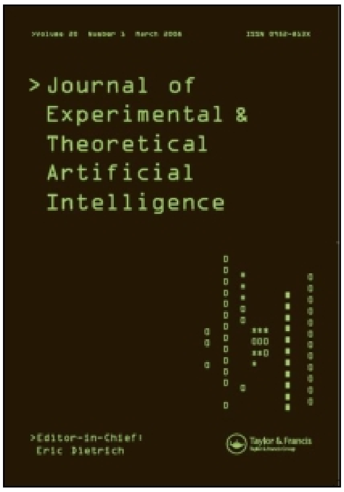

Journal of Experimental & Theoretical Artificial Intelligence
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713723652

Finding abstract commonalties of category members

Takashi Yamauchi [a]
[a] Department of Psychology, Texas A&M University, College Station, TX, USA

First Published:September2009

To cite this Article Yamauchi, Takashi(2009)'Finding abstract commonalties of category members',Journal of Experimental & Theoretical Artificial Intelligence,21:3,155 — 180

To link to this Article: DOI: 10.1080/09528130802113299

URL: http://dx.doi.org/10.1080/09528130802113299

Journal of Experimental & Theoretical Artificial Intelligence
Vol. 21, No. 3, September 2009, 155–180

Finding abstract commonalties of category members

Takashi Yamauchi*

Department of Psychology, Texas A&M University, College Station, TX, USA

(*Received 7 May 2007; final version received 23 September 2007*)

To form a coherent conceptual category and use it for inductive inference, the cognitive system needs to discover commonalties among different objects. How does the system accomplish this task? This study compares two of the most commonly used functions of categories – classification and feature inference – and examines their effectiveness in finding abstract commonalties of category members. The results from two experiments show that a classification task is not very useful for abstraction. In contrast, a feature inference task is advantageous in extracting abstract commonalties. However, this advantage is limited. Finding abstract commonalities becomes burdensome when category labels are absent in the feature inference task. These results underscore the importance of category membership information for abstraction. It is suggested that this advantage comes from the fact that category labels help form structured representation and facilitate structural alignment.

Keywords: categories; abstraction

1. Introduction

The ability to find commonalties among different instances is among the most remarkable aspects of human cognition. Medical diagnosis, solving complex physics problems, finding a solution for social problems or learning new concepts, all require the discovery of abstract commonalties among drastically different instances.

In this article, I focus on one type of abstraction – finding abstract commonalities among category members. Although much categorisation research has suggested that 'commonalties' are the building block of conceptual categories (e.g. Rosch and Mervis 1975), it is unclear how the cognitive system discovers them. In this study, I contrast classification and feature inference tasks – two of the most important functions of categories (Smith 1994; Yamauchi and Markman 1998; Murphy 2002; Markman and Ross 2003), and present an insight into the mechanism of abstraction. The specific goal of this article is to explore the following hypothesis: category labels help form structured representations and facilitate structural alignment. In brief, I derived this hypothesis from theories developed in three interrelated areas – inductive inference, categorisation and similarity judgment (see, in particular, Markman and Gentner 1993a, 1993b; Lassaline 1996; Gentner and Markman 1997; Lassaline and Murphy 1998; Markman and Wisniewski 1997; Yamauchi and Markman 2000b; Yu and Yamauchi 2006).

*Email: tya@psyc.tamu.edu

ISSN 0952–813X print/ISSN 1362–3079 online

DOI: 10.1080/09528130802113299
http://www.informaworld.com

1.1. *Background and previous studies*

Categorical concepts, such as *dogs*, *mammals* and *tools*, are represented by a collection of features, exemplars and/or their summaries (Medin and Schaffer 1978; Rosch 1978; Nosofsky 1986; Anderson 1990; Kruschke 1992; Love, Medin and Gureckis 2004). The basis of this idea is that conceptual categories are basically 'containers' storing information about their members, and category labels provide the names for the containers. In this approach, inductive inference and classification arises from the similarity-based statistical process (Anderson 1990; Sloman 1993; Tenenbaum 1999; Tenenbaum and Griffiths 2001; Pothos 2005).

However, there is considerable evidence that category labels play a distinct role in inductive generalisation; category labels help organise underlying attribute information and guide inductive inferences (Yamauchi, Kohn and Yu 2007; Yamauchi and Yu in press; but see also Sloutsky 2003; Sloutsky and Fisher 2004 for a different view). Even when categories are organised arbitrarily with geometric stimuli, category labels generate an expectation that members of a category have some attributes in common (Yamauchi, Love and Markman 2002; Pothos, Chater and Stewart 2004). Children between the ages of two and four understand noun labels as representing a group of related objects, rather than a group of attributes (Markman and Hutchinson 1984; Gelman and Markman 1986; Markman 1999). Research further indicates that both children and adults have special expectations about biological categories as well as personal and social categories (Gelman 2003; Gelman and Coley 1990; Gelman and Heyman 1999; Medin and Atran 2004; Wisniewski and Medin 1994; Wattenmaker 1995; Yamauchi 2005; Yamauchi and Yu, 2008; Yamauchi et al. 2007).

In this article, I ask whether this inductive potential of categorical noun labels can be extended to the abstraction of commonalties. I postulate that category labels lead people to form structured representations, encourage high-order cognitive processes such as structural alignment, and help find abstract commonalties among category members.

The effectiveness of structural alignment in inductive generalisation has been well documented in research in analogy, similarity judgment and inductive inference (Gentner 1983; Markman and Gentner 1993a; Medin, Goldstone and Gentner 1993; Holyoak and Thagard 1995; Goldstone 1996; Lassaline 1996; Gentner and Markman 1997; Hummel and Holyoak 1997; Lassaline and Murphy 1998; Yamauchi and Markman 2000b). This article aims to extend this idea to the abstraction of category members. In the following sections, I first introduce the operational definitions of the technical terms used in this article, and then lay out a possible mechanism that links category labels and structural alignment.

1.2. *Operational definitions and a mechanism of abstraction*

In this article, *abstract commonalties* are defined as common characteristics that go beyond the surface appearance of individual stimuli. For example, in Figure 1, the exact appearance of these schematic insects is different. However, these insects share 'abstract commonalties'. All the pictures labelled as 'monek' commonly *have long horns, round heads, dotted bodies, eight legs* and *short tails* although these commonalties are not immediately obvious. Similarly, all the pictures labelled as 'plaple' commonly have *short horns, angular heads, striped bodies, four legs* and *long tails*. I define these common characteristics that relate multiple instances beyond their surface appearance as '*abstract commonalties*' (see also Goldstone, Medin and Gentner 1991; Lassaline 1996; Gentner and Medina 1998; Markman 1999; for a similar definition). *Classification* is defined as a practice in which the category membership of an item (e.g. 'Is John a liberal?') that is predicted on the basis of the item's known features (e.g. 'John reads the New York Times', and 'John likes wine'; Posner and Keele 1968; Medin and Schaffer 1978). *Feature inference* is defined as a practice in which an unknown characteristic of an item is predicted ('Does John read the New York Times?') on the basis of the item's other known features ('John likes wine') and its category membership ('John is a liberal'; Murphy and Ross 1994; Yamauchi and Markman 1998, 2000a).

Translating these definitions into an experimental setting, the classification task in this study required the prediction of the category label ('monek' or 'plaple') of a stimulus on the basis of a sample stimulus (Figure 2(a) and (b)); the feature inference task required the prediction of the value of a hidden feature of a stimulus on the basis of the same sample stimulus (Figures 2(c) and (d)). Thus, the two tasks were formally equivalent if category labels can be treated just as another feature (Anderson 1990; Sloman 1993). The only difference was that the membership information was available in the feature inference task, but not in the classification task. However, I argue that this difference is pivotal in the discovery of abstract commonalties, because category membership information, especially when stated by verbal noun labels, can help form structured representation and facilitate structural alignment.

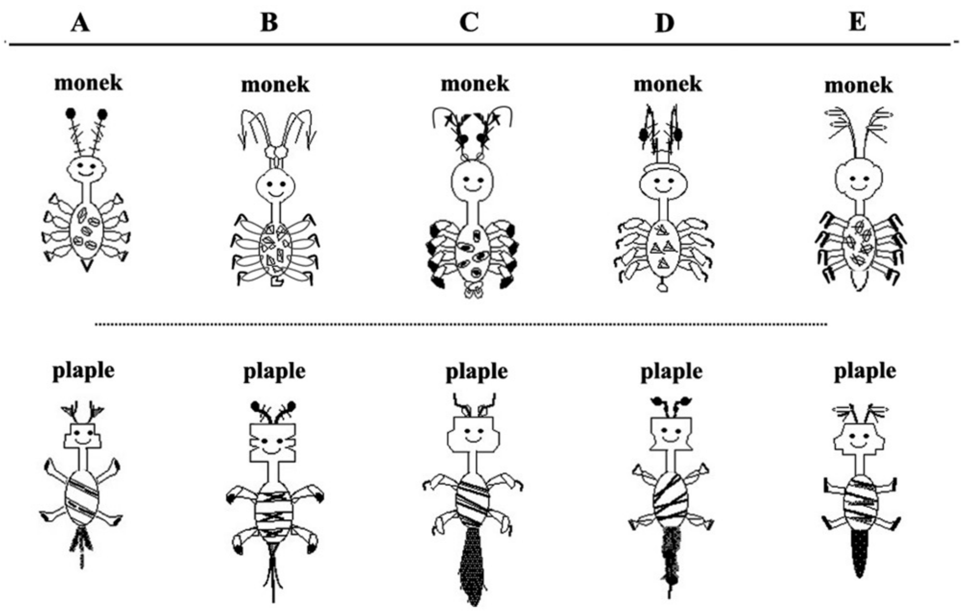

Figure 1. Ten prototypes that were used to produce test stimuli in the two experiments. Note that those labelled as 'monek/plaple' share 'abstract commonalties'. For example, all 'monek' prototypes have 'long horns', 'round heads', 'dotted bodies', 'eight legs', and 'short tails'. All 'plaple' prototypes have 'short horns', 'angular heads', 'striped bodies', 'four legs', and 'long tails'. Prototype set A was used to depict sample stimuli, as well as some of the test stimuli. Prototype sets B–E were used to depict test stimuli only.

