BACKGROUND KNOWLEDGE, CATEGORY LABELS, AND SIMILARITY JUDGMENT

A Dissertation

by

NA YUNG YU

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 2010

Major Subject: Psychology

BACKGROUND KNOWLEDGE, CATEGORY LABELS,

AND SIMILARITY JUDGMENT

A Dissertation

by

NA YUNG YU

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Takashi Yamauchi
Committee Members,	Ricardo Gutierrez-Osuna
	Steve Smith
	Jyotsna Vaid
Head of Department,	Les Morey

August 2010

Major Subject: Psychology

ABSTRACT

Background Knowledge, Category Labels, and Similarity Judgment. (August 2010)

Na Yung Yu, B.A.; M.A., Pusan National University, Korea

Chair of Advisory Committee: Dr. Takashi Yamauchi

Labels are one source of our judgments. By assigning labels to objects, we not only create references but we also group prior and current experiences together. The goal of this research is to investigate how labels influence our judgments. Previous research on inductive generalization shows that labels can be more important than physical characteristics (the labeling effect), but the mechanism for this effect remains unclear. There are two differing views regarding the role of labels. One view proposes that labels are not essentially different from physical features: shared labels increase overall similarity between two items in the same way as shared physical features. The other view suggests that people have a naïve theory that shared labels are more special than shared physical features. The goal of this dissertation is to provide evidence that complements these conflicting views. I suggest that the role of labels varies depending on the background knowledge: types of categories (living things vs. man-made objects), amount of knowledge (number of exemplars people could list for the category), and types of labels (categorical vs. indexical). The results from four experiments showed that, for living things, the labeling effect is strong and depends less on the amount of

knowledge; for man-made objects, the labeling effect is weak and depends on the amount of knowledge.

ACKNOWLEDGMENTS

I wish to thank everyone for a rewarding graduate school experience. I am grateful to my committee members, Ricardo Gutierrez-Osuna, Steve Smith and Jyotsna Vaid, for providing fascinating courses and insightful comments. I am in great debt to my ever patient advisor, Takashi Yamauchi, who sincerely and persistently helped me to get through this course of journey. Without his guidance, I would never have arrived this far. I also thank my adopted family, Sue and Malcolm, for their generosity and kindness, and Jenna for her enthusiasm and curiosity in research. To Chaitra, I send my deepest thanks for walking the path with me as a friend and a comrade. Finally my deepest appreciation goes to my parents and siblings for their unconditional faith and love. I owe this degree to all of them.

TABLE OF CONTENTS

Page

LIST OF FIGURES

FIGURE Page

INTRODUCTION

The ability to generalize is one of the most important aspects of human cognition. Research suggests that similarity (Sloman, 1993), category membership (Osherson, Smith, Wilkie, López, & Shafir, 1990), and previous experiences (Heit, 1998, 2000) influence inductive generalization. For example, when two objects share many features, we project an attribute of one stimulus to the other (Rips, 1975). When two items are in the same category, we also infer that stimuli would have features in common (Yamauchi & Markman, 2000). Prior experiences and background knowledge are also used in inductive generalization. If items have been previously observed together, we assume that they will occur together in the future. These assumptions guide our inductive learning and allow for coherent and ad-hoc explanations about the world (Murphy & Medin, 1985; Rehder & Hastie, 2004; Sloman, 1994). For example, certain crimes such as assault or robbery typically happen in dark alleys. Therefore, people associate danger with dark settings, and this assumption leads them to avoid walking alone at night.

One important issue in inductive generalization is the role of shared category labels. When two objects carry the same label, we tend to assume that these objects have some important characteristics in common; when objects carry different labels, we tend to presume that they have some distinctive characteristics (i.e., labeling effect). Much

---

This dissertation follows the style of *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

research has suggested that the strong tendency to use labels for generalization emerges early in infancy (DeLoache, 2004; Keates & Graham, 2009; Lupyan, Rakison, & McClelland, 2007; Waxman & Braun, 2005; Waxman & Markow, 1998) and spans across cultures (López, Atran, Coley, Medin & Atran, 2004; Roberson, Davies, & Davidoff, 2000; Saalbach & Imai, 2007).

