How to Reason without Words: Inference as Categorization

Ronaldo Vigo
rvigo@indiana.edu

Colin Allen
colallen@indiana.edu

Cognitive Science Program
Indiana University, Bloomington

Draft for comments. Email welcome.
Please do not circulate or cite without permission.
References are incomplete.
In prep. for *Cognitive Processing* 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 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 categorization; animal cognition; animal learning; animal reasoning

1 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. It is reported 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 three-choice searching task. Dogs in this experiment 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 4-6 yr 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" (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 & Nudds 2006, and the review by Watanabe & Huber 2006). McGonigle & Chalmers (2002), who argue that monkeys *are* rational (McGonigle & Chalmers 1992), report, "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 & Pylyshyn 1988; Piaget 1971)." 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 depending essentially on linguistically represented premises and conclusions. This conception makes it hard to credit animals with reasoning capacities. As José 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 to the subsymbolic perceptual processes underlying similarity assessment, discrimination, and categorization. The framework is provided by the Modal Similarity Theory (MST) of Vigo (2007), which we outline in section 5. MST introduces the concept of modal similarity and shows how one may reduce the propositional connectives to 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 don't 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.

2 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' is "the power of the mind to think, understand, and form judgments by a process of logic" (Oxford American Dictionary 2005). Neither of these definitions mention 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 another sentence. Semantic conceptions of inference are grounded in semantic relationships between premises and conclusion: namely, those of validity (truth preservation) and inductive strength (probable truth-enhancing). 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.

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 non-human 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 the following section we argue that inference is not necessarily language-dependent, or even symbol-dependent. Instead, we argue, inference is a sub-symbolic process grounded in the cognitive meaning of the logical connectives themselves. But what is the cognitive meaning of the connectives? We answer this question by first asking the question, what is their utility? Or, in stronger words, what are they good for?

Intuitively, the connectives are relations that indicate which of a set of alternatives are possible members of a category. Take, for instance, material implication. The statement $p \supset q$ designates the elimination of one (p and not-q) of the four possible states that the two entities or features (not necessarily propositions), p and q, may be in. In the most general sense, the states of interest concern the presence or absence of discernable features. Thus, in the theory of inference that will be presented under section five, instead of propositions, we assume a far more general domain of application for the logical operators, where the objects of inference will be generalized to *stimuli*, and in particular, to relations among *features of stimuli.*

3 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 & 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 2006) and, in section 8 we will describe results that suggest that the notations themselves can become the objects of modal categorization processes.

But language is no more necessary for logical reasoning than possession of a word for pain is necessary to feel pain. Case studies of aphasic patients provide evidence in support of this view. 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). Hence, aphasia researchers tacitly suggest that perhaps the inferential capacity of humans is exercised at a subsymbolic 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 linguistic 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 (Reeves 1985). 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. If this doubt is correct, then it would seem 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 we would not be able to learn a language in the first place. That is, we need 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 we would only have within our 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 pre-linguistic 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 pre-linguistic process.
5. Hence, inference is a pre-linguistic process.

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. Neither 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 (TI) in animals, which anchor our argument that nonhuman animals are capable of some form of inferential reasoning. Controversy about TI 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 & Huber 2006.) MST provides the sufficiently rigorous and empirically tractable conceptual underpinning sufficient to reorient the millennia-old debate about animal reasoning.

4 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 2007) and rodents (Dusek & Eichenbaum 1997), birds including pigeons (Zentall 2001), jays, (Bond, Kamil & Balda 2003), and chickens (Beaugrand, Hogue & Lagüe 1997), and Siamese fighting fish (Grosenick, Clement, & Fernald 2007) all seem capable of responding to novel pairings as one would predict using TI. Such capacities may contribute to the fitness of individuals fitness of the individuals living in dominance hierarchies (Seyfarth & Cheney 2002; Allen 2006; Grosenick et al. 2007).

