P}}(\varphi)) = 1$
5. Maximal and minimal modal similarity property. Maximal and minimal similarity values for the relational measure are given by: $\mathrm{Max}(\mathrm{Sim}_M(\mathbf{E}, \Psi, \varphi)) = \mathrm{Sim}_M(\mathbf{E}, \mathbf{E}, \varphi)$ and $\mathrm{Min}(\mathrm{Sim}_M(\mathbf{E}, \Psi, \varphi)) = \mathrm{Sim}_M(\mathbf{E}, \overline{\mathbf{E}}, \varphi)$ where $\overline{\mathbf{E}} = \langle (p_{\mathsf{P}}(\varphi), q_{\mathsf{A}}(\varphi)), (p_{\mathsf{A}}(\varphi), q_{\mathsf{P}}(\varphi)) \rangle$.
6. Zero property. The modal similarity value for the empty vector $\langle\rangle$ or $\varnothing$ is zero: $\mathrm{Sim}_M(\mathbf{E}, \varnothing, \varphi)) = 0$.
7. Contextual reversal property. $\mathrm{sim}(p_{\mathsf{P}}(\varphi), q_{\mathsf{P}}(\varphi)) = \alpha_1$ and $\mathrm{dis}(p_{\mathsf{P}}(\varphi), q_{\mathsf{A}}(\varphi)) = \alpha_2$, if $\mathbf{E} \not\subseteq \Psi$; $sim(p_{\mathsf{P}}(\varphi), q_{\mathsf{P}}(\varphi)) = \alpha_2$ and $\mathrm{dis}(p_{\mathsf{P}}(\varphi), q_{\mathsf{A}}(\varphi)) = \alpha_1$, if $\mathbf{E} \subseteq \Psi$.
8. Existence property II (of metafunction $\widetilde{\mathfrak{M}}$). For any pair of objects $(p(\varphi), q(\varphi)) \in \{(p_{\mathsf{P}}(\varphi), q_{\mathsf{P}}(\varphi)), (p_{\mathsf{A}}(\varphi), q_{\mathsf{A}}(\varphi)), (p_{\mathsf{P}}(\varphi), q_{\mathsf{A}}(\varphi)), (p_{\mathsf{A}}(\varphi), q_{\mathsf{P}}(\varphi))\}$, there exists a metafunction $\widetilde{\mathfrak{M}}$ such that:
$$\widetilde{\mathfrak{M}}_\Psi((p(\varphi), q(\varphi))) = \begin{matrix} 1 & \text{if}(p(\varphi), q(\varphi)) & \text{is present in}\,\Psi \\ 0 & \text{if}(p(\varphi), q(\varphi)) & \text{is absent in}\,\Psi \end{matrix}$$

Note that postulate 6 simply states that the empty modal relation $\varnothing$ expresses neither a degree of similarity nor a degree of dissimilarity: this means that in some sense it expresses "comparative" neutrality—or perhaps better yet, it expresses nothing in respect to similarity or dissimilarity. Postulate 7 simply states that the measure must do what we expected it to do: namely, to assign to each modal category corresponding to each of the logical connectives a unique degree of modal similarity.

The modal similarity measure

$$\mathrm{Sim}_M(\mathbf{E}, \Psi, \varphi) = \begin{matrix} \beta\left[\alpha_1 \widetilde{\mathfrak{M}}_\Psi(\mathbf{E}(1)) + (1-\alpha_1)\widetilde{\mathfrak{M}}_\Psi(\mathbf{E}(2))\right] & \\ -(1-\beta)\left[\alpha_2 \widetilde{\mathfrak{M}}_\Psi(\mathbf{E}'(1)) + (1-\alpha_2)\widetilde{\mathfrak{M}}_\Psi(\mathbf{E}'(2))\right] & \text{if } \widehat{\Psi} \in \{8, \ldots, 15\} \\ \beta\left[\alpha_2 \widetilde{\mathfrak{M}}_\Psi(\mathbf{E}(1)) + (1-\alpha_2)\widetilde{\mathfrak{M}}_\Psi(\mathbf{E}(2))\right] & \\ -(1-\beta)\left[\alpha_1 \widetilde{\mathfrak{M}}_\Psi(\mathbf{E}'(1)) + (1-\alpha_1)\widetilde{\mathfrak{M}}_\Psi(\mathbf{E}'(2))\right] & \text{if } \widehat{\Psi} \in \{0, , 7\} \end{matrix}$$