How do labels influence our judgments? One view suggests that categorical labels (i.e., labels that present categories of instances) have a special status in inductive generalization (Gelman, 2003; Gelman & Markman, 1986; Murphy, 2003; Yamauchi & Markman, 2000; Yamauchi, 2009; Yamauchi & Yu, 2008): people believe that shared categorical labels indicate groups of things (Gelman & Markman, 1986; Waxman & Braun, 2005) and that categorical labels are qualitatively different from physical features (Yamauchi & Markman, 2000). The other view argues that there is nothing special about the labels (Anderson, 1990; Sloutsky & Fisher, 2004; Sloutsky, Kloos, & Fisher, 2007): the matching categorical labels contribute to the overall similarity in the same way as matching physical features. Evidence from both views demonstrates the strong tendency to use categorical labels in inductive generalization; however, they differ on the extent to which these labels are used in inductive generalization.

In this dissertation, I investigate the role of category labels in inductive generalization. The goal is to incorporate background knowledge and complement these two opposing views. In four experiments, I manipulated three types of knowledge from which the labeling effect may originate– types of categories (e.g., living things and man-made objects, Warrington & Shallice, 1984), the amount of knowledge (how much people know about a given category; Heit, 1998, 2000) and types of labels (labels for category membership vs. labels for single items; Rhemtulla & Hall, 2008; Waxman & Markow, 1995). I propose that the labeling effect stems from the interaction among these three factors.

In the following section, I briefly review two major claims regarding the role of labels in inductive generalization. Then, I review previous research on the role of background knowledge in inductive generalization. This review will show that the link between category labels and background knowledge requires further exploration.

Role of Labels in Inductive Generalization

Why do people use labels to make inductive judgments? The naïve theory view states that people tend to make a naïve assumption that category labels are qualitatively different from physical features (Gelman, 2003; Gelman & Markman, 1986; Gelman & Waxman, 2007; Keates & Graham, 2009; Lupyan, 2008; Murphy, 2003; Waxman & Braun, 295; Yamauchi, 2005; Yamauchi, Kohn & Yu, 2007; Yamauchi & Markman, 2000a; Yamauchi & Yu, 2008; see also Markman & Ross, 2003 for discussion). A wide range of developmental studies showed that when children were presented with physically dissimilar stimuli that shared the same label, they assumed that similar features existed among those stimuli (Gelman & Coley, 1990; Gelman & Heyman, 1999; Jaswal, 2007; Keates & Graham, 2009; Özçalişkan, Goldin-Meadow, Gentner, & Mylander, 2009; Rhemtulla & Hall, 2008; Waxman & Booth, 2001; Waxman & Markow, 1995). For example, Gelman and Markman (1986) investigated children's inductive judgments while physical similarity of stimuli was pitted against category membership, which was depicted by having the same category labels (Figure 1a). Participants in their study were presented with three line drawings: a target at the bottom and two base pictures at the top (Figure 1a). One base picture was dissimilar in appearance to the target, but shared the same label ("fish") as the target. The other base picture was similar to the target but was shown with a different label ("dolphin"). Four to five-year-old children judged that the target and the dissimilar base would breathe in the same way because they share the label "fish." This result suggests that young children base their inductive generalization on shared category labels without relying on physical features.

Adult studies further demonstrated that people rely on shared category labels when the shared category labels indicate shared category membership (Yamauchi, 2009; Yamauchi et al., 2007; Yamauchi & Markman, 2000a; Yamauchi & Yu, 2008). In their study, the matching label (e.g., monek) indicated either category membership (e.g., both are monek) or feature information (e.g., both eat monek). Adult participants used the label that represented category membership more than the labels that indicated feature information. Their further analysis showed that participants viewed the shared category labels more often than physical features, and the participants viewed the labels in earlier stages of inductive generalization (Yamauchi, Kohn & Yu, 2007). Studies have demonstrated that category membership information is important in inductive generalization (Heit & Heyes, 2005; Kaminski, Sloutsky, & Heckler, 2008; Malt, Ross & Murphy, 1995; Murphy & Ross, 1994; Ross & Murphy, 1996), which may explain