The experimental basis for the claim that nonhuman animals engage in TI originates from an experiment originally conducted by Piaget to study the cognitive development of children (Piaget 1954). 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-labelled) stimuli: $(a+, b-)$ and $(b+, c-)$. Here the plus sign stands for positive reinforcement if that stimulus is selected by the subject, and the minus sign stands for no reinforcement. The letters a through c are 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 TI. Hence, the rewarded stimulus is never uniformly on the left or the right, for instance. When trained to criterion via operant conditioning, pigeons and rats are highly likely to select stimulus a when presented with the novel pair (a, c). However, this particular result admits of a very simple associative explanation. In 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 industry standard for laboratory investigations of transitive inference. In the 5-element procedure, due to Bryant & 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 and 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 & Chalmers 2001; Treichler, Raghanti, & Tilburg 2007).

What exactly allows animals to induce new relationships among familiar stimuli? Treichler et al. invoke the category of "emergents", following Rumbaugh, Washburn & Hillix (1996), who introduced this notion as a new category of acquired behavior, to supplement the standard behaviorist's toolkit of respondents and operants. McGonigle & Chalmers (2002) invoke the idea of "private codes" to anchor "relational primitives" which can order stimuli into the sequences required for TI — e.g. *bigness* as an anchor for size relations. For reasons that would take us too far afield to explain here, we believe that neither of these suggestions has the degree of specificity offered by MST (although we believe that McGonigle & Chalmers are on the right track). So, without further ado, we introduce MST next.

5 Modal Similarity Theory

MST was developed by Vigo (2007) to show the relationship between the propositional connectives, similarity assessment, and categorization. A modal category is a relational category consisting of pairs of items defined by the absence or presence of attended features. 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. The central claim of MST is that the propositional connectives as used by human agents in language are expressions of degrees of similarity between the modal states (presence or absence of features) of pairs of stimuli. MST thus builds a foundational bridge between similarity-based categorization and rule-based categorization processes. It provides an understanding of the role of the logical connectives in rule-based processing accounts of categorization, based upon a similarity measure that is more distributive and relational in nature than current representations of similarity.

Vigo's axioms for MST (see box 1) enable a precise statement of the role that perceptual features play in similarity and dissimilarity judgments. Modal similarity is a higher order similarity relation between the modal categories defined by the sixteen logical connectives and the standard modal category (**E**) defined by the logical equivalence connective (i.e., the biconditional), with respect to a feature φ. **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. 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, a.k.a. 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 of the feature φ, whereas $\mathbf{E}'$ consists of those pairs whose elements are dissimilar in respect of φ. The categories formed of these pairs are thus maximally dissimilar from each other.

Using the componential notion of similarity captured in MST, it is possible to define a structured measure of the degree of similarity between **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 the feature φ within a modal category, and the relative empha-

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 (2007). 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, and φ is a property or feature.

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, $sim_p(p_{\mathsf{P}}, q_{\mathsf{P}})$, $sim_p(p_{\mathsf{A}}, q_{\mathsf{A}})$, $dis_p(p_{\mathsf{P}}, q_{\mathsf{A}})$, and $dis_p(p_{\mathsf{A}}, q_{\mathsf{P}})$.

   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:
   $\mathrm{Sim}_M(\mathbf{E}, \Psi, \varphi) = f(sim_p(p_{\mathsf{P}}, q_{\mathsf{P}}), sim_p(p_{\mathsf{A}}, q_{\mathsf{A}}), dis_p(p_{\mathsf{P}}, q_{\mathsf{A}}), dis_p(p_{\mathsf{A}}, q_{\mathsf{P}}))$

3. Order Property. The primitive similarity and dissimilarity measures are ordered as follows: $sim_p(p_{\mathsf{P}}, q_{\mathsf{P}}) > sim_p(p_{\mathsf{A}}, q_{\mathsf{A}})$ and $dis_p(p_{\mathsf{A}}, q_{\mathsf{P}}) > dis_p(p_{\mathsf{P}}, q_{\mathsf{A}})$