any instance of **E** contains two kinds of pairs of objects that are internally similar with respect to some first-order feature φ that is present in both items or absent in both. and these two kinds have a higher-order relationship of similarity between them in that the relationship between the members of each pair are identical (i.e., the members are the same as each other with respect to φ). **E** also excludes two kinds of pairs that are internally dissimilar but identical at the higher-order in that the members making up these excluded pairs are different from each other with respect to φ. This is why the other modal categories are suitably compared to **E**; these cross-category comparisons in effect involve a within-category comparison of the pairs that do and the pairs that do not belong to them.

Vigo's axioms for MST (Box 1) enable a precise statement of the contribution made by first-order similarities and dissimilarities to the overall higher-order similarity among the modal categories. The first-order similarities and dissimilarities are given at the level of the states of presence or absence of the feature φ in the paired objects making up the modal categories. Modal similarity is a higher order similarity relation, defined with respect to the feature φ as holding between the standard modal category **E** specified by the logical equivalence connective (i.e., the biconditional) and all 16 modal categories specified by the logical (truth-functional) connectives. **E** can be represented as a vector containing the two pairs of objects corresponding to modal identity, i.e., the two pairs of objects p and q where φ is either present in both, or absent in both objects (see Fig. 1). In MST, **E** is not a prototype for similarity judgments. Rather, **E** represents an upper boundary for maximal similarity. Correspondingly, $\mathbf{E}'$, the complement of **E** (which corresponds to negation of biconditional, i.e., exclusive or), represents the lower limit of similarity. This illustrates the distributive and relational nature of the theory: **E** is the category consisting of exactly those pairs whose elements are exactly similar to each other in respect to the feature φ and excluding those pairs

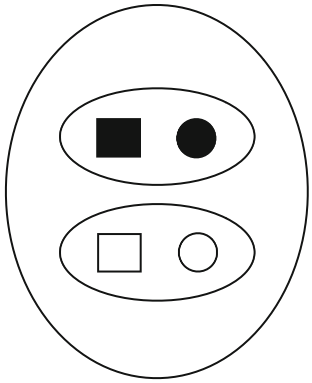

Fig. 1 Visual representation of the structure of the modal category defined by logical equivalence. The feature (*shading*) is present in both or absent in both members of the pairs belonging to the category

whose elements are exactly dissimilar with respect to φ, whereas $\mathbf{E}'$ consists of those pairs whose elements are dissimilar in respect to φ and excluding those pairs whose elements are exactly similar. The categories formed of these pairs are thus maximally dissimilar from each other.

Using the componential notion of similarity captured in MST, primitive similarity and dissimilarity with respect to the presence or absence of a feature φ are combined to define a structured measure of the degree of similarity between $\mathbf{E}$ and any modal category (represented by a vector Ψ). This measure contains parameters whose values reflect the differing degrees of salience between presence and absence of φ within a modal category, and the relative emphasis of similarity over dissimilarity between modal categories. With appropriately chosen values for these parameters, Vigo (2008) was able to generate and test modal similarity predictions for human subjects (see Table 1).

Subjects in Vigo's experiment were presented with iconic stimuli representing modal categories (see Fig. 2). (It is important for readers to bear in mind that these iconic representations are not used to"picture" the modal categories, but are instances of the categories which possess the appropriate structure.) Subjects were told nothing about the stimuli except that they represented the hat-wearing behaviors of married couples in different cultures. Each pair belonging to a modal category consisted of iconic representations of a male and a female, and each member of the pair was shown either wearing a hat or hatless. Subjects in one group were presented with stimuli representing the modal categories for equivalence (biconditional), conjunction, implication (material conditional), disjunction and negation of biconditional (exclusive or). Subjects in another group were presented with representations of biconditional, negation of disjunction, negation of conjunction, negation of implication, and negation of biconditional. Subjects were shown these iconic representations alongside of that for equivalence and asked to rate on a scale of 1–10 the similarity of each "culture" to the "culture" corresponding to $\mathbf{E}$ where both wear a hat or neither do. The average similarity ratings returned by subjects in this experiment provided an ordering exactly matched by the theory (Table 1). Thus, human subjects were sensitive to modal similarity even though they had not been given explicit instruction in logic, and were not provided with linguistic labels for the logical connectives. On the basis of these results, we argue that the Boolean operators (logical connectives) should be thought of as