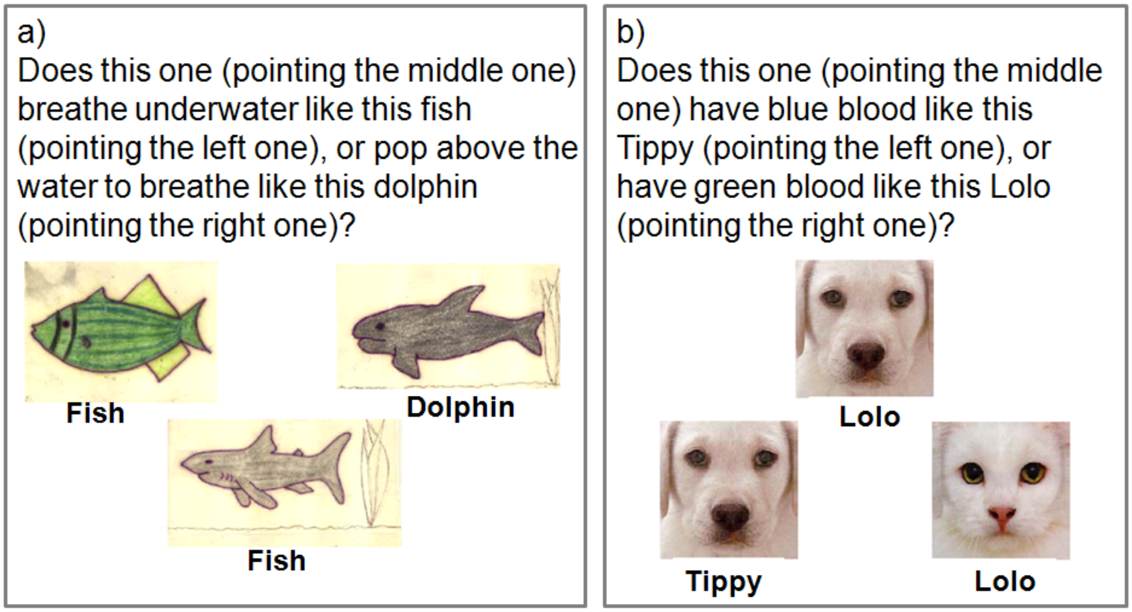


Figure 1. Sample trials from previous studies. (a) Sample trial from the study by Gelman and Markman (1986); (b) sample trial from the study by Sloutsky and Fisher (2004a).

why people first categorize stimuli based on shared category labels.

The other dominant view, however, argues that there is no indication that people have such a naïve theory about category labels (Anderson, 1990; Fisher & Sloutsky, 2005; Napolitano & Sloutsky, 2004; Robinson & Sloutsky, 2004; Sloutsky & Fisher, 2004a, 2004b; Sloutsky, Kloos, & Fisher, 2007; Sloutsky & Spino, 2004). According to this view, referred to here as the similarity-based approach, matching labels increase the perception of overall similarity just like matching physical features. For example, when two pictures have the matching physical features (e.g., both animals are white), people perceive the two pictures as similar (Tversky, 1977). Likewise, when the pictures carry the matching labels (e.g., both are called "Lolo"), people perceive the two pictures as similar.

Sloutsky and Fisher (2004a) adopted a triad task similar to the one used by Gelman and Markman (1986) and corroborated this claim. In their study, 4 to 5-years-olds were presented with a target on the top and two base pictures at the bottom using labels that were pitted against the physical similarity (Figure 1b). Participants judged whether the target was more likely to have blue blood, like the similar base picture with the different label, or green blood, like the dissimilar base picture with the same label. They found that children judged that the target would have green blood like the dissimilar base picture because it shared the label "Lolo." Sloutsky and Fisher created a computational model where labels and physical features did not qualitatively differ; though they assigned larger attention weights for the labels than for the physical features (but also see Heit & Hayes, 2005). They argue that labels are not special because their model shows that people evaluate matching labels as a type of matching feature that increases the overall similarity of items.