4. Summation Postulate. Partial similarity adds up to 1 and partial dissimilarity adds up to 1: $sim_p\,(p_{\mathsf{P}}, q_{\mathsf{P}},\ \varphi) + sim_p(p_{\mathsf{A}}, q_{\mathsf{A}}, \varphi) = 1$ $\qquad dis_p(p_{\mathsf{P}}, q_{\mathsf{A}}, \varphi) + dis_p(p_{\mathsf{A}}, q_{\mathsf{P}}, \varphi) = 1$

5. Maximal and Minimal Modal Similarity Property. Maximal and minimal similarity values for the relational measure are given by:
   $Max(Sim_M(\mathbf{E}, \Psi, \varphi)) = Sim_M(\mathbf{E}, \mathbf{E}, \varphi)$ and $Min(Sim_M(\mathbf{E}, \Psi, \varphi)) = Sim_M\big(\mathbf{E}, \overline{\mathbf{E}}, \varphi\big)$
   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: $Sim_M(\mathbf{E}, \varnothing, \varphi)) = 0$.

7. Contextual Reversal Property. $sim_p(p_{\mathsf{P}}, q_{\mathsf{P}}) = \alpha_1$ and $dis_p(p_{\mathsf{P}}, q_{\mathsf{A}}) = \alpha_2$, if $\mathbf{E} \notin \Psi$; $sim_p(p_{\mathsf{P}}, q_{\mathsf{P}}) = \alpha_2$ and $dis_p(p_{\mathsf{P}}, q_{\mathsf{A}}) = \alpha_1$, if $\mathbf{E} \in \Psi$.

8. Existence Property II (of metafunction $\widetilde{\mathfrak{M}}$). For any pair of objects p and q and for any property or feature φ there exists a metafunction $\widetilde{\mathfrak{M}}$ such that:

   $$\widetilde{\mathfrak{M}}_\Psi(\,(p(\varphi), q(\varphi))) = \begin{array}{l} 1 \text{ if } (p(\varphi), q(\varphi)) \text{ is present in } \Psi \\ 0 \text{ if } (p(\varphi), q(\varphi)) \text{ is absent in } \Psi \end{array}$$

Please note that postulate 6 simply states that the empty modal relation $\varnothing$ expresses neither a degree of similarity or 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. On the other hand, 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

$$Sim_M(\mathbf{E}, \Psi, \varphi) \;=\; \begin{array}{ll} \beta\Big[\alpha_1\widetilde{\mathfrak{M}}_\Psi(\mathbf{E}(1)) + (1-\alpha_1)\,\widetilde{\mathfrak{M}}_\Psi(\mathbf{E}(2))\Big] - & \\ (1-\beta)\Big[\alpha_2\widetilde{\mathfrak{M}}_\Psi\Big(\mathbf{E}'(1)\Big) + (1-\alpha_2)\,\widetilde{\mathfrak{M}}_\Psi\Big(\mathbf{E}'(2)\Big)\Big] & \text{if } \widehat{\Psi} \in \{8, \ldots, 15\} \\ \beta\Big[\alpha_2\widetilde{\mathfrak{M}}_\Psi(\mathbf{E}(1)) + (1-\alpha_2)\,\widetilde{\mathfrak{M}}_\Psi(\mathbf{E}(2))\Big] - & \\ (1-\beta)\,\Big[\alpha_1\widetilde{\mathfrak{M}}_\Psi\Big(\mathbf{E}'(1)\Big) + (1-\alpha_1)\,\widetilde{\mathfrak{M}}_\Psi\Big(\mathbf{E}'(2)\Big)\Big] & \text{if } \widehat{\Psi} \in \{0, \ldots, 7\} \end{array}$$

sis 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 box 2). Subjects in Vigo's experiment were presented with iconic stimuli representing modal categories (see figure 5.1). They were told nothing about the stimuli except that they represented the kinds of married couples found 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 either shown wearing a hat, or hatless. Subjects in one group were presented with pictorial stimuli representing the modal categories for biconditional (equivalence), 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. All subjects were asked to rate the similarity on a scale of 1 to 10 of each "culture" to the "culture" corresponding to **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. Thus, human subjects were sensitive to modal similarity even though they lacked explicit instruction in logic, or linguistic labels for the logical connectives. On the basis of these results, Vigo suggests that the Boolean operators (logical connectives) should be thought of as expressing degrees of modal similarity.