Table 1 Modal similarity measure predictions when $\beta = 0.55$

$\mathfrak{M}_\mu(p_\mathsf{P}, q_\mathsf{P})$	$\mathfrak{M}_\mu(p_\mathsf{A}, q_\mathsf{A})$	$\mathfrak{M}_\mu(p_\mathsf{P}, q_\mathsf{A})$	$\mathfrak{M}_\mu(p_\mathsf{A}, q_\mathsf{P})$	Connective	$\mathrm{Sim}_M(p,q,\varphi)$
1	1	0	0	$\equiv$	0.55
1	0	0	0	$\wedge$	0.38
1	1	1	0		0.37
0	1	0	0	$\neg(\vee)$	0.33
1	1	0	1	$\supset$	0.28
1	0	1	0		0.21
0	1	0	1		0.20
1	0	0	1		0.12
1	1	1	1		0.10
0	1	1	0		0.02
0	0	0	0		0
1	0	1	1	$\vee$	−0.07
0	1	1	1	$\neg(\wedge)$	−0.12
0	0	0	1		−0.14
0	0	1	0	$\neg(\supset)$	−0.32
0	0	1	1	$\neg(\equiv)$	−0.45

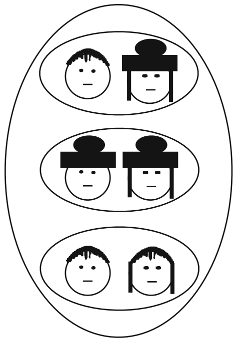

Fig. 2 Stimulus used by Vigo (2008) for the "conditional culture" category

expressing degrees of modal similarity, and that modal similarity assessment is the cognitive process underlying the use of the connectives. The stimuli in Vigo's experiment represent the structural alternatives for relations between presence and absence of an attended feature. It is sensitivity to relationships among these structures and their similarity to the biconditional modal category that is revealed by the experiment. These structural alternatives are specified by the Boolean operators, which express the similarity between the categories they define and the biconditional category **E**. To our knowledge, no one has previously suggested that there is any relationship between Boolean operators and similarity assessment, let alone provided a theory that empirically predicts the similarity judgments of human subjects. It is on this basis that we claim that the logical connectives are reducible to modal similarity.

MST and inference

We have now explained the basis for the first premise of our master argument that the logical connectives define modal similarity categories. Our second premise is that inference is reducible to conditional categorization. Given Vigo's empirical demonstration that humans seem to be making modal similarity comparisons using **E** as a template, we suggest the following theoretical analysis. Inference is a process of conditional categorization and modal state elimination. The Deduction Theorem, $\Gamma \cup \{\mathcal{A}\} \vdash \mathcal{B} \Leftrightarrow \Gamma \vdash (\mathcal{A} \supset \mathcal{B})$, establishes the equivalence of the entailment relationship to a conditional. We propose to generalize beyond sentences to modal states as follows. Consider an entailment relationship where the presence of feature φ in an object p entails the presence of φ in another object q; i.e., if $p_{\mathbf{P}}(\varphi)$ then $q_{\mathbf{P}}(\varphi)$. Three of the four possible combinations of modal states of p and q with respect to φ are members of this conditional category (present in both, absent in both, and absent in p but present in q). The recognition that such a conditional category is in place can support various expectations for unobserved modal states. Thus, for instance, observing the presence of φ in p rules out two of the three instance pairs, leaving as the only possibility that φ is also present in q. Likewise, noticing the absence of φ in q (i.e., $q_{\mathbf{A}}(\varphi)$) reduces the modal state possibilities to just one. However, noticing that $q_{\mathbf{P}}(\varphi)$ leaves two possibilities, with no immediate resolution for the state of p with respect to φ. Hence, *modus ponens* and *modus tollens* are subsumed under a single general form of explanation which does not depend on explicitly encoding a rule. This is not to say that any organism capable of *modus ponens* will be automatically capable of *modus tollens*, or that it will find the latter kind of inference as easy as the former. This is because MST treats the presence and absence of features independently, with primitive similarity with respect to the presence of a feature being greater (more salient) than similarity with respect to its absence, and conversely for primitive dissimilarity judgments at the featural level. (This asymmetry between presence and absence of φ is represented by the parameters in the modal similarity measure; see Box 1.)