In both the naïve theory and the similarity-based approaches, category labels are expected to influence inductive generalization, but the naïve theory approach specifies that labels are used to form assumptions, while the similarity-based approach suggests that labels are used as a type of feature. In the following section, I review the research on three types of background knowledge and suggest that this knowledge can complement the naïve theory approach and the similarity-based approach.

Research on Three Types of Background Knowledge

Research has shown the influence of background knowledge in many aspects of inductive generalization (Ahn, 1998; Coley, Hayes, Lawson & Moloney, 2004; Heit, 1998, 2000; Heit & Rubinstein, 1994; Heyes & Thompson, 2007; Keil, 1989; Medin, Coley, Storms, & Hayes, 2003; Murphy & Medin, 1985; Rehder, 2006; Wattenmaker, Dewey, Murphy, & Medin, 1986; Wisniewski & Medin, 1994). The goal of this study is to address the link between background knowledge and the labeling effect.

Types of categories. Studies of patients suffering from semantic dementia suggest that people may organize background knowledge based on the types of categories such as living things versus man-made objects. Patients who had bilateral damage to the temporal limbic system often could not recognize pictures of living things (e.g., animals, plants), but they could name pictures of man-made objects (e.g., tools, towels) (Damasio, Grabowski, Tranel, Hichwa & Damasio, 1996; Hillis & Caramazza, 1991; Warrington & Shallic, 1984). The opposite pattern of dissociation for these two types of categories was also observed for patients with damage to the frontoparietal region. These patients could not recognize pictures of man-made objects, whereas they could name living things (Hillis & Caramazza, 1991; Sacchett & Humphreys, 1992; Warrington & McCarthy, 1983). This dissociation suggests that living things and man-made objects are organized differently, each involving separate brain regions (for review, see Caramazza & Shelton, 1998; Moss, Tyler, Durrant-Peatifield & Bunn, 1998).

The distinction between living things and man-made objects is also shown in developmental research. Children's understanding of living things relies on internal features (e.g., DNA and maturation), while their understanding of man-made objects relies on external features (e.g., physical appearances; Barton & Komatsu, 1989; Carey, 1985; Chao, Chen, Roisman, & Hong, 2007; Hong, Chao, & No, 2009; Keil, 1989; Medin & Ortony, 1989; Murphy & Medin, 1985; Rips, 1989; Wattenmaker, 1995; for review, see Medin, Lynch & Solomon, 2000; Rips, 2001). For example, participants in the Rips (1989) study were presented with a story of an animal called "sob" that originally looked like a bird but changed to look like an insect. When the transformation in appearance was explained as a result of maturation (internal feature), they thought that the transformed animal was no longer a sob. In a control condition, when the transformation in appearance was due to chemical hazards, participants judged that the transformed animal was still a sob. Physical appearance was considered an important factor for judgments about living things only when it was caused by internal features. These results suggest that, for living things, physical appearances are not always considered important.

Amount of knowledge. Although studies have showed that experts form inductive generalizations differently than novices (e.g., Johson & Mervis, 1997), the amount of knowledge that novices have has not been well explored in the research on the labeling effect. The main reason is the difficulty of quantifying "background knowledge." In the similarity-based approach, similarity between items is quantitatively measured by counting the number of matching and mismatching features between two items (Tversky, 1977), and the measured similarity is used to explain the pattern of inductive generalization (e.g., Sloutsky & Fisher, 2004a; Osherson et al., 1990). Although the models of inductive generalization have tried to quantify background knowledge, the knowledge was measured in terms of similarity (Heit, 1998, 2000; McDonald, Samuels, & Rispoli, 1996; Smith, Shafir, & Oshersonat, 1993; Tenenbaum & Griffiths, 2001). In Heit's Bayesian model of inductive generalization (1998, 2000), for example, prior knowledge of two items was measured based on the matching/mismatching features that people already knew for the items, and this measured prior knowledge was used to explain the pattern of inductive generalization. Since background knowledge was measured in the same way as similarity, a method that quantifies background knowledge differently than similarity should be tested.