1.3. *Structural alignment and category labels*

Observe (A) and (B) in Figure 3. It is not difficult to see the structural relationship between the two figures. In Figure 3(A), the length of the three lines increases in an ascending order. In Figure 3(B), the size of the circles increases in an ascending order. Structural alignment theory of similarity (Markman and Gentner 1993; Gentner and Markman 1997) explains that this relational commonalty becomes transparent as a result of 'structural alignment'. The lower panel of Figure 3 illustrates this process. In Figure 3(A′), two predicates, *Left* and *Smaller*, and one function, *length*, express the relationship between the three lines. In Figure 3(B′), two predicates, *Left* and *Smaller*, and another function, *size*, also express the relationship between the three circles. For example, $a \rightarrow Left \rightarrow b$ means that a is located on the left of b, and $a \rightarrow Smaller \rightarrow b$ means that a is smaller than b. Functions *length* and *size* each take one argument and return relevant values. By aligning these two representations with respect to their common predicates, *Left* and *Smaller*, the relationship between (A) and (B) becomes explicit, irrespective of the exact appearance of a, b, c, a', b', and c'. A similar alignment process is likely to mediate the abstraction of category members.

Observe Figure 4. In this figure, an arbitrary category, 'monek', is structurally organised by its label, *Monek*, along with its components *Horns*, *Head*, *Body*, *Legs* and *Tail*, and corresponding individual instances, {a, b, c, d, e} and {a′, b′, c′, d′, e′}. The small boxes represent the specific properties of these instances, such as 'long', 'curvy',

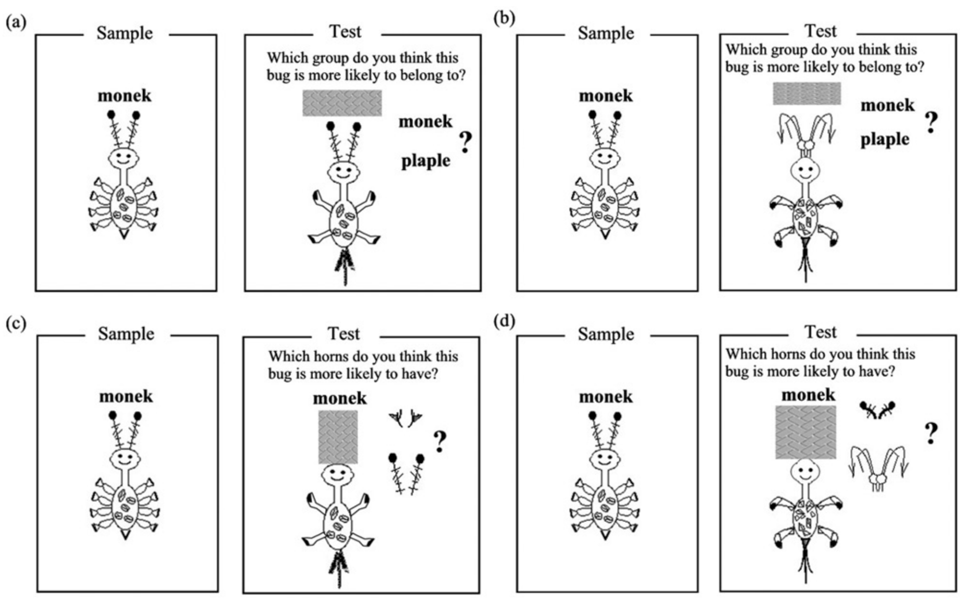

Figure 2. Examples of classification trials, (a) and (b), and inference trials, (c) and (d). The test stimuli in (a) and (c) were produced from the same feature set as the one used for the sample stimuli (set A – the exact instance condition). The test stimuli in (b) and (d) were produced from a different feature set from the one used for the sample stimuli (set B – the abstract instance condition). 'Consistent response' is defined as selecting the label/feature consistent with the sample stimulus. For example, selecting the 'monek' lable is a consistent response in (a) and (b), and selecting the long horns is a consistent response in (c) and (d).

'round,' and so on, and the arrows specify the whole-part relationship between them. For example, *Monek* → *Horns* means that *Monek* 'has-an-instance-of' *Horns*; *Horns* → *a* means that *Horns* 'have-an-instance-of' *a*; *a* → *Box i* means that *a* 'has-an-instance-of' the property specified by the box (e.g. 'long'). To predict the value of an unknown feature of a test stimulus (the lower panel in Figure 4), the structured representation of the sample stimulus is aligned to that of the test stimulus with respect to their corresponding predicates, and an inference can be made by projecting the arrows from a base concept (the sample stimulus) to a target concept (the test stimulus) (Lassaline 1996). In this case, the attribute of the horns in the sample stimulus is projected to that in the test stimulus, favouring the long horns over the short horns (Figure 4). In this manner, the inference task can be solved, and abstract commonalties binding two concepts become transparent.

In contrast to the feature inference task, structural alignment would be difficult in the classification task (Figure 5). Because category labels are absent in this task (they should be predicted), it is difficult to form structured representations. Even if structural alignment is applied in the classification task, its alignment would be shallower. For example, an alignment can be made by placing individual components – *Horns, Head, Body, Legs and Tail* – in correspondence (Figure 5 and see Lassaline and Murphy 1998). But these predicates are shallow as they take individual *instances* as arguments (e.g. first-order predicates), rather than other predicates as arguments (e.g. high-order predicates, *Monek* (*have*(*horns*(a, b)). Research in analogical transfer has shown that alignments based on first-order predicates are much weaker than alignments based on

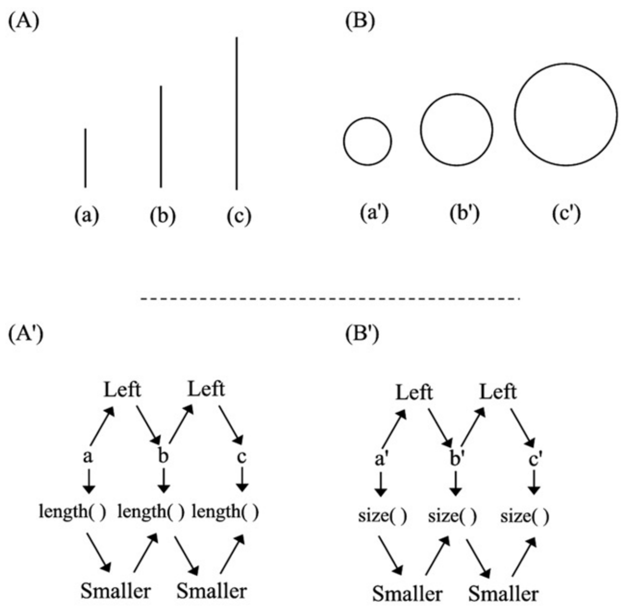

Figure 3. Three lines – (A), and three circles – (B) share a relational structure. In (A), the length of the lines changes in an ascending order; in (B), the size of the circles changes in an ascending order. Two possible structured representations of these figures were shown in (A′) and (B′). In both cases, two predicates, *Left* and *Smaller*, specify the relationship between individual instances (lines or circles).

higher-order predicates (i.e. 'systematicity principle' and see Gentner 1983; Falkenhainer, Forbus and Gentner 1989).

The classification task can be solved by focusing on a single salient feature (e.g. the number of legs). This focus on a salient feature, however, would not facilitate the discovery of abstract commonalties, unless it helps to form structured representation. In Figure 5, for example, the attention to a single feature (e.g. legs) can accentuate that particular feature. However, the level of alignment applied in this situation is not deep enough as long as this feature remains at the same hierarchical level as the other features.

Consistent with these observations, category learning research has shown that a classification task is extremely unwieldy in handling stimuli depicted with varying instances (Medin, Dewey and Murphy 1983; Yamauchi and Markman 2000b; Markman and Maddox 2003; Brooks and Hannah 2006). Computational and empirical evidence