Our claim that modal similarity categorization is a prelinguistic capacity is supported by the recognition that MST is about relations among perceivable features of stimuli independent of their symbolic or linguistic representation. Vigo's "hat" experiment shows implicit sensitivity to logical categories constructed from such features. Modal similarity capacities are within the range of many nonhuman animals since they rely on the same categorization, conceptualization, and attentional capacities that are necessary for many other behaviors. There is, of course, a substantial literature on animal concepts, pro and con, but we believe that the preponderance of the evidence supports the attribution of conceptual abilities to languageless animals (Herrnstein and Loveland 1964; Herrnstein 1990; Allen 1999; Stephan 1999). For our present purposes, however, it is possible to sidestep this controversy by pointing out that attention and categorization are essential for all but the most primitive forms of conditioning. Animals learn and generalize reward relations among stimuli by learning which features to attend to, and they do so without the benefit of language.

This completes, although by no means settles, our "master" argument for the claim that inference is a prelinguistic process. Although the conclusion needs further support, specifically in the form of direct empirical evidence that animals really are capable of modal similarity judgments, we believe that, with the help of MST, we have outlined a framework for hope, contra Bermúdez (2003), that an inference-based conception of rationality can be applied at the nonlinguistic level.

Transitive inference and MST

The argument that inference is a species of modal similarity categorization has the status of an empirical theory which identifies a class of models to be developed for specific cases of inference. However, along with Thomas (2002), we think that experimental tests of conditional reasoning in nonhuman animals have generally failed to distinguish it adequately from conjunctive reasoning because they focus only on the case where both objects possess the relevant feature (but see Templeton 1998 for a suggestive study in which birds learned by observing other birds experience the absence of a food reward in a foraging choice task). Because transitive inference provides a much richer domain of examples, our next task is to attempt an analysis of transitive inference in the terms provided by MST. In its most generic form, a transitive inference has the following conditionalized structure, where $\mathcal{R}$ is a binary relation:

$$[\mathcal{R}(a_1, a_2) \wedge \cdots \wedge \mathcal{R}(a_{n-1}, a_n)] \supset \mathcal{R}(a_i, a_j) \quad \text{where } i,j \in \{1, \ldots, n\},\ i<j \tag{1}$$

For instance, if $\mathcal{R}$ is the greater-than relation, the following provides an example of transitive inference: $[(a_1 > a_2) \wedge (a_2 > a_3)] \supset (a_1 > a_3)$. Of course, not all transitive arguments of the form shown in (1) are deductively valid in nature. Some, in fact, are only inductively strong, in that the transitivity of $\mathcal{R}$ may pertain only to a subset of stimuli encountered or to be encountered. However, the same general structure of the inference applies.

In transitive inference experiments, nonhuman animals show themselves to be sensitive to the relation of reward presence to one member of each pair, and reward absence to the other. To keep our presentation relatively simple, we will analyze the simplest version of the experiment, which uses just two training pairs ($a+$, $b-$) and ($b+$, $c-$), even though the more complicated 5-element experiment may be required to rule out simple associative explanations. If we denote the property of a reward being present in respect to a stimulus α as $\alpha(r) = \mathbf{P}$, abbreviated as $\alpha_{\mathbf{P}}(r)$, and the property of a reward being absent as $\alpha(r) = \mathbf{A}$, or $\alpha_{\mathbf{A}}(r)$, we can represent the putative inference as:

$$\left[\widehat{\mathcal{R}}(a_{\mathbf{P}}(r), b_{\mathbf{A}}(r)) \wedge \widehat{\mathcal{R}}(b_{\mathbf{P}}(r), c_{\mathbf{A}}(r))\right] \supset \widehat{\mathcal{R}}(a_{\mathbf{P}}(r), c_{\mathbf{A}}(r)) \tag{2}$$

In (2), the relation $\widehat{\mathcal{R}}$ is not simply the relation between stimuli that is $\mathcal{R}$, but a relation between the stimuli and the property of a reward being either present, $_{\mathbf{P}}(r)$, or absent, $_{\mathbf{A}}(r)$. That is, the relation $\widehat{\mathcal{R}}$ is the higher order relation between the presence of a reward in one of the stimuli and the absence of a reward in the other.