One way to measure the amount of background knowledge is to ask participants to list as many exemplars of a category as possible (Barsalou, 1985; Smith, Ward, Tindell, Sifonis, & Wilkenfeld, 2000; Ward, Patterson, Sifonis, Dodds, & Saunders, 2002). Smith and colleagues (2000) asked a group of participants to list as many exemplars that they could think of for 10 different categories (e.g., birds). The researchers also asked a second group of participants to study words (e.g., eagle, flamingo) that belong to the same 10 categories and then recall the words they studied. Smith et al found that people falsely recollected nonpresented items (e.g., sparrow) that were also the exemplars that the first group of participants frequently listed. Their study demonstrates that the frequently listed members have greater accessibility within the category, and this conceptual knowledge influences false memory. Given that the amount of background knowledge influences false memory, it is also possible that the amount of background knowledge influences inductive generalizations.

Types of labels. Studies suggest that categorical labels (labels representing category membership) guide inductive generalization even when they are pitted against physical similarity (Gelman & Coley, 1990; Özçalişkan et al., 2009; Rhemtulla & Hall, 2008; Waxman & Booth, 2001; Waxman & Markow, 1995; Yamauchi, 2005; Yamauchi et al., 2007; Yamauchi & Markman, 2000; Yamauchi & Yu, 2008; Yu, Yamauchi, & Schumacher, 2008). Children used noun labels that depicted category membership more than adjective labels (Waxman & Booth, 2001) or verb labels (Gelman & Heyman, 1999) in inductive generalizations (Özçalişkan et al., 2009; Rhemtulla & Hall, 2008). Gelman and Heyman (1999) found that when someone was described with a noun label (e.g., Linda is a carrot eater), children assumed that the described characteristic was more stable over time and context than the characteristic described with a verb label (e.g., Linda eats carrots a lot). This results suggests that young children can distinguish between syntactic cues associated with labels (e.g., nouns and verbs) in order to understand category membership.

Given that the labeling effect only occurs for certain types of labels (e.g., categorical labels), it is problematic to compare studies that use different types of labels. Although the overall settings of the tasks in Gelman and Markman (1986) and Sloutsky and Fisher (2004a) were identical (Figure 1a and 1b), the types of labels were different. Gelman and Markman (1986) presented pictures with real animal labels such as “fish” and “dolphin.” Because participants already knew these labels, they could think of many exemplars of these categories. In contrast, Sloutsky and Fisher (2004a) showed pictures with novel, arbitrary names such as “Lolo” and “Tippy” (i.e., indexical labels). Because participants could not think of other exemplars of “Lolo” or “Tippy,” they could have relied less on these labels. The discrepancy between these two studies might be partly due to types of labels: it is possible that indexical labels are considered to be another matching feature while categorical labels are considered to be more important than matching physical features.

This review suggests that the influence of the labels can be based on three types of background knowledge: types of categories (e.g., living things vs. man-made objects), amount of knowledge (e.g., listed exemplars of a category), and types of labels (categorical labels vs. indexical labels). I hypothesize that the labeling effect is stronger for living things than for man-made objects, stronger for the categories for which people have more background knowledge, and stronger for categorical labels than for indexical labels. In the following four experiments, I manipulated these factors and tested these hypotheses.

EXPERIMENT 1: SIMILARITY JUDGMENT TASK

Experiment 1 employed a similarity judgment task to demonstrate the labeling effect with an experimental setting that has been used in previous studies (Gelman & Markman, 1986; Sloutsky & Fisher, 2004). However, different questions were used in this experiment. The questions used in the studies by Gelman and Markman (1986) and Sloutsky and Fisher (2004a) may be one reason for their different results (Figure 1). Gelman and Markman used more realistic questions, such as whether the target breathes like a fish or a dolphin; whereas Sloutsky and Fisher used less realistic questions, such as whether the target has green blood or blue blood (Figure 1). The type of questions can be one confound in these studies.

The goal of Experiment 1 is to examine the effect of labels while controlling for the type of questions. Instead of having participants answer realistic/unrealistic questions, participants in Experiment 1 were asked to judge the similarity of animal faces (Figure 2). They were presented with three animal face pictures: a target picture placed at the top and two base pictures placed at the bottom (Figure 2a). In each stimulus frame, the target was more similar looking to one base picture ("similar base picture") than to the other base picture ("dissimilar base picture") (Figure 2a). The task was to judge which base picture, left or right, is more similar to the target.