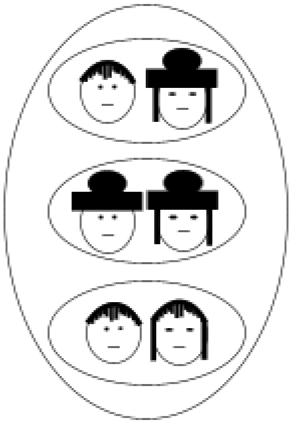

Figure 5.1: Stimulus used by Vigo (2008) for the "conditional culture" category.

Box 2: Modal similarity measure predictions when $\beta_1 = .55$

$\mathfrak{M}_\mu(p_\mathsf{A}, q_\mathsf{A})$	$\mathfrak{M}_\mu(p_\mathsf{A}, q_\mathsf{A})$	$\mathfrak{M}_\mu(p_\mathsf{A}, q_\mathsf{A})$	$\mathfrak{M}_\mu(p_\mathsf{A}, q_\mathsf{A})$	connective	$Sim_M(p, q, \varphi)$
1	1	0	0	$\equiv$	.55
1	0	0	0	$\wedge$	.38
1	1	1	0		.37
0	1	0	0	$\neg(\vee)$	.33
1	1	0	1	$\supset$	.28
1	0	1	0		.21
0	1	0	1		.20
1	0	0	1		.12
1	1	1	1		.10
0	1	1	0		.02
0	0	0	0		0
1	0	1	1	$\vee$	-.07
0	1	1	1	$\neg(\wedge)$	-.12
0	0	0	1		-.14
0	0	1	0	$\neg(\supset)$	-.32
0	0	1	1	$\neg(\equiv)$	-.45

6 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 Equivalence 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 possesses φ, then so does q. 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, reduces the modal state possibilities to just one. However, noticing the presence of φ in q 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 presence and absence of features independently, with primitive similarity with respect to presence of a feature always being greater (more salient) than similarity with respect to absence, and conversely for primitive dissimilarity judgments at the featural level.

Our claim that modal similarity categorization is a pre-linguistic 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 & Loveland 1964; Herrnstein 1990; Allen 1995; 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 pre-linguistic 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, that an inference-based conception of rationality can be applied at the nonlinguistic level.

6.1 Transitive Inference and Modal Similarity Theory

Note: this section is too long and formalistic. It will be condensed. Skip if you wish.

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 adequately to distinguish it 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 (TI) provides a much richer domain of examples, our next task is to attempt an analysis of TI 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:

$$[\mathcal{R}(a_1, a_2), \ldots \wedge \mathcal{R}(a_{n-1}, a_n)] \supset \mathcal{R}(a_i, a_j) \qquad \text{where } i, j \in \{1, \ldots, n\}, i < j \quad (6.1)$$

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 6.1 are deductively valid in nature. Some, in fact, are only inductively strong, in that the transitivity of $\mathcal{R}$ may pertain only to the subset of stimuli encountered so far. However, the same general structure of inference applies.

In TI 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. For the simplest experiment consisting of the training pairs $(a+, b-)$ and $(b+, c-)$, if we denote the property of a reward being present in respect to a stimulus α as $\alpha(r) = \mathbf{P}$ or $\alpha_{\mathbf{P}}(r)$ and the property of a reward not being present as $\alpha(r) = \mathbf{A}$ or $\alpha_{\mathbf{A}}(r)$, we can represent the putative inference as:

$$\left[\widehat{\mathcal{R}}(a_{\mathbf{P}}(r), b_{\mathbf{A}}(r)) \wedge \widehat{\mathcal{R}}(b_{\mathbf{P}}(r), c_{\mathbf{A}}(r))\right] \supset \widehat{\mathcal{R}}(a_{\mathbf{P}}(r), c_{\mathbf{A}}(r)) \quad (6.2)$$