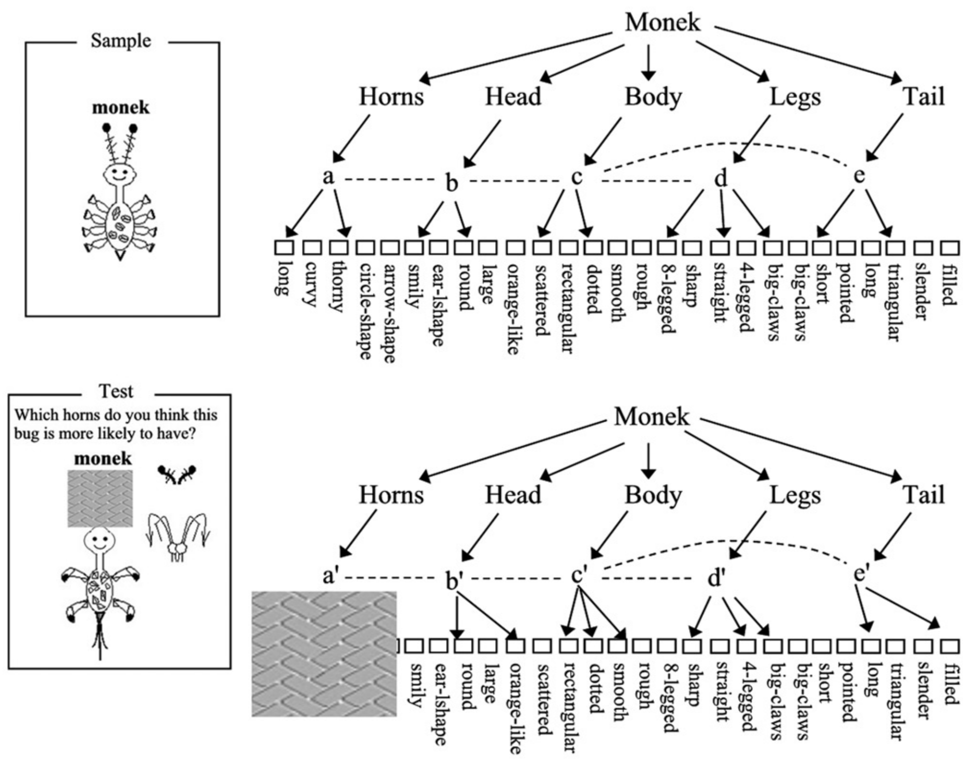

Figure 4. Structured representations of a sample stimulus (the upper panel) and a test stimulus (the lower panel) in an inference trial. In these representations, the category label 'monek' is used to organise individual components of a stimulus. In this manner, an arbitrary category, 'monek', is organised by its label, *Monek*, and components *Horns, Head, Body, Legs,* and *Tail,* and corresponding instances, {a, b, c, d, e} and {a′, b′, c′, d′, e′}. Each small box represents an attribute, such as 'long', 'curvy', 'round', and so on, and the arrows specify the relationship between arguments. For example, *Monek* → *Horns* means that *Monek* 'has-an-instance-of'' *Horns*. *Horns* → *a* means that *Horns* 'have-an-instance-of' *a*. *a* → *Box X* means that *a* 'has-an-instance-of' of the property specified by the box (e.g., 'long'). The dotted lines represent the spatial relationships between instances. For example, *a* - - - *b* means that *a* '*has-an-attached-part*' with *b*, and vice versa.

also reveals that people use a recognition heuristic to make classifications (Ross, Perkins and Tenpenny 1990; Nosofsky 1986; Ashby and Lee 1991). In the context of medical diagnosis, Brooks, LeBlanc and Norman (2000) demonstrated that even medical experts with an average of 10 years of practice failed to identify key pathological features of patients without proper information about their diagnostic class.

In other words, category labels, which denote category membership information, provide a basis to form a structured representation of a category, encourage structural alignment and help find abstract commonalties. This idea was tested in two experiments.

1.4. *Overview of experiments*

The stimuli used in the experiments were schematic illustrations of fictitious insects. They were composed of five feature dimensions with binary values (horns = long/short, head = round/angular, torso = dotted/striped, legs = 8 legs/4 legs, tail = short/long), along with category labels (‘monek’/’plaple’) (Figure 1). All participants received a sample stimulus and a test stimulus side-by-side and performed either a classification task or a feature inference task on the basis of the sample stimulus (Figure 2). The sample stimuli

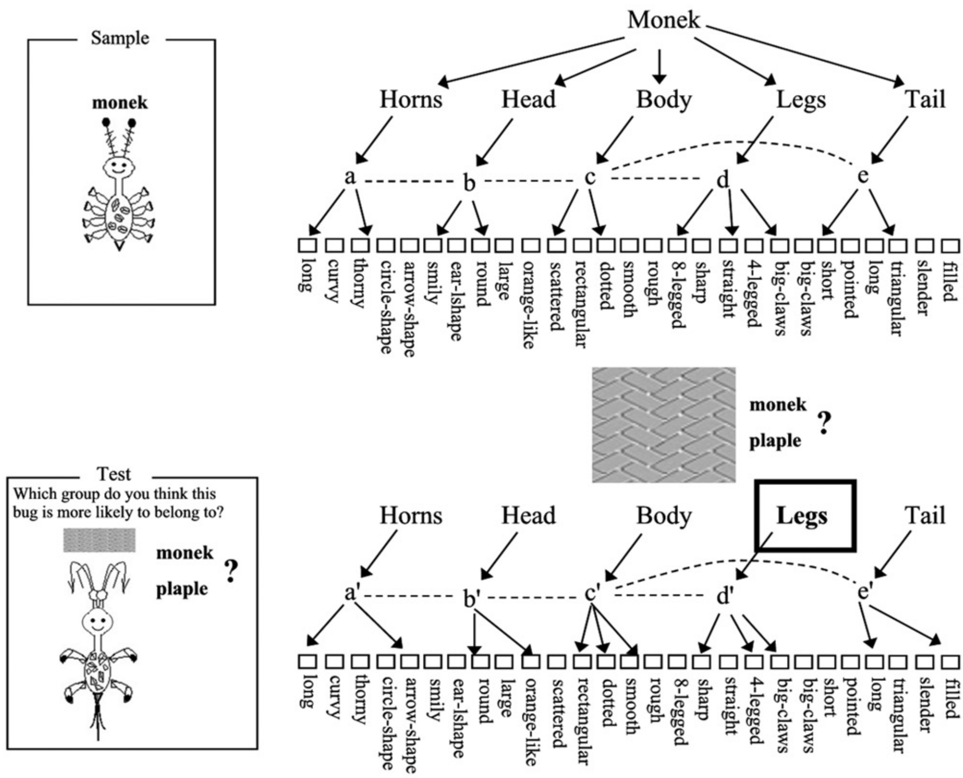

Figure 5. Structured representations of a sample stimulus (the upper panel) and a test stimulus (the lower panel) in a classification trial. In the test stimulus, the label is covered by a mask, therefore the representation of the test stimulus is shallower than that of the test stimulus in an inference trial (Figure 4). The black box enclosing the ‘legs’ feature suggests extra attention given to this feature.

shown next to test stimuli were prototypes of the corresponding category, 'monek' or 'plaple' (Table 1).

Among the 60 test stimuli that were presented to each participant, 40 stimuli were produced from the feature instances that were different from those used for the sample stimuli, but they shared abstract commonalties (sets B–E in Figure 1 and also see Figure 2(b) and (d)). The remaining 20 test stimuli were produced from the same feature instances that were used for the sample stimuli (set A in Figure 1 and also see Figures 2(a) and (c)).

I measured the proportion of participants selecting the feature instances consistent with the sample stimuli to assess the extent to which abstract commonalties affect the two tasks. For example, in Figure 2(b), the proportion of participants selecting the label ('monek') consistent with the sample stimulus was counted in the classification questions. In Figure 2(d), the proportion of participants selecting the feature value consistent with the sample stimulus ('long horns') was tallied. These two measures were formally equivalent if category labels are equivalent to another feature. For example, the classification question in Figure 2(b) can be expressed schematically as Sample $(1, 1, 1, 1, 1) =$ (long horns, round head, dotted body, eight legs, short tail, monek label), vs. Test $(1, 1, 1, 0, 0, ?) =$ (long horns, round head, dotted body, four legs, long tail, ?monek/plaple label), where a response of 1 ('monek') was defined as a 'consistent response'. Similarly, the inference question in Figure 2(d) can be expressed schematically as Sample $(1, 1, 1, 1, 1)) =$ (long horns, round head, dotted body, eight legs, short tail, monek label) vs. Test $(?, 1, 1, 0, 0, 1) =$ (?long/short horns, round head, dotted body, four legs, long tail, monek label), where a response of 1 (long horns) was defined as a 'consistent response'.

Table 1. The structure of the test stimuli used in Experiment 1.

Monek							Plaple						
Stimuli	Horns	Head	Body	Legs	Tail	Labels	Stimuli	Horns	Head	Body	Legs	Tail	Labels
High													
M1	***1***	1	1	1	0	1	P1	***0***	0	0	0	1	0
M2	1	1	1	0	***1***	1	P2	0	0	0	1	***0***	0
M3	1	1	0	***1***	1	1	P3	0	0	1	***0***	0	0
M4	1	0	***1***	1	1	1	P4	0	1	***0***	0	0	0
M5	0	***1***	1	1	1	1	P5	1	***0***	0	0	0	0
Low													
M6	***1***	1	1	0	0	1	P6	***0***	0	0	1	1	0
M7	1	1	0	0	***1***	1	P7	0	0	1	1	***0***	0
M8	1	0	0	***1***	1	1	P8	0	1	1	***0***	0	0
M9	0	0	***1***	1	1	1	P9	1	1	***0***	0	0	0
M10	0	***1***	1	1	0	1	P10	1	***0***	0	0	1	0
M0 (sample stimulus)	1	1	1	1	1	1	P0 (sample stimulus)	0	0	0	0	0	0

Notes: 'High' and 'Low' stand for the two levels of the feature level factor. The numbers shown with bold italics are target features queried in inference trials. 'Consistent responses' were defined as those consistent with the values indicated in the target features and the target labels. M0 and P0 are the protototypes of 'monek' and 'plaple' categories. They were shown next to each test stimulus as a 'sample'. $1/0 =$ long/short horns, round/angular head, dotted/striped body, eight/four legs, short/long tail.

Experiment 1 contrasted the classification task and the feature inference task. Experiment 2 introduced three different inference tasks and examined the role of category labels in the discovery of abstract commonalties.

2. Experiment 1

As discussed earlier, the main distinction between the classification and feature inference tasks is the presence of category labels. In the feature inference task, the category membership information was specified with an arbitrary label ('monek'/'plaple'), while the classification task involved the prediction of the value of these labels (therefore, the label was covered by a mask). This subtle distinction should be crucial for the discovery of abstract commonalties, if, as hypothesised, category labels facilitate the formation of structured representation. The responsiveness to abstract commonalties was assessed by the proportion of participants selecting labels/features consistent with the sample stimuli (i.e. a *consistent score*). For example, selecting the 'monek' label was a *consistent* response in Figure 2(b) and selecting the long horns was a *consistent* response in Figure 2(d). Because of the presence of category labels, the inference task should yield higher consistent scores than the classification task.

2.1. *Method*

2.1.1. *Participants*

Two-hundred and twenty-five undergraduate students participated in the experiment for course credit. These participants were randomly assigned to either a classification condition ($N = 136$) or an inference condition ($N = 119$).