The effect of labels was examined by comparing two conditions: no-label and same-label conditions. In the no-label condition, all pictures were presented without labels (Figure 2a). In the same-label condition, the dissimilar base picture had the same animal label (e.g., "horse") as the target (Figure 2b). In both conditions, the

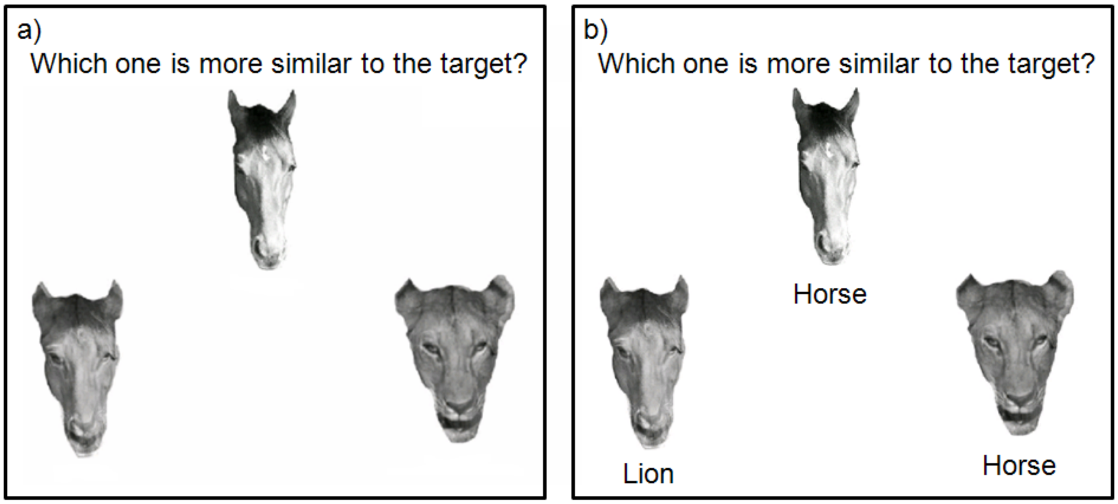


Figure 2. Sample trials in Experiment 1. In each trial, a target picture was placed at the top and two base pictures were placed at the bottom. One base picture was less similar to the target (“dissimilar base picture”). In the no-label condition (a), pictures were presented without labels. In the same-label condition (b), the target and the dissimilar base picture on the bottom right shared the same labels; the target and the other base picture on the bottom left had different labels.

proportions of participants judging the dissimilar base picture as similar to the target (i.e., the proportion of participants judging pictures with the same category labels as similar) were recorded. I hypothesize that labels will influence similarity judgments. That is, the proportion of participants judging the dissimilar base picture as similar to the target will be significantly higher in the same-label condition than in the no-label condition.

Method

Participants. One hundred one undergraduate students participated for course credit. They were randomly assigned to one of two conditions: no-label (n = 52) and same-label (n = 49).

Materials. In Experiment 1, participants received triads of animal face pictures (Figure 2). To create stimuli, five pairs of original animal face pictures were selected. Each pair of original animal faces was merged using MorphMan 4.0 (2003) software. Altogether, 90 morphed pictures were created from five sets of original pictures (18 degrees of morphed pictures for each of five pairs; see Appendix A for samples of morphed stimuli).

From the morphed pictures, three levels of physical difference—low-, medium-, and high-difference—were created based on the degree of merging of the two original pictures (Figure 3). In the low-difference condition, the two base pictures were not very different (Figure 3a); in the medium-difference condition, the two base pictures were moderately different (Figure 3b); and in the high-difference condition, the two base pictures were highly different (Figure 3c). Two sets of base pictures were randomly selected at each level of physical difference and were combined with one of two original pictures in each pair, yielding 12 triads for each pair (a total of 60 triads = 5 pairs × 12 triads).