In 6.2, 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 6.2 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 non-reward 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 particularly evident when one considers that the formulation in 6.2 makes clear the role of the relation $\widehat{\mathcal{R}}$ as a preference relation for the presence over the absence of the reward feature, but does not articulate the existence of a higher order ordinal relation between the arguments of the relations: namely, the chain property $\mathfrak{T}$ that for every pair in a sequence of pairs of stimuli the unrewarded member of the previous pair is the rewarded member of the following pair. The chain property is the key property underlying transitive inference. Clearly, to recognize this meta-property the nonhuman animal must recognize the similarity of one member of a pair to a member of another pair in the sequence. Hence, once again we see the importance of similarity assessment in the process of stimulus recognition. Moreover, by making similarity comparisons between the already acquired reward vs. non-reward categories, the category of the chain property may be formed. Without this category, the nonhuman animal would not be able to have a concept or category of transitivity.

In view of the above, not only does the simple three-item, two-pair experiment fail to provide proof that some species of nonhuman animals have a concept of transitivity, but more importantly, it does not prove that some species of nonhuman animals have reasoning capacities. Even when animals form (via similarity assessment) the category corresponding to relation $\widehat{\mathcal{R}}$ and the category corresponding to $\mathfrak{T}$, they are still not necessarily engaged in a process of transitive inference. That fact emerges when we recognize that in addition to the chain meta-property $\mathfrak{T}$ between the possible arguments of $\widehat{\mathcal{R}}$ (which makes $\widehat{\mathcal{R}}$ transitive) there exists the logical component of 6.2: namely, the higher order logical relation $\mathfrak{L}$ whose form is given by $(x_1 \wedge \cdots \wedge x_n) \supset y$.

In the five-element TI paradigm of Bryant & Trabasso (1971), subjects are required to recognize the $\mathfrak{L}$-type structure shown here:

$$\Big[\widehat{\mathcal{R}}(a_{1\mathbf{P}}(r), a_{2\mathbf{A}}(r)) \wedge \widehat{\mathcal{R}}(a_{2\mathbf{P}}(r), a_{3\mathbf{A}}(r)) \wedge \widehat{\mathcal{R}}(a_{3\mathbf{P}}(r), a_{4\mathbf{A}}(r)) \wedge \widehat{\mathcal{R}}(a_{4\mathbf{P}}(r), a_{5\mathbf{A}}(r))\Big] \supset \widehat{\mathcal{R}}(a_{2\mathbf{P}}(r), a_{4\mathbf{A}}(r)) \quad (6.3)$$

Now, 6.3 expresses a rule that 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 6.3 exemplifies transitivity. Another way of understanding this category is as the category consisting of all the sub-categories 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 Modal Similarity Theory (MST). Two common ways of representing categories in theories of human category learning have been (i) by logical rules and (ii) 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 6.3 $\mathfrak{L}$ 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 sub-symbolically by the process of similarity assessment and categorization.

Under MST, the logical glue (represented by $\mathfrak{L}$) for the transitive inference rule described in 4.7 above, is simply a modal similarity judgment on on the part of the cognitive agent. More specifically, modal similarity theory posits that cognitive agents would not be able to describe categories symbolically (in terms of rules) if it were not for the simple fact that the logical connectives, the glue of our rules, are nothing more than the representatives of the process of modal similarity assessment and discriminability under symbolic cover. Moreover, since the inferential component $\mathfrak{L}$ of TI is simply a modal similarity comparison between a pair of stimuli in respect to some feature or property, it is, at its core, a sub-symbolic process. Because of its language independence, the process is best understood through a more general ability to categorize.