2.1.2. *Materials*

The stimulus materials were schematic illustrations of cartoon insects produced from five sets (A, B, C, D, and E) of prototypes (Figure 1). Each stimulus was composed of a combination of five feature dimensions with binary values (horns = long/short, head = round/angular, torso = dotted/striped, legs = 8 legs/4 legs, tail = short/long), along with category labels ('monek'/'plaple') (Table 1). The depiction of these features varied across the five stimulus sets, while they maintained abstract commonalties. For example, the monek prototypes in all the five sets had long horns, round heads, dotted torsos, 8 legs and short tails, while the exact appearance of these components was different across the sets.

Individual trials consisted of pairs of a sample stimulus and a test stimulus (Figure 2). We created 100 test stimuli from the five sets, A, B, C, D, and E, (20 stimuli from each set – Figure 1). These test stimuli were produced systematically by swapping individual components of the two prototypes (Table 1). For example, in the high-feature match stimuli (M1–M5 and P1–P5 in Table 1), four features were taken from the prototype of one category, and one feature was taken from the prototype of the other category. In the low-feature match stimuli (M6–M10 and P6–P10 in Table 1), three features were taken from the prototype of one category, and two features were taken from the prototype of the other category. The sample stimuli, which were shown right next to test stimuli, were the two prototype stimuli of set A.

2.1.3. *Procedure*

At the beginning of the experiment, participants were randomly assigned to one of two conditions – a classification condition or an inference condition, in which either 60 classification or 60 inference questions were given. For each question, the computer displayed the sample stimulus on the left and the test stimulus on the right (Figure 2). The classification question was phrased as 'Which group do you think this bug is more likely to belong to?' (Figures 2(a) and (b)). An inference question was phrased as 'Which *feature* do you think this bug is more likely to have?', where feature was replaced with 'horns', 'head', 'body', 'legs', or 'tail' (Figures 2(c) and (d)). Participants in the inference condition predicted the value of all the four feature dimensions (Table 1).

Participants indicated their responses by clicking one of the two designated buttons. The order of presenting stimuli was determined randomly for each participant. The same sample stimuli were shown to all participants throughout the experiment. Participants were tested individually at their own pace. The entire experiment took about 20 min.

2.1.4. *Design*

The design of the experiment was a 2 (question type: classification and inference; between-subjects) $\times$ 2 (feature instance: exact vs. abstract; within-subjects) $\times$ 2 (feature level: high vs. low; within-subjects) $\times$ 2 (stimulus version: version 1 vs. version 2; between-subjects) factorial. The feature-instance factor, which had two levels – exact and abstract, manipulated the surface correspondence between sample and test stimuli. The test stimuli in the exact condition were produced from set A. Thus, the sample and test stimuli had the same feature instances in the exact condition (the sample stimuli were the two prototype stimuli in set A). The test stimuli in the abstract condition were produced from sets B and C (version 1) or sets D and E (version 2). Thus, the sample and test stimuli had different feature instances in the abstract condition but they were related to each other by abstract commonalties.

The dependent measure was the proportion of participants making consistent responses (see Table 1 for the definition). For example, given test stimulus M1 in Table 1, selecting the label 'Monek' was defined as a consistent response in the classification condition and selecting 'long horns' was defined as a consistent response in the inference condition. Because the two versions of the stimuli materials did not show any systematic differences, the stimulus version factor was collapsed in the data analysis. The main focus of the data analysis involved an interaction effect between question-type and feature-instance.

2.2. *Results and discussion*

2.2.1. *Overall results*

The three-way interaction effect of question type (classification vs. inference), feature instance (exact, abstract) and feature level (high, medium) was marginally significant; $F(1, 253) = 3.25$, $MSE = 0.001$, $p = .07$, $\eta^2 = 0.01$ (Table 2). Pair-wise t-tests revealed that the mean consistent score obtained in the inference condition was significantly higher than that in the classification condition in the stimuli with abstract instances at the low-feature level (Table 2). This result suggests that the classification performance was far less receptive to abstract commonalties than the inference performance. A similar interaction effect was observed in the Lassaline study (Experiment 2, 1996), in which the impact of structural alignment was pronounced particularly when two propositional concepts had few attributes in common.

A main effect of question type as well as a two-way interaction effect between question type and feature level were also significant; the main effect, $F(1, 253) = 11.08$, $MSE = 0.11$, $p < 0.001$, $\eta^2 = 0.042$; the interaction effect, $F(1, 253) = 10.41$, $MSE = 0.02$, $p < 0.001$ $\eta^2 = 0.04$.

2.2.2. *Controlling decision criteria*

The high consistent scores observed in the inference condition relative to those in the classification condition might have been a reflection of different decision criteria employed in the two tasks. For example, classification judgments were made with a conservative decision criterion, while inference judgments were made with a less conservative decision criterion. To rule out this explanation, an *ANCOVA* (analysis of covariance) was applied to the abstract feature instance stimuli after controlling the performance for the exact feature instance stimuli as a covariate. If the disparity of the two tasks simply reflected different decision criteria, the difference between the two tasks should disappear with this measure. Even with this corrective measure, the two conditions were significantly different in the abstract instance stimuli; $F(1, 252) = 7.74$, $MSE = 0.02$, $p < 0.01$, $\eta^2 = 0.03$.

2.2.3. *Focusing on a single diagnostic feature in the classification task*

Participants in the classification condition might have used a single diagnostic feature to make classification judgments. For example, by attending to the leg feature, the consistent label (e.g. 'monek' in M1) may be selected when the sample and test stimuli had the same leg feature (e.g. M1 in Table 1) but not when the sample and test stimuli had different leg features (e.g. M2 in Table 1). To examine this possibility, we removed the data from the test stimuli that had inconsistent values in each feature dimension (e.g. we removed the

Table 2. A summary of the results from Experiment 1.

		High	Low	Average
			Abstract feature-instance	
Inference	*M*	0.64	0.64	0.64
Classification	*M*	0.62	0.54	0.58
	t-score	0.89	4.74	3.03
	p-value	0.370	0.000	0.003
	d	0.11	0.60	0.38
			Exact feature-instance	
Inference	*M*	0.74	0.70	0.72
Classification	*M*	0.69	0.61	0.65
	t-score	2.02	3.88	3.24
	p-value	0.040	0.000	0.001
	d	0.25	0.49	0.41

Notes: The *t*-scores, *p*-values, and *d*'s are based on the comparisons between the inference condition and the classification condition along the same column. 'High' and 'Low' correspond to the two levels of the feature level factor (see Table 1). M = mean, d = Cohen's d (effect size).

Table 3. Comparing classification and inference performance after excluding inconsistent stimuli in Experiment 1.

		Horns	Head	Body	Legs	Tail
Inference	*M*	0.63	0.62	0.65	0.67	0.58
Classification	*M*	0.50	0.52	0.59	0.59	0.51
	t-score	4.30	2.73	1.92	3.26	5.50
	p-value	0.000	0.012	0.068	0.004	0.000
	d	1.43	0.89	0.92	1.30	1.72

Notes: The *t*-scores, *p*-values, and *d*'s are based on the comparisons between the inference condition and the classification condition after excluding stimuli with inconsistent feature values. For example, in the column under 'Horns', the test stimuli, M5, M9, M10, P5, P9 and P10 were removed because these stimuli have the horns inconsistent with the ones shown in the sample stimuli. The classificaton and inference performance was compared after removing these stimuli. M = mean, d = Cohen's d (effect size).

data from M2, M6, M7, P2, P6 and P7 in assessing the effect of the leg feature, we removed the data M3, M7, M8, P3, P7 and P8 in assessing the effect of the body feature, and so on) and compared the performance for the two tasks. Table 3 summarises the results from this analysis. As the table shows, the inference performance is still consistently higher than the classification performance even after removing inconsistent stimuli in all dimensions.

Taken together, the results from Experiment 1 indicate that there was a significant disparity between the inference task and the classification task in terms of their capacity to handle abstract feature information. This disparity appears to stem from the fact that category labels were present in the inference task but not in the classification task. Experiment 2 tested directly the role of category labels in finding abstract commonalties.

3. Experiment 2

The hypothesis that category labels help construct structured representations and promote structural alignment suggests that the abstraction of category members should become difficult when category labels are missing or when the labels do not carry category membership information. Figure 6 illustrates this idea. In Figure 6(a), the label is removed and a new feature, wings, is attached. In Figure 6(b), the label is still present but the label is associated with another feature (the label is assumed to represent the shape of wings). The extraction of abstract commonalties should be difficult in these conditions, if, as hypothesised, category labels help form structured representation and facilitate structural alignment.

In Experiment 2, all participants answered the same inference questions in one of three conditions – a category-label condition, an attribute-label condition, or an attribute condition. The category-label condition and the attribute-label condition were identical except for their instructions. The instructions in the category-label condition characterised the two labels, 'monek' and 'plaple', as signifying two 'types' of bugs. The instructions in the attribute-label condition characterised the same labels as signifying two different 'shapes' of wings. Thus, the two arbitrary labels were associated with attribute information in the attribute-label condition. The two conditions were identical except

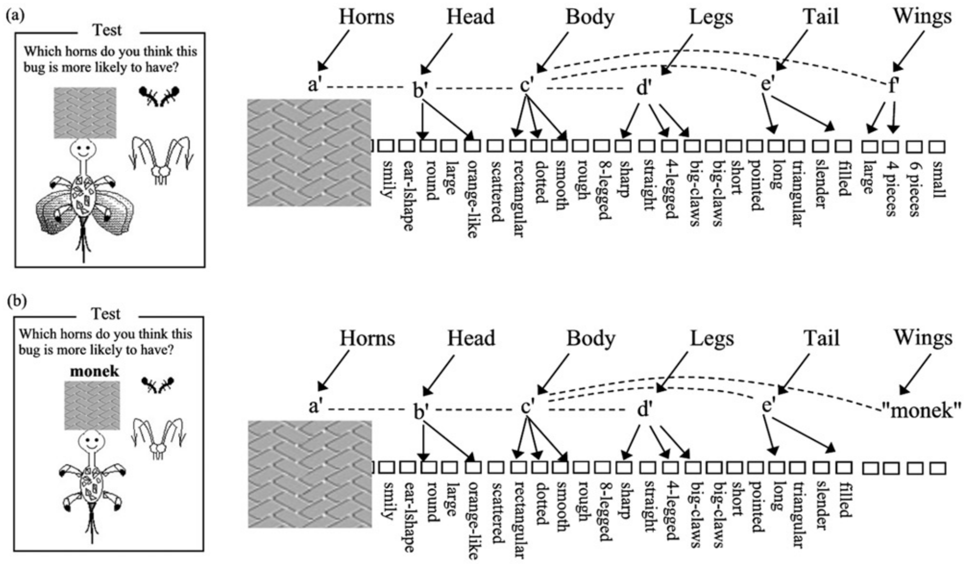

Figure 6. The structured representation of two inference trials when labels are replaced with wings – (a), and when verbal labels are associated with an attribute (wings) – (b). Note that the labels are replaced with actual wings in (a).

for this single point. In the attribute condition, the two verbal labels were removed from all sample and test stimuli, and were replaced with two different kinds of wings (large vs. small wings, Figure 8).