The formulation in (2) enables us to focus on the key issue of whether transitive inference rather than some merely associative process underlies nonhuman animal decision making in these experiments. If such a relation as $\widehat{\mathcal{R}}$ is to be recognized by the animal, then the category of reward versus the category of nonreward must be formed as a function of the learning trials. For, suppose that these categories were not formed: how then would the animal know that the novel pair of stimuli is anything more than another pair like those encountered in the learning trials before the learning took place? So, the nonhuman animal must form categories of pairs of stimuli related via the presence of a reward for one and the absence of a reward for the other. For example, the category corresponding to the cases where one element of the pair of stimuli (a, b) is always rewarded and the other is never rewarded. These relational categories are formed by the process of making similarity judgments between the stimuli pairs presented on each trial. Once the categories for ($a+$, $b-$) and ($b+$, $c-$), respectively, have been formed, the nonhuman animal may be able to recognize higher order patterns through the process of similarity assessment on these relational categories.

In the experiment which uses only two training pairs, because animals are always given a reward when a is chosen and never given a reward when c is chosen it is reasonable to interpret this behavior as being merely associative in nature. This is because the choice of a may be due to the fact that the animal remembers that $a+$ and $c-$ or, in other words, the animal associates a reward with stimulus a and associates a lack of reward with stimulus c regardless of the relations that these two stimuli may bear to other stimuli. Thus, the subjects in such an experiment might be insensitive to the property that makes the relation $\widehat{\mathcal{R}}$ transitive in the first place. This is the chain property $\mathfrak{T}$ that for every pair in a set of pairs of stimuli the unrewarded member of the previous pair is the rewarded member of exactly one (distinct) member of another pair. The chain property is the key property underlying transitive inference.

Genuine transitive inference occurs when recognition of the chain meta-property $\mathfrak{T}$ between the possible arguments of $\widehat{\mathcal{R}}$ (which makes $\widehat{\mathcal{R}}$ transitive) is combined with the logical component of (2); namely, the higher order logical relation $\mathfrak{L}$ whose general form is given by:

$$\left[\widehat{\mathcal{R}}(a_{1\mathbf{P}}(r), a_{2\mathbf{A}}(r)) \wedge \cdots \wedge \widehat{\mathcal{R}}(a_{n-1\mathbf{P}}(r), a_{n\mathbf{A}}(r))\right] \supset \widehat{\mathcal{R}}(a_{i\mathbf{P}}(r), a_{j\mathbf{A}}(r)) \quad \text{where } i,j \in \{1, \ldots, n\},\ i<j \tag{3}$$

The rule expressed in (3) describes a relational category. Once again, it corresponds to the category consisting of those stimuli pairs where one stimulus in each pair has the feature of the presence of a reward while the other stimulus in each pair does not have this feature, and the higher order ordinal feature that the unrewarded stimulus in the previous pair is the rewarded stimulus in the following pair. Thus the higher order category described in (3) exemplifies transitivity. Another way of understanding this category is as the category consisting of all the subcategories corresponding to the relation $\widehat{\mathcal{R}}$, or in other terms, the set of all instances of the relation $\widehat{\mathcal{R}}$. As mentioned, along with property $\mathfrak{T}$, this higher order concept defines the transitive property. But how does the transitive property relate to the process of inference?

The answer to this question lies in MST. Two common ways of representing categories in theories of human category learning have been (1) by logical rules and (2) by exemplars organized by a similarity measure. MST was devised with the purpose of bridging these two representational paradigms. Recall that the key to the theory lies in the meaning of the connectives. In MST the connectives are in fact degrees of similarity in disguise, produced by the form of similarity assessment called modal similarity. Hence, according to MST, the conditional relationship shown in (3), $\mathfrak{L}$, expresses a degree of similarity in respect to the presence or absence of some attended feature φ, namely the relation $\widehat{\mathcal{R}}$. Hence, the entire inferential process can be characterized subsymbolically by the process of similarity assessment and categorization.

Since the inferential component $\mathfrak{L}$ of transitive inference is a modal similarity comparison between a pair of stimuli in respect to some feature or property, it is, at its core, a subsymbolic process. Because of its language independence, the process is best understood through a more general ability to categorize. And because it provides for mathematically precise models of how primitive similarity judgments can be combined into higher-order modal similarity judgment, we regard the approach as providing a promising tool for analyzing and explaining empirical results that are not easily captured within a traditional associationist framework, such as the ability of monkeys to rapidly link three 5-element lists into a 15-element serial list (Treichler 2007).