Whether or not one accepts MST as plausible cognitive interpretation of transitive inference, it cannot be denied that in these experiments on transitive inference, nonhuman animals recognize the presence and/or absence of the reward and that it is this property along with the transitive property that is the key to performing a transitive inference. Contrary to the common view that nonhuman animals are incapable of higher order relations, they are very much capable and must be capable of higher order relations (whether they are also aware of the possession of such relations is both unlikely and inconsequential to our argument). The ability to recognize the presence or absence of a feature must lie at the most primitive and universal levels of cognitive capacity. Indeed, this is as basic as a discriminatory capacity gets. Without it, and without the ability to categorize (to the extent imposed by the limitations of memory storage), virtually no animal would be able to survive. As an example, consider that most animals must be able to discriminate and recognize not only food sources but their many possible locations.

To recapitulate, we argued in this section that transitive inference is more complex a process than it may at first seem. Furthermore, and in view of this fact, we considered what would plausibly have to be the case in order for animals to be engaging in a process of transitive inference. We concluded that 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 $\widehat{\mathcal{R}}$ 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 sub-symbolic 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.

7 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 theory logical negation plays the role of a simplification function. To explain, consider the category $\mathbb{C}=\langle\{e_1 \ldots e_n\}, \mathbb{S}\rangle$ defined by its exemplars $e_1 \ldots e_n$ 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. So $\varphi' \equiv \{e_i | e_i(\varphi) = \mathbf{A}\}$. 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 the ability 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 (versus 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.

8 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 & Goldstone (2006a,b) who demonstrate experimentally that algebraic knowledge 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 Modal Similarity Theory. 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.

9 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.

Acknowledgement. *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. Both authors benefitted from the comments and questions following presentation of our ideas to the IU Logic Seminar, and CA also acknowledges the useful questions and comments from audiences at the Universities of Bonn and Memphis.*

References (incomplete for now)

1. Allen, C. (1995). Animal concepts revisited. *Erkenntnis.*
2. Allen, C. (2006). Rational vs. Associative Processes. In S. Hurley & M. Nudds (eds.) *Rational Animals?* Oxford University Press; pp. 173-184.
3. Battista, M. T., & Clements, D. H. (1991). Using spatial imagery in geometric reasoning. Arithmetic Teacher, 39(3), 18-21.
4. Beaugrand, J.P., Hogue, M.E., and Lagüe, P.C. (1997). Utilisation cohérente de l'information obtenue par des poules domestiques assistant à la défaite de leur dominante contre une étrangère: s'agit-il d'inférence transitive? Processus cognitifs et ajustement écologique. Société Franaise pour l'Étude du Comportement Animal. Presses de l'Université Paul Sabatier, Toulouse, pp. 131-137.
5. Benard J, Satch S, Giurfa M (2006). Categorization of visual stimuli in the honeybee *Apis mellifera*. Animal Cognition 9: 257-270.
6. Bond, A.B., Kamil, A.C., and Balda, R.P. (2003). Social complexity and transitive inference in corvids. *Animal Behaviour* 65:479-487.
7. Bower, G. H., & Clark, M. C., (1969). Narrative stories as mediators for serial learning. Psychonomic Science, 14, 181-182.
8. Bourne, L.E. (1974). An inference model for conceptual rule learning. In R. Solso (Ed.), Theories in Cognitive Psychology (pp. 231-256). Washington: Erlbaum.
9. Brunswick E (1955) "Ratiomorphic" models of preception and thinking. Acta Psychologica 11: 108-109.