Participants answered inference questions when sample and test stimuli shared the same labels (i.e. a matched condition, and see Figure 7(a)) and when sample and test stimuli had different labels (i.e. a mismatched condition, and see Figure 7(b)). In the attribute condition, participants answered inference questions when sample and test stimuli had the same wings (i.e. a matched condition, and see Figure 8(a), and when sample and test stimuli had different wings (i.e. a mismatched condition, and see Figure 8(b)).

3.1. ***Predictions***

3.1.1. *Prediction 1 (consistent score)*

The extraction of abstract commonalties should be difficult when category labels are absent (attribute condition) or when the labels convey attribute information (attribute-label condition). Thus, the proportion of consistent response in the attribute and attribute-label condition should be substantially lower than that in the category-label condition.

3.1.2. *Prediction 2 (polarity effect)*

If category labels promote structural alignment (aligning predicates rather than matching feature values), the matching/mismatching status of labels should dictate the response

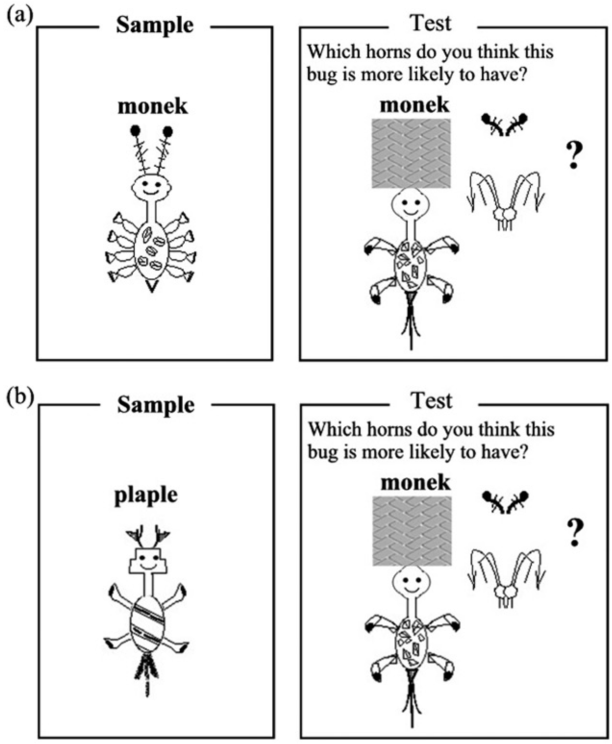

Figure 7. Two samples of stimulus frames used in the category-label condition and in the attribute-label condition in Experiment 2. In (a), the sample and test stimuli have the same label ('monek') – a matched condition. In (b), the sample and test stimuli have different labels – a mismatched condition.

patterns in the category-label condition. For example, the proportion of participants selecting consistent features should go up or down drastically depending on whether the sample and test stimuli have the same label (Figure 7(a)) or different labels (Figure 7(b)) (polarity effect). This tendency should be reduced substantially in the attribute and attribute-label conditions.

3.1.3. *Prediction 3 (uniformity effect)*

If structural alignment is applied in the category-label condition, responses made in the category-label condition should be highly homogeneous and uniform. Research has shown that structural alignment guides attention to the relationship between objects, transforming an attribute-based context-dependent representation to an abstract uniform representation are based on object structure (Gentner and Medina 1998; Gentner and Namy 1999). The uniformity of representation was examined by comparing the extent to which the responses made by individual participants correlated (the detail of this analysis procedure is described in the result section).

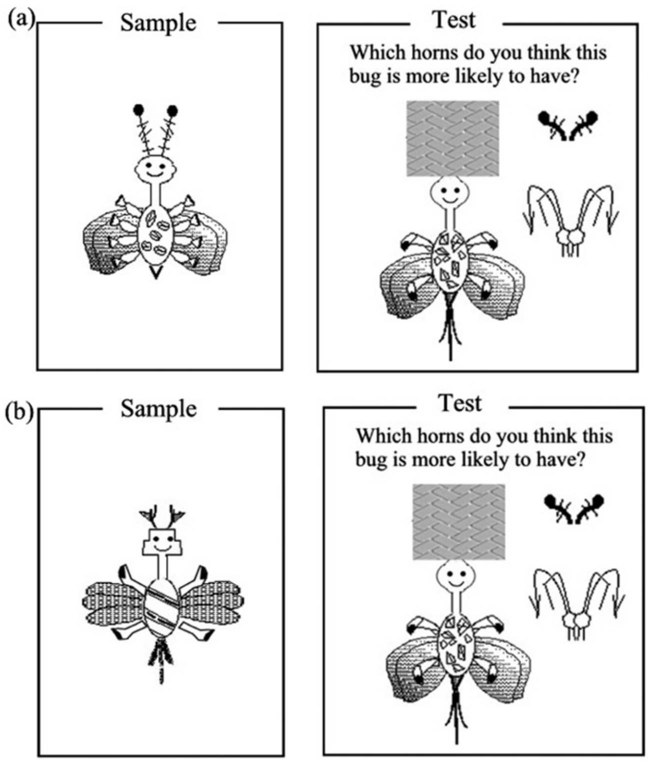

Figure 8. A sample stimulus frame used in the attribute condition in Experiment 2. In the matched condition – (a), the sample and test stimuli have the same wings. In the mismatched condition – (b), the sample and test stimuli have different wings.

3.2. *Method*

3.2.1. *Participants*

A total of 138 undergraduate students participated in the experiment for course credit. These participants were assigned to one of two conditions – a category-label condition ($N = 49$), an attribute-label condition ($N = 41$), or an attribute condition ($N = 48$).

3.2.2. *Materials*

Individual trials consisted of pairs of a sample stimulus and a test stimulus (Figures 7 and 8). All test stimuli had two features consistent with the prototype of one category and two features consistent with the prototype of the other category and one feature was masked for an inference question (Table 4). One version of the stimulus materials was produced from prototype sets A, B and C. The other version was produced from prototype sets A, D and E (Figure 1). In the attribute condition, the two labels were replaced with actual wings (Figure 8).

Table 4. The structure of the test stimuli used in Experiment 2.

	Horns	Head	Body	Legs	Tail	Labels
M6	***?/1***	1	1	0	0	1
M7	1	1	0	0	***?/1***	1
M8	1	0	0	***?/1***	1	1
M9	0	0	***?/1***	1	1	1
M10	0	***?/1***	1	1	0	1
M0 (sample stimulus)	1	1	1	1	1	1
P6	***?/0***	0	0	1	1	0
P7	0	0	1	1	***?/0***	0
P8	0	1	1	***?/0***	0	0
P9	1	1	***?/0***	0	0	0
P10	1	***?/0***	0	0	1	0
P0 (sample stimulus)	0	0	0	0	0	0

Notes: '?' indicates target questions. The responses consistent with the value shown in the corresponding sample stimulus (M0/P0) were defined as 'consistent responses'. In the matched trials, the sample and test stimuli both had the same labels. In the mismatched trials, the sample and test stimuli had different labels. For example, test stimulus M6 was shown twice. In a matched trial, it was paired with M0. In a mismatched trial it was paired with P0. One version of test stimuli was produced from prototype sets A, B and C, and the other version was produced from prototype sets A, D, and E (see Figure 1).

In the category-label condition, the instructions characterised the two labels ('monek' and 'plaple') as representing 'two types' of bugs. In the attribute-label condition, the instructions characterised the two labels ('monek' and 'plaple') as representing two different 'shapes of wings'. In the attribute condition, the labels were removed and replaced with actual wings. Excerpts of the instructions from the three conditions are shown below:

> *Category-label condition*. In this experiment, we are interested in the way you make judgments... Each bug belongs to two types – 'monek' and 'plaple'. These bugs are depicted with 5 different body parts – horns, head, body, legs, and tail, along with a tag, 'monek' or 'plaple'... Based on the sample shown on the left side of the screen, please answer each question as accurately as you can...
>
> *Attribute-label condition*. In this experiment, we are interested in the way you make judgments... Each bug is depicted with 6 different body parts – horns, head, body, legs, tail and wings. Because the wings of the bugs are folded on their back, we were not able to show them, so that we specified them with two names – 'monek' and 'plaple', which roughly stand for two different shapes of wings... Based on the sample shown on the left side of the screen, please answer each question as accurately as you can....
>
> *Attribute condition*. In this experiment, we are interested in the way you make judgments... These bugs are depicted with 6 different body parts – horns, head, body, legs, wings, and tail... Based on the sample shown on the left side of the screen, please answer each question as accurately as you can...

3.2.3. *Procedure*

For each trial, participants were shown a pair of sample and test stimuli on a computer screen, and were instructed to select one of two feature values for the body part in question.