To recapitulate, we have argued in this section that for animals to engage in transitive inference, they must: (1) recognize the relation $\widehat{\mathcal{R}}$ of the absence and/or presence of a reward that exists between stimuli in pairs of learning trials; (2) recognize the chain property $\mathfrak{T}$ of the pairs of stimuli (i.e., the property that the rewarded element of the next pair is the unrewarded element of the previous pair); and (3) recognize the conditional relation $\mathfrak{L}$ that exists between the consequent of the transitive formula and its antecedent in terms of the absence and/or presence of the property $\widehat{\mathcal{R}}$ in the novel pair. What this means is that nonhuman animals should recognize the likelihood that if all the pairs of stimuli in the chain have property $\widehat{\mathcal{R}}$ then the novel pair will also have property $\widehat{\mathcal{R}}$. We then showed how each of the processes in steps 1, 2, and 3 are in fact rooted in similarity assessment, discrimination, and categorization, thereby making transitive inference a subsymbolic perceptual process. This, added to the hypothesis (to be supported in the next section) that nonhuman animals possess a robust ability for categorization well within the demands of 1, 2, and 3, indicates that nonhuman animals are capable of at least some forms of inference.

Negation without language

Thus far, we have argued that the categorical basis of inference makes plausible the idea that animals are capable of some forms of inference. But is it possible to understand negation without language? Negation is widely regarded as crucial to logical reasoning, and inexpressible without the full semantic apparatus of language. In our discussion of transitive inference, it was not necessary to deal with negation since it does not play a direct role in its formulation. But there is a broad sense of negation that we must define in order to complete our picture of how logical rules are merely expressions of degrees of modal similarity. In MST, logical negation plays the role of a simplification function. To explain, consider the category $\mathbb{C} = \langle \{e_1 \cdots e_n\}, \mathbb{S} \rangle$ defined by its exemplars $e_1 \cdots e_n$ and a similarity measure $\mathbb{S}$ between them. Suppose that an agent wishes to exclude exemplars from the category $\mathbb{C}$ in respect to a particular feature φ. To say that a certain subset of the exemplar set does not have a certain feature φ is to say that for some exemplars, φ is absent. In some cases, it is much easier to exclude elements from a category (to specify which do not have the feature) than to specify the elements with the feature. That negation is implied in modal categories by way of exclusion can be seen by the fact that each logical connective indirectly excludes possible members of the category: For instance, the conjunction excludes three possible instance pairs, while disjunction excludes but one. This suggests that the cognitive significance of negation is the exclusion of elements from categories.

While humans may use logical negation as a means to express and understand categorical exclusions, nonhuman animals may possess limited cognitive capabilities in this regard. For instance, the number of modal alternatives that the nonhuman animal can discriminate may be considerably lower than the number of modal alternatives that humans can discriminate. But in a dog's world, this smaller number may be sufficient. In the most basic case of *modus tollendo ponens*, only two alternatives have to be discriminated, while in the Classical telling of the "rule of dogs" it requires three. These modal alternatives prompt the animal to make a choice based strictly on similarities and the eventual exclusion of the road that fails the modal similarity test. With their capacity for language, humans have a convenient way of expressing these exclusions. Without language, nonhuman animals have a categorical way of processing the same problem. The different capacities in memory storage and processing speed will, however, make a difference in the complexity of the eventual exclusions.

This perspective also helps us understand why Watson (2005) argues that MTP is of particular interest for the study of animal reasoning. He writes that MTP is important because "this syllogistic frame (vs. modus ponens or modus tollendo tollens) readily lends itself to an operational distinction from what one would expect or predict from an associative learning perspective." Watson's point seems to be that apparent instances of reasoning in accordance with *modus ponens* and *modus tollens* may not be empirically distinguishable from associative mechanisms operating on contingently established connections between antecedent and consequent of the conditional premise, whereas MTP essentially involves a retrospective reference to the particular state of affairs establishing the disjunction—in other words, one cannot eliminate the possibilities without having a representation of the conditions establishing those possibilities.