10. Candland, D. K., Feral Children and Clever Animals: Reflections on Human Nature, (1993) from Oxford University Press.
11. Church, A. (1956). Introduction to Mathematical Logic. Princeton: Princeton University Press.
12. Clark, A. (2006) ...
13. Dusek, A. and Eichenbaum, H. (1997). The hippocampus and memory for orderly stimulus relations. *Proceedings of the National Academy of the Sciences* USA 94:7109-7114.
14. Fersen, L. von, Wynne, C.D.L., Delius, J.D. and Staddon, J.E.R. (1991). Transitive inference formation in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes* 17: 334-341.
15. Garner, W. R. (1974). The Processing of Information and Structure. New York: Wiley.
16. Grosenick, L., Clement, T.S. & Fernald, R.D. (2007) Fish can infer social rank by observation alone. Nature 445, 429-432.
17. Herrnstein, R. J., & Loveland, , D. H. (1964). Complex visual concept in the pigeon. Science, 146, 549-551.
18. Herrnstein, R. J. (1990). Levels of stimulus control: A functional approach. Cognition, 37, 133-166.
19. Hurley, S. and Nudds, M. (2006). *Rational Animasl?* Oxford: Oxford University Press.
20. Kohler, Wolfgang. Gestalt Psychology. New York: Horace Liveright, 1929.
21. Landy, D. and Goldstone R. (multiple)
22. Lubow, R. E. (1974). High-order concept formation in the pigeon. Journal of the Experimental Analysis of Behavior, 21, 475-483.
23. McGonigle, B. and Chalmers, M. (1992). Monkeys are rational! *Quarterly Journal of Experimental Psychology* 45B: 198-228.

24. McGonigle, B. and M. Chalmers (2002). The growth of cognitive structure in monkeys and men. In *Animal cognition and sequential behaviour: Behavioral, biological and computational perspectives.* S. B. Fountain, M. D. Bunsey, J. H. Danks and M. K. McBeath (eds.). Boston, Kluwer Academic: 269-314.

25. Nosofsky, R. M., Clark, S. E., & Shin, H. J. (1989). Rules and exemplars in categorization, identification, and recognition. Journal of Experimental Psychology: Learning, Memory, & Cognition, 15(2), 282-304.

26. Rumbaugh, D. M., Washburn, D. A., & Hillix, W. A. (1996). Respondents, operants, and emergents: Toward an integrated perspective on behavior. In K. Pribram & J. King (Eds.), *Learning as a self-organizing process.* Hillsdale,NJ: Erlbaum; pp. 57-73.

27. Seyfarth, R.M., and Cheney, D.L. (2002). The structure of social knowledge in monkeys. In Bekoff, M., Allen, C., and Burghardt, G.M. (eds.) *The Cognitive Animal: Empirical and Theoretical Perspectives on Animal Cognition.* Cambridge, MA: The MIT Press, pp. 379-384.

28. Shepard, R., Hovland, C. L. & Jenkins, H. M. (1961). Learning and Memorization of Classifications. Psychological Monographs: General and Applied, 75(13), (pp. 1-42).

29. Siegal M, Varley R, Want SC., Mind over grammar: reasoning in aphasia and development. Trends Cog. Sci. 2001 Jul 1;5(7):296-301.

30. Stephan, A. (1999). Are Animals Capable of Concepts? *Erkenntnis* 51, 583-596.

31. Templeton, J. (1998) Learning from others' mistakes: a paradox revisited. *Animal Behaviour* 55: 79-85.

32. Thomas, R. (2002). Conditional Discrimination Learning and Conditional Reasoning by Nonhuman Animals. Updated version of Thomas, R. K. (1991). Misuse of conditional reasoning in animal research with special reference to evolution of language. Southern Society for Philosophy & Psychology, Atlanta, GA. http://rkthomas.myweb.uga.edu/Conditional.htm accessed 2008-01-07.

33. Treichler, R.

34. Vigo, R. (2007). Modal Similarity. Journal of Experimental and Theoretical Artificial Intelligence, under review.

35. Vigo, R., (2008) Predicting Degrees of Similarity in Modal Categories. In progress.

36. Watanabe S, Huber L (2006) Animal logics: Decisions in the absence of human language. Animal Cognition 9: 235-245.

37. Watson, J.S., 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. *Journal of Comparative Psychology* 115: 219-226.

38. Watson 2005 http://www.interdisciplines.org/causality/papers/7/4/1#_4

39. Wilkie, D. M., Wilson, R. J., & Kardal, S. (1989). Pigeons discriminate pictures of a geographic location. Animal Learning and Behavior, 17, 163-171.

40. Zentall, T. R. (2001). The case for a cognitive approach to animal learning and behavior. *Behavioural Processes* 54, 65-78.