Thirty test stimuli were shown twice. In one case, a test stimulus was paired with a sample stimulus that had the same label (i.e. *match* condition) (Figure 7(a)). In the other case, the same test stimulus was paired with a sample stimulus that had a different label (i.e. *mismatch* condition) (Figure 7(b)). For example, stimulus M6 in Table 4 was shown twice, once with the sample stimulus M0 (this is called a *matched* trial because M6 and M0 had the same label 'monek') and once with the sample stimulus P0 (this is called a *mismatched* trial because M6 and P0 had different labels, 'monek' and 'plaple'). Each participant received a total of 60 trials. In the attribute condition, matched/mismatched trials were produced by matched/mismatched wings (Figure 8).

3.2.4. *Design*

The design of the experiment was 3 (label status – category-label, attribute-label, attribute; between-subjects) $\times$ 2 (feature instance – exact vs. abstract; within-subjects) $\times$ 2 (match status – matched vs. mismatched; within-subjects) $\times$ 2 (stimulus version – version 1 vs. version 2; between-subjects) factorial. *Stimulus version* did not interact with *label status*; therefore this factor was collapsed in subsequent data analyses. *Feature instance* stands for the correspondence of the feature sets used to depict sample stimuli and test stimuli. As in Experiment 1, the sample and test stimuli had the same feature instances in the exact condition (they were produced from set A). The sample and test stimuli had different feature instances in the abstract condition (the test stimuli were produced from sets B and C or from sets D and E).

The dependent measure of the experiment was the proportion of consistent responses – selecting the feature value consistent with the corresponding sample stimulus.

3.3. ***Results and discussion***

3.3.1. *Prediction 1: Consistent scores*

This prediction was tested by applying ANOVAs to the data obtained from the abstract instance stimuli and the exact instance stimuli separately.

First, given the abstract instance stimuli, there was an interaction effect between label type and match status; $F(2, 135) = 31.0$; $MSE = 0.04$, $p < 0.001$, $\eta^2 = 0.32$. In the matched stimuli, the mean consistent score observed in the category-label condition was substantially higher than that in the attribute condition and in the attribute-label condition; in all comparisons, t's > 4.90, p's < 0.001, d's > 1.00 (Table 5). The effect of category label was significant even after controlling the performance for the stimuli with the exact feature instance; $ANCOVA$, $F(2, 134) = 35.05$, $MSE = 0.02$, $p < 0.001$, $\eta^2 = 0.34$. In the mismatched stimuli, the mean consistency score in the category-label condition was significantly lower than that in the attribute condition and in the attribute-label condition; in all comparisons, t's > 2.50, p's < 0.01, d's > 0.55 (Table 5). This effect is significant even after controlling the performance for the stimuli with the exact instance; $ANCOVA$, $F(2, 134) = 16.82$, $MSE = 0.02$, $p < 0.001$, $\eta^2 = 0.20$. Note that this result is in accord with the prediction that participants in the category-label condition use the labels for predicate alignment. In the mismatched condition, the labels shown in the sample and test stimuli were mismatched, so the proportion of participants selecting the feature consistent with the sample stimulus should decline accordingly.

Table 5. A summary of the results from Experiment 2 (the stimuli with abstract feature-instance only).

		Abstract feature-instance		
		Matched	Mismatched	Polarity = matched – mismatched
Category-label	*M*	0.77	0.29	0.48
Attribute	*M*	0.50	0.48	0.02
Attribute-label	*M*	0.57	0.42	0.15
Category-label vs. attribute	*t*-score	9.07	4.71	7.86
	p-value	0.000	0.000	0.000
	d	1.84	0.96	1.60
Category-label vs. attribute-label	*t*-score	4.95	2.69	4.49
	p-value	0.000	0.009	0.000
	d	1.05	0.57	0.95

Note: M = mean, d = Cohen's d (effect size).

Given the exact feature instance stimuli, there was an interaction effect between feature level and label status; $F(1, 135) = 4.94$, $MSE = 0.46$, $p < 0.01$, $\eta^2 = 0.07$ (Table 6). Overall, this difference was particularly pronounced between the category-label condition and the attribute-label condition. In the matched stimuli, participants in the category-label condition made consistent responses more often than participants in the attribute-label condition; $t(88) = 2.85$, $d = 0.60$, $p < 0.01$; this tendency was reversed in the mismatched stimuli; $t(88) = 2.34$, $d = 0.50$, $p < 0.05$. The difference between the category-label condition and the attribute condition was not significant after Bonferonni adjustments.

Taken together, these results support the hypothesis that finding abstract commonalties was particularly difficult when category labels were missing or when labels conveyed attribute information.

3.3.2. *Prediction 2: Polarity effect*

This prediction derives from the idea that structural alignment would generate direct mapping of labels, creating drastically different feature inferences depending on the matched/mismatched status of the two labels. That is, when sample and test stimuli have the same label, then the feature value consistent with the sample would be selected. In contrast, when sample and test stimuli have mismatched labels, then the projection should decline sharply, yielding highly polarised response patterns. To test this prediction, 'polarity scores' were calculated for individual participants by subtracting the performance for the mismatched stimuli from that for the matched stimuli.

In the abstract instance stimuli, the mean polarity score in the category-label condition was significantly higher than those observed in the attribute-label condition and in the attribute condition; $t's > 4.48$, $p's < 0.001$, $d's > 0.95$ (Table 5). The difference between the category-label condition and the attribute-label and attribute conditions was substantial even after controlling their performance in the exact instance stimuli; *ANCOVA* controlling the exact instance stimuli as a covariate; $F(2, 134) = 35.67$, $MSE = 0.05$, $p < 0.001$, $\eta^2 = 0.35$.

Table 6. A summary of the results from Experiment 2 (the stimuli with exact feature-instance only).

		Exact feature-instance		
		Matched	Mismatched	Polarity = matched − mismatched
Category-label	M	0.79	0.27	0.52
Attribute	M	0.69	0.30	0.39
Attribute-label	M	0.64	0.41	0.23
Category-label vs. attribute	t-score	2.03	0.56	1.47
	p-value	0.045	0.578	0.146
	d	0.41	0.11	0.30
Category-label vs. attribute-label	t-score	2.85	2.34	3.30
	p-value	0.005	0.021	0.001
	d	0.60	0.50	0.70

Note: M = mean, d = Cohen's d (effect size).

In the exact instance stimuli, the polarity score in the category-label condition was significantly higher than that in the attribute-label condition (Table 6). The polarity scores observed in the attribute condition was statistically indistinguishable from that in the category condition (Table 6).

3.3.3. *Prediction 3: Uniformity effect*

Another way of testing the application of structural alignment is to check the response patterns of individual participants. If structural alignment is applied uniformly, then participants' responses should be highly correlated. Gentner and Medina (1998) suggest that structural alignment creates more abstract uniform representation, resulting in homogeneous rule-based response patterns, rather than 'similarity-based' responses. To test this idea, I measured the extent to which the response patterns obtained from individual participants correlated. This was carried out with the following steps (see Figure 9 for an illustration of the procedure). *Step 1*, the entire responses of every participant were transcribed to a vector of 60 dimensions (each dimension represents a response score (1 or 0) in one trial, and there were 49, 48 and 41 vectors respectively in the category-label condition, the attribute condition, and the attribute-label condition, respectively); *Step 2*, in each label status condition, 30 response vectors were randomly selected, and these 30 vectors were randomly divided into two groups of 15 vectors; *Step 3*, the individual elements of the 15 vectors were averaged over each dimension, yielding two group-vectors of 60 dimensions; *Step 4*, Pearson's correlation coefficient score was measured along the two group-vectors. Steps 2–4 were repeated 1000 times in each label status condition and mean correlation scores were calculated from a sample of 1000 (see Johansen and Kruschke 2005; Rosch and Mervis 1975 for a similar analysis). Overall, the response patterns observed in the category-label condition were highly correlated; the mean correlation score of 1000 samples in the category-label condition, $M = 0.87$, $SD = 0.04$; as compared to the attribute condition, $M = 0.56$, $SD = 0.13$ and the attribute-label condition, $M = 0.30$, $SD = 0.13$ (Figure 10).

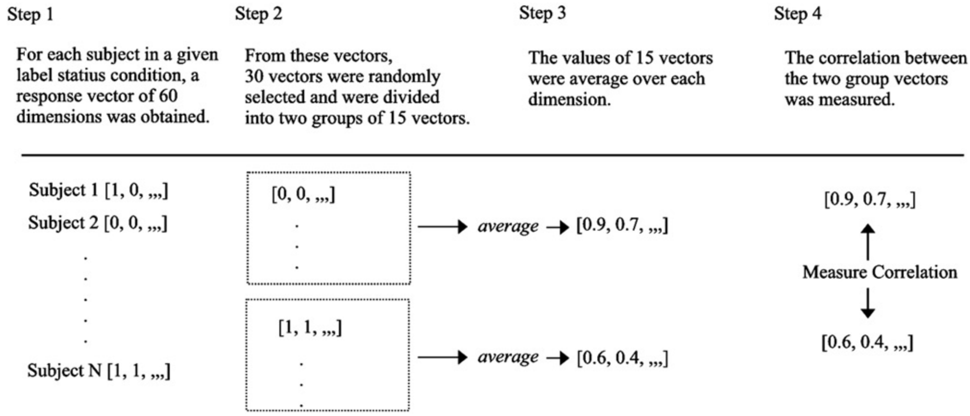

Figure 9. An illustration of the correlation analysis used in Experiment 2.

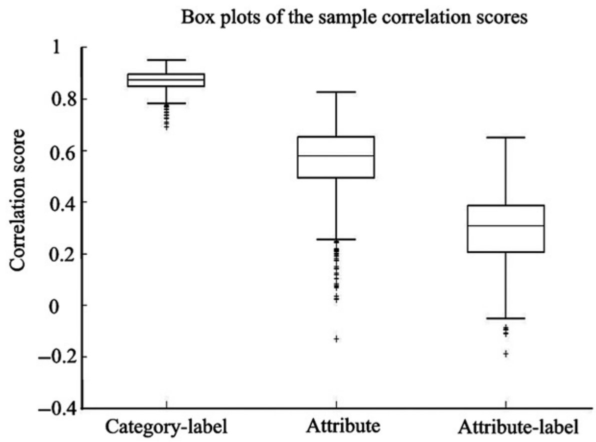

Figure 10. Three box plots of sample correlation scores obtained in Experiment 2. The tops and bottoms of each 'box' are the 25th and 75th percentiles of the samples, respectively. The whiskers enclose the entire samples except outliers, which are shown with '+' signs. The outliners are defined as those that are more than 1.5 times the interquartile range away from the top or bottom of the box.