From ethological to logical

How do the capacities of nonhuman animals bear on the formally rich reasoning abilities of humans? We believe the answer lies in reconceptualizing our own reasoning capacities, along lines we have already suggested in this paper. An independent line of evidence comes from Landy and Goldstone (2007) who demonstrate experimentally that algebraic competence is surprisingly fragile in the face of quite small differences in the way formulas are written, and that reasonably competent algebraic and logical reasoners appear to exploit what should be semantically irrelevant properties of formulas, such as white space, to support their competency.

We believe that the perceptual capacities underlying formal reasoning are grounded in the categorization behavior presented in this paper and described in MST. In this paper we have argued that inference can be understood as a process of making conditional similarity judgments, where new inferences proceed by recognizing the modal similarity between the premises of the new instance and the premises of familiar instances. Humans, by virtue of culture, education, and biology have a particular capacity for perceiving and manipulating a wide range of finely variegated symbolic structures, and are thus able to construct very sophisticated chains of formal reasoning. The actual processes supporting such reasoning in humans are nonetheless subsymbolic, detecting and exploiting modal similarities inherent in public notational systems. Insofar as formally irrelevant features like spacing do in fact provide regularities in the environment that can be exploited to assist perceptual systems to recognize important similarities, then modal relationships among such subsymbolic features will be detected and exploited automatically by the perceptual systems of reasoners.

Precisely how these perceptual mechanisms operate remains to be shown. But we believe that MST holds out a very real prospect of a unified account of categorization and inference. This account would cover symbolic reasoning, natural-language syllogisms, and generalization from perceptions of non-symbolic parts of the world. To reserve terms like "reasoning" and "inference" for just the first two of these would be an acceptable terminological move, but one which obscures, we believe, the essential underlying similarity among all three.

Conclusion

We have argued that the process of drawing inferences is not necessarily language driven and that at the heart of reasoning as well as associative learning lies the more fundamental processes of similarity assessment, discrimination, and categorization. We conclude from this that there is no reason in principle that languageless animals should be incapable of some forms of reasoning, including transitive inference. However, direct tests of modal similarity capabilities in nonhuman animals are needed.

Acknowledgments We wish to thank Junko Obayashi and members of the Indiana University Biology Studies Reading Group for their suggestions and encouragement while preparing this manuscript. Elisabeth Lloyd and Robert Treichler provided helpful written comments. Both of the authors benefitted from the comments and questions following presentation of these ideas to the IU Logic Seminar, and C.A. also acknowledges the useful questions and comments from audiences at the Universities of Bonn, Memphis, Pennsylvania, and Alabama. We thank Holger Lyre at University of Bonn and the journal referees for encouragement and comments. Finally, C.A. deeply regrets that he never met Brendan McGonigle in person nor did Professor McG. live to see a long-promised draft of this paper due to his sudden death at the end of 2007. His energetic, enthusiastic, and challenging emails on the topic of transitive inference will be missed, but his seminal work remains.

References

Allen C (1999) Animal concepts revisited. Erkenntnis 51:33–40
Allen C (2006) Transitive inferences in animals: reasoning or conditioned associations? In: Hurley S, Nudds M (eds) Rational animals? Oxford University Press, Oxford, pp 173–184

Battista MT, Clements DH (1991) Using spatial imagery in geometric reasoning. Arith Teacher 39:18–21

Bermúdez J (2003) Thinking without words. Oxford University Press, New York

Bond AB, Kamil AC, Balda RP (2003) Social complexity and transitive inference in corvids. Anim Behav 65:479–487

Bower GH, Clark MC (1969) Narrative stories as mediators for serial learning. Psychon Sci 14:181–182

Brunswick E (1955) "Ratiomorphic" models of perception and thinking. Acta Psychol 11:108–109

Bryant P, Trabasso T (1971) Transitive inferences and memory in young children. Nature 232:456–458

Chomsky N (1980) Rules and representations. Basil Blackwell, Oxford

Clark A (1998) Magic words: how language augments human computation. In: Carruthers P, Boucher J (eds) Language and thought: interdisciplinary themes. Cambridge University Press, Cambridge, pp 162–183

Dictionary (2008) New Oxford American dictionary. Accessed via Dictionary.app for Mac OS X, Cupertino, Apple Inc.