To summarise, the results from Experiment 2 suggest that when category labels were missing or labels did not carry category information, abstraction of category members became particularly difficult; furthermore, when the two labels carried category information, participants' responses became polarised and uniform, supporting the view that category labels help form structured representation and facilitate structural alignment.

4. General discussion

How does the cognitive system find abstract commonalties among category members? I suggest that category labels facilitate a high-order cognitive strategy such as *structural alignment* and help find abstract commonalties of category members.

Experiment 1 compared the classification and inference tasks and showed a severe disadvantage in the classification task. The performance for the classification task was significantly affected by the varying instances of individual features, while this was not the case in the inference task. Experiment 2 investigated the source of this discrepancy and showed that feature inference became difficult particularly when the verbal labels were absent or when the labels carried attribute information. When the labels carried category information, participants made inferences in a rule-like fashion, relying largely on matched/mismatched status of labels. These results suggest that category labels play a pivotal role in the abstraction of category members. I propose that category labels are used to organise category information, and facilitate high-order cognitive processes such as structural alignment.

Several studies compared classification and feature inference in the context of category learning (Yamauchi and Markman 2000b, Markman and Maddox 2003). They reported a similar disadvantage in classification learning. The inductive potential of verbal labels was also reported in developmental studies (Gelman and Markman 1986; Sloutsky and Fisher 2004). The results from the present study extend their studies by demonstrating how structural alignment can be applied in the abstraction of category members and how category labels can aid this process.

4.1. *Theoretical implications*

Why is classification disadvantageous in finding abstract commonalties? Classification judgments may be made primarily based on an inexpensive recognition heuristic (Ross et al. 1990; Gigerenzer and Goldstein 1996). For example, Ross et al. (1990) showed that being reminded of category exemplars is a key variable for classification judgments. Because of the frequent use of a recognition heuristic, the classification task may be particularly sensitive to the exact appearance of category members, making it inflexible in dealing with abstract commonalties. In line with this idea, Brooks and colleagues (2000) present a stunning demonstration that even medical experts with an average of 10 years of experience fail to identify clinical signs without proper diagnostic information. In their study, specialists in internal medicine were given face-to-shoulder photographs of patients. These photographs displayed clear-cut clinical features related to particular ailments. However, without additional pathological class information, the expert clinicians were unable to identify these visual signs in nearly 40% of the cases. This is surprising because many of these photographs were taken from medical textbooks, and most of the pathological features that were depicted in the photographs were 'typical' and 'clear-cut'.

Why are category labels useful for forming structured representations and structural alignment? Research in analogical transfer and problem solving showed that finding structural commonalties between concepts is not automatic; it often requires domain-specific background knowledge and expertise (Gick and Holyoak 1980, 1983; Chi, Feltovich and Glaser 1981). I argue that mere reference to category labels is sufficient to provide background knowledge and expectation about how members of a category are organised (Brown 1957; Markman and Hutchinson 1984; Gelman and Markman 1986; Flagnnagan, Fried and Holyoak 1986; Markman 1989; Gelman and Heyman 1999; Yamauchi and Markman 2000a; Yamauchi et al. 2002; Pothos et al. 2004; Yamauchi 2005; Yamauchi and Yu 2008; but also see Sloutsky 2003; Sloutsky and Fisher 2004 for an opposing view). Studies have shown that people have an implicit idea about how attributes of a category are interrelated (Sloman, Love and Ahn 1998), and use that knowledge to make a classification or feature inference (Rehder and Hastie 2001). In other words, mere reference to 'category labels' is likely to create a predisposition to seek structural commonalties among category members. This predisposition may help form structured representations.

Finally, the findings from the present study have important implications into the general issue of abstraction and generalisation. The issue of abstraction and generalisation is one of the richly studied yet most controversial topics in cognitive science. Stimulus generalisation (Shepard 1987), category learning (Love et al. 2004), case-based reasoning in artificial intelligence (Mitchell 1997), and artificial grammar learning (Vokey and Brooks 1992) all speak to the agenda relevant to abstraction and generalisation of stimuli. These research programs propose memory-intensive statistical learning as a key mechanism for abstraction. In this view, abstraction and generalisation result from the memory store that accumulates a large number of stimulus information. For example, in Hintzman's MINERVA 2 (1986), the abstraction of stimulus information arises from accumulated memory 'traces' of individual instances. In Love et al.'s category learning model (Love et al. 2003), the abstraction of category members occurs due to the incremental adjustments of exemplar information. In this sense, abstraction and generalisation are a gradual and incremental process, interleaving new stimulus information into the existing knowledge base (McClelland, McNaughton and O'Reilly 1995; Rogers and McClelland 2004).

The present study showed that such a lengthy accumulation process is not always necessary for abstraction to happen. In Experiments 1 and 2, participants made judgments on the basis of a single sample stimulus, rather than a large number of multiple stimuli. For the discovery of abstract commonalties, memory-based statistical learning is probably not very effective. In conjunction with some appropriate cognitive strategies, such as structural alignment, this lengthy process can be shortened, and abstract commonalties can be identified relatively easily.

4.2. *Limitations*

Up to this point, I have used the terms 'category information', 'category membership' or 'category labels' without clearly defining them. People are flexible enough to form almost any category. For example, 'penguin' signifies a group of animals, rather than a specific instance of an animal. The same group of animals can be 'categorised' by a different criterion, such as 'bird-living-in-the-South-Pole'. It is unclear if the effect of category labels observed in this study can be extended to these *ad hoc* categories. At this point, I limit my definition of 'category labels' to noun labels that are used as names of natural kinds.

4.3. *Conclusion*

The effectiveness of structure-mapping/alignment in abstraction and generalisation of concepts has been investigated in other research programs. These studies include forming analogies (Holyoak and Thagard 1995; Gentner and Markman 1997; Hummel and Holyoak 1997), making similarity judgments (Markman and Gentner 1993; Medin et al.

1983; Goldstone 1996), decision making (Markman and Medin 1995, 2002; Zhang and Markman 2001), inductive inference (Lassaline 1996; Yamauchi and Markman 2000b; Hummel and Holyoak 2003), and category learning (Lassaline and Murphy 1998). The results from the present study extend these findings by showing that structural alignment can be also at work in finding abstract commonalties among category members. The present study further suggests that category information, which is strengthened by verbal labels, can promote this process.

Acknowledgements

This research was supported by the College Faculty Research Enhancement Award, the Glasscock Center Faculty Fellow Award, and the Developmental Grant by the Mexican American and U.S. Latino Research Center (Texas A&M University). The author would like to thank Art Markman, Na-Yung Yu and Wookyoung Jung for their helpful comments.

References

Anderson, J.R. (1990), *The Adaptive Character of Thought*, Hillsdale, NJ: Lawrence Erlbaum.

Ashby, F.G., and Lee, W.W. (1991), 'Predicting Similarity and Categorization from Identification', *Journal of Experimental Psychology: General*, 120, 150–172.

Brooks, L.R., and Hannah, S.D. (2006), 'Instantiated Features and the Use of Rules', *Journal of Experimental Psychology: General*, 135, 133–151.

Brooks, L.R., LeBlanc, V.R., and Norman, G.R. (2000), 'On the Difficulty of Noticing Obvious Features in Patient Appearance', *Psychological Science*, 11, 112–117.

Brown, R. (1957), 'Linguistic Determinism and the Part of Speech', *Journal of Abnormal and Social Psychology*, 55, 1–5.

Chi, M.T.H., Feltovich, P.J., and Glaser, R. (1981), 'Categorization and Representation of Physics Problems by Experts and Novices', *Cognitive Science*, 5, 121–152.

Falkenhainer, B.F., Forbus, K.D., and Gentner, D. (1989), 'The Structural Mapping Engine: Algorithm and Example', *Artificial Intelligence*, 41, 1–63.

Flagnnagan, M.J., Fried, L.S., and Holyoak, K.J. (1986), 'Distributional Expectations and the Induction of Category Structure', *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 241–256.

Gelman, S. (2003), *The Essential Child: Origins of Essentialism in Everyday Thought*, New York: Oxford University Press.

Gelman, S.A., and Coley, J.D. (1990), 'The Importance of Knowing a Dodo is a Bird: Categories and Inference in 2-year-old Children', *Developmental Psychology*, 26, 796–804.

Gelman, S., and Heyman, G.D. (1999), 'Carrot-eaters and Creature-Believers: The Effects of Lexicalization on Children's Inferences About Social Categories', *Psychological Science*, 10, 489–493.

Gelman, S.A., and Markman, E.M. (1986), 'Categories and Induction in Young Children', *Cognition*, 23, 183–209.

Gentner, D. (1983), 'Structure-mapping: A Theoretical Framework for Analogy', *Cognitive Science*, 7, 155–170.

Gentner, D., and Markman, A.B. (1997), 'Structure Mapping in Analogy and Similarity', *American Psychologist*, 52, 45–56.

Gentner, D., and Medina, J. (1998), 'Similarity and the Development of Rules', *Cognition*, 65, 263–297.

Gentner, D., and Namy, L.L. (1999), 'Comparison in the Development of Categories', *Cognitive Development*, 14, 487–513.

Gick, M.L., and Holyoak, K.J. (1980), 'Analogical Problem Solving', *Cognitive Psychology*, 12, 306–355.

Gick, M.L., and Holyoak, K.J. (1983), 'Schema Induction and Analogical Transfer', *Cognitive Psychology*, 15, 1–38.

Gigerenzer, G., and Goldstein, D.G. (1996), 'Reasoning the Fast and Frugal Way: Models of Bounded Rationality', *Psychological Review*, 103, 650–669.

Goldstone, R.L. (1996), 'Alignment-based Nonmonotonicities in Similarity', *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 988–1001.

Goldstone, R.L., Medin, D.L., and Gentner, D. (1991), 'Relational Similarity and the Nonindependence of Features in Similarity Judgments', *Cognitive Psychology*, 23, 222–262.

Hintzman, D.L. (1986), 'Schema Abstraction in a Multiple-trace Memory Model', *Psychological Review*, 93, 411–428.

Holyoak, K.J., and Thagard, P. (1995), *Mental Leaps*, Cambridge, MA: MIT Press.

Hummel, J.E., and Holyoak, K.J. (1997), 'Distributed Representations of Structure: A Theory of Analogical Access and Mapping', *Psychological Review*, 104, 427–466.

Hummel, J., and Holyoak, K.J. (2003), 'A Symbolic-Connectionist Theory of Relational Inference and Generalization', *Psychological Review*, 110, 220–264.

Johansen, M.K., and Kruschke, J.K. (2005), 'Category Representation for Classification and Feature Inference', *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 1433–1458.

Kruschke, J.K. (1992), 'ALCOVE: An Exemplar-Based Connectionist Model of Category Learning', *Psychological Review*, 99, 22–44.

Lassaline, M.E. (1996), 'Structural Alignment in Induction and Similarity', *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 754–770.

Lassaline, M.E., and Murphy, G.L. (1998), 'Alignment and Category Learning', *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 144–160.

Love, B.C., Medin, D.L., and Gureckis, T.M. (2004), 'SUSTAIN: A Network Model of Category Learning', *Psychological Review*, 111, 309–332.

Markman, A.B. (1999), *Knowledge Representation*, New Jersey: Lawrence Erlbaum.

Markman, A.B., and Gentner, D. (1993a), 'Structural Alignment During Similarity Comparisons', *Cognitive Psychology*, 25, 431–467.

Markman, A.B., and Gentner, D. (1993b), 'Splitting the Differences: A Structural Alignment View of Similarity', *Journal of Memory and Language*, 32, 517–535.

Markman, A.B., and Maddox, W.T. (2003), 'Classification of Exemplars with Single and Multiple Feature Manifestations: The Effects of Relevant Dimension Variation and Category Structure', *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 107–117.

Markman, A.B., and Medin, D.L. (1995), 'Similarity and Alignment in Choice', *Organizational Behavior and Human Decision Processes*, 63, 117–130.

Markman, A.B., and Medin, D.L. (2002), 'Decision Making', in *Stevens Handbook of Experimental Psychology* (Vol. 2, 3rd ed.) eds. D.L. Medin and H. Pashler, New York: John Wiley and Sons, pp. 413–466.

Markman, A.B., and Ross, B. (2003), 'Category Use and Category Learning', *Psychological Bulletin*, 129, 592–613.

Markman, A.B., and Wisniewski, E.J. (1997), 'Similar and Different: The Differentiation of Basic Level Categories', *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23, 54–70.

Markman, E.M. (1989), *Categorization and Naming in Children: Problems of Induction*, Cambridge, MA: MIT Press.

Markman, E.M., and Hutchinson, J.E. (1984), 'Children's Sensitivity to Constraints on Word Meaning: Taxonomic Versus Thematic Relations', *Cognitive Psychology*, 16, 1–27.

McClelland, J.L., McNaughton, B.L., and O'Reilly, R.C. (1995), 'Why there are Complementary Learning Systems in the Hippocampus and Neocortex: Insights from the Successes and Failures of Connectionist Models of Learning and Memory', *Psychological Review*, 102, 419–457.

Medin, D.L., and Atran, S. (2004), 'The Native Mind: Biological Categorization and Reasoning in Development and Across Culture', *Psychological Review*, 111, 960–983.

Medin, D.L., Dewey, G.I., and Murphy, T.D. (1983), 'Relationships Between Item and Category Learning: Evidence that Abstraction is Not Automatic', *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9, 607–625.

Medin, D.L., Goldstone, R.L., and Gentner, D. (1993), 'Respects for Similarity', *Psychological Review*, 100, 254–278.

Medin, D.L., and Schaffer, M.M. (1978), 'Context Theory of Classification', *Psychological Review*, 85, 207–238.

Mitchell, T.M. (1997), *Machine Learning*, New York: McGraw Hill.

Murphy, G.L. (2002), *The Big Book of Concepts*, Cambridge, MA: MIT Press.

Murphy, G.L., and Ross, B.H. (1994), 'Predictions from Uncertain Categorizations', *Cognitive Psychology*, 27, 148–193.

Nosofsky, R.M. (1986), 'Attention, Similarity, and the Identification-Categorization Relationship', *Journal of Experimental Psychology: General*, 115, 39–57.

Posner, M.I., and Keele, S.W. (1968), 'On the Genesis of Abstract Ideas', *Journal of Experimental Psychology*, 77, 353–363.

Pothos, E.M. (2005), 'The Rules versus Similarity Distinction', *Behavioral and Brain Sciences*, 28, 1–49.

Pothos, E.M., Chater, N., and Stewart, A.J. (2004), 'Information About the Logical Structure of a Category Affects Generalization', *British Journal of Psychology*, 95, 371–386.

Rehder, B., and Hastie, R. (2001), 'Causal Knowledge and Categories: The Effects of Causal Beliefs on Categorization, Induction, and Similarity', *Journal of Experimental Psychology: General*, 130, 323–360.

Rogers, T.T., and McClelland, J.L. (2004), *Semantic Cognition: A Parallel Distributed Processing Approach*, Cambridge, MA: MIT Press.

Rosch, E. (1978), 'Principles of Categorization', in *Cognition and Categorization*, eds. E. Rosch and B.B. Lloyd, Hillsdale, NJ: Lawrence Erlbaum, pp. 27–48.

Rosch, E., and Mervis, C.B. (1975), 'Family Resemblances: Studies in the Internal Structure of Categories', *Cognitive Psychology*, 7, 573–605.

Ross, B.H., Perkins, S.J., and Tenpenny, P.L. (1990), 'Reminding-Based Category Learning', *Cognitive Psychology*, 22, 460–492.

Shepard, R.N. (1987), 'Toward a Universal Law of Generalization for Psychological Science', *Science*, 237(4820), 1317–1323.

Sloman, S.A. (1993), 'Feature-based Induction', *Cognitive Psychology*, 25, 231–280.

Sloman, S.A., Love, B.C., and Ahn, W.K. (1998), 'Feature Centrality and Conceptual Coherence', *Cognitive Science*, 22, 189–228.

Sloutsky, V.M. (2003), 'The Role of Similarity in the Development of Categorization', *Trends in Cognitive Sciences*, 7, 246–558.

Sloutsky, V.M., and Fisher, A.V. (2004), 'Induction and Categorization in Young Children', *Journal of Experimental Psychology: General*, 133, 166–188.

Smith, E.E. (1994), 'Concepts and Categorization', in *An Invitation to Cognitive Science*, eds. E.E. Smith and D.N. Osherson, Cambridge, MA: MIT Press, pp. 3–33.

Tenenbaum, J.B. (1999), 'Bayesian Modeling of Human Concept Learning', in *Advances in Neural Information Processing Systems*, eds. M.S. Keams, S.A. Solia and D.A. Cohn, Cambridge, MA: MIT Press, pp. 59–65.

Tenenbaum, J.B., and Griffiths, T.L. (2001), 'Generalization, Similarity, and Bayesian Inference', *Behavioral and Brain Sciences*, 24, 629–640.

Vokey, J.R., and Brooks, L.R. (1992), 'Salience of Item Knowledge in Learning Artificial Grammars', *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 328–344.

Wattenmaker, W.D. (1995), 'Knowledge Structures and Linear Separability: Integrating Information in Object and Social Organization', *Cognitive Psychology*, 28, 274–328.

Wisniewski, E.J., and Medin, D.L. (1994), 'On the Interaction of Theory and Data in Concept Learning', *Cognitive Science*, 18, 221–281.

Yamauchi, T. (2005), 'Labeling Bias and Categorical Induction: Generative Aspects of Category Information', *Journal of Experimental Psychology: Learning, Memory and Cognition*, 31, 538–553.

Yamauchi, T., Kohn, N., and Yu, N. (2007), 'Tracking Mouse Movement in Feature Inference: Category Labels are Different from Feature Labels', *Memory and Cognition*, 35, 852–863.

Yamauchi, T., Love, B.C., and Markman, A.B. (2002), 'Learning Non-linearly Separable Categories by Inference and Classification', *Journal of Experimental Psychology: Learning, Memory and Cognition*, 28, 585–593.

Yamauchi, T., and Markman, A.B. (1998), 'Category-learning by Inference and Classification', *Journal of Memory and Language*, 39, 124–148.

Yamauchi, T., and Markman, A.B. (2000a), 'Inference using Categories', *Journal of Experimental Psychology: Learning, Memory and Cognition*, 26, 776–795.

Yamauchi, T., and Markman, A.B. (2000b), 'Learning Categories Composed of Varying Instances: The Effect of Classification, Inference and Structural Alignment', *Memory and Cognition*, 28, 64–78.

Yamauchi, T. and Yu, N. (2008), 'Category Labels versus Feature labels: Category Labels Polarize Inferential Predictions', *Memory and Cognition*, 36(3), 567–577.

Yu, N. and Yamauchi, T. (2006), 'Selective Attention or Structured Representation? The Effect of Category Information in Similarity Judgment', in *Paper Presented at the 28th Annual Meeting of the Cognitive Science Society*, Mahwah, NJ: Lawrence Erlbaum, p. 2636.

Zhang, S., and Markman, A.B. (2001), 'Processing Product-unique Features: Alignment and Involvement in Preference Construction', *Journal of Consumer Psychology*, 11, 13–27.