Dusek A, Eichenbaum H (1997) The hippocampus and memory for orderly stimulus relations. Proc Natl Acad Sci 94:7109–7114

Fersen L von, Wynne CDL, Delius JD, Staddon JER (1991) Transitive inference formation in pigeons. J Exp Psychol Anim B 17:334–341

Fodor J, Pylyshyn Z (1988) Connectionism and cognitive architecture. Cognition 28:3–71

Grosenick L, Clement TS, Fernald RD (2007) Fish can infer social rank by observation alone. Nature 445:429–432

Herrnstein RJ (1990) Levels of stimulus control: a functional approach. Cognition 37:133–166

Herrnstein RJ, Loveland DH (1964) Complex visual concepts in the pigeon. Science 146:549–551

Hogue ME, Beaugrand JP, Lagüe PC (1996) Coherent use of information by hens observing their former dominant defeating or being defeated by a stranger. Behav Process 38:241–252

Hurley S, Nudds M (2006) Rational animals? Oxford University Press, Oxford

Inhelder B, Piaget J (1958) The growth of logical thinking from childhood to adolescence. Basic Books, New York

Köhler W (1929) Gestalt psychology. Horace Liveright, New York

Landy D, Goldstone RL (2007) Formal notations are diagrams: evidence from a production task. Mem Cognit 35:2033–2040

McGonigle B, Chalmers M (1992) Monkeys are rational! Q J Exp Psychol 45B:198–228

McGonigle B, Chalmers M (2002) The growth of cognitive structure in monkeys and men. In: Fountain SB, Bunsey MD, Danks JH, McBeath MK (eds) Animal cognition and sequential behaviour: behavioral, biological and computational perspectives. Kluwer Academic, Boston, pp 269–314

Pernet C, Franceries X, Basan S, Cassol E, Démonet JF, Celsisa P. (2004) Anatomy and time course of discrimination and categorization processes in vision: an fMRI study. Neuroimage 22:1563–1577

Piaget J (1971) Biology and knowledge. University of Chicago Press, Chicago

Rumbaugh DM, Washburn DA, Hillix WA (1996) Respondents, operants, and emergents: toward an integrated perspective on behavior. In: Pribram K, King J (eds) Learning as a self-organizing process. Erlbaum, Hillsdale, pp 57–73

Seyfarth RM, Cheney DL (2002) The structure of social knowledge in monkeys. In: Bekoff M, Allen C, Burghardt GM (eds) The cognitive animal: empirical and theoretical perspectives on animal cognition. MIT Press, Cambridge, pp 379–384

Siegal M, Varley R, Want SC (2001) Mind over grammar: reasoning in aphasia and development. Trends Cogn Sci 5:296–301

Sorabji R (1993) Animal minds and human morals: the origins of Western debate. Cornell University Press, Ithaca

Stephan A (1999) Are animals capable of concepts? Erkenntnis 51:583–596

Templeton J (1998) Learning from others' mistakes: a paradox revisited. Anim Behav 55:79–85

Thomas RK (2002) Conditional discrimination learning and conditional reasoning by nonhuman animals. Updated version of Thomas RK (1991) Misuse of conditional reasoning in animal research with special reference to evolution of language. Southern Society for Philosophy and Psychology, Atlanta, GA. http://rkthomas.myweb.uga.edu/Conditional.htm. Accessed 07 January 2008

Treichler FR (2007) Monkeys making a list: checking it twice? In: Washburn DA (ed) Primate perspectives on behavior and cognition. American Psychological Association, Washington DC, pp 143–160

Treichler FR, Raghanti MA, Van Tilburg DN (2003) Linking of serially ordered lists by Macaque monkeys (*Macaca mulatta*): list position influences. J Exp Psychol Anim B 29:211–221

Vigo R (2008) Modal similarity. J Exp Theor Artif Intell (in press)

Watanabe S, Huber L (2006) Animal logics: decisions in the absence of human language. Anim Cogn 9:235–245

Watson JS (2005) Causal logic and the intentional stance: retrospection and logical thought. Interdisciplines. http://www.interdisciplines.org/causality/papers/7/4/1#_4. Accessed 27 January 2008

Watson JS, Gergely G, Csanyi V, Topal J, Gacsi M, Sarkozi Z (2001) Distinguishing logic from association in the solution of an invisible displacement task by children (*Homo sapiens*) and dogs (*Canis familiaris*): using negation of disjunction. J Comp Psychol 115:219–226

Zentall TR (2001) The case for a cognitive approach to animal learning and behavior. Behav Process 54:65–78