In each trial, the target was always an original picture, and the two base pictures (see Figure 2) were morphed images of two original pictures (see Figure 3).

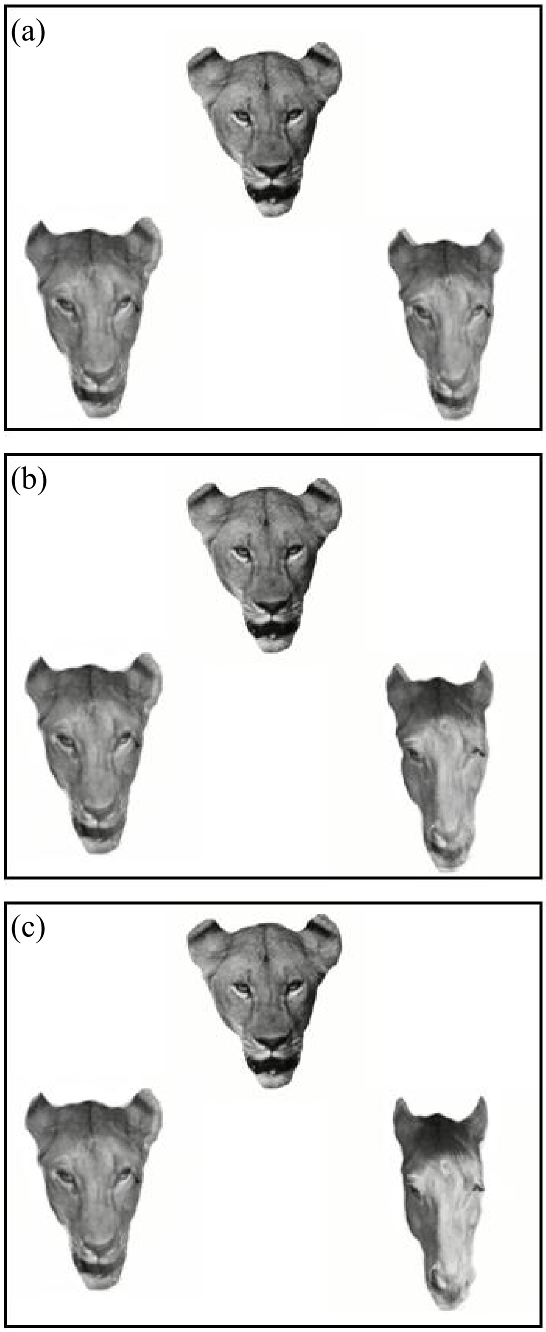


Figure 3. Three levels of physical difference. The two base pictures were very similar in the low-difference condition (a), moderately similar in the medium-difference condition (b), and very dissimilar in the high-difference condition (c).

Design. Experiment 1 had a 2 (Label Condition; no-label vs. same-label conditions; between-subjects) × 3 (Physical Difference; low-difference, medium-difference, high-difference; within-subjects) factorial design. The dependent measure was the proportion of participants selecting the dissimilar base pictures as more similar to the target than the other base pictures (Figure 2).

Procedure. Sixty triads of pictures were presented to participants one at a time in the center of the computer screen. Participants were asked to select the base picture that they judged to be more similar to the target than the other base picture. They indicated their responses by pressing the left or right arrow key on the keyboard. E-Prime 1.1 (Psychology Software Tools Inc., 2002) was used to present individual trials. The order of presented stimuli was determined randomly. The dissimilar base picture was presented on the left or the right side an equal number of times. The experiment lasted approximately 10 minutes.

Results and Discussion

Figure 4 summarizes the main results from Experiment 1. In Experiment 1, participants were asked to judge similarity of pictures to control for the effect of type of questions while testing for the labeling effect. As predicted, when labels were attached to the pictures, similarity judgments of animal face pictures changed considerably. The proportion of participants selecting the dissimilar base pictures was significantly higher in the same-label condition ($M = 0.29$) than in the no-label condition ($M = 0.15$), $F(1, 99) = 28.57$, $MSE = 0.05$, $p < .001$, $\eta^2 = .22$ (see also Appendix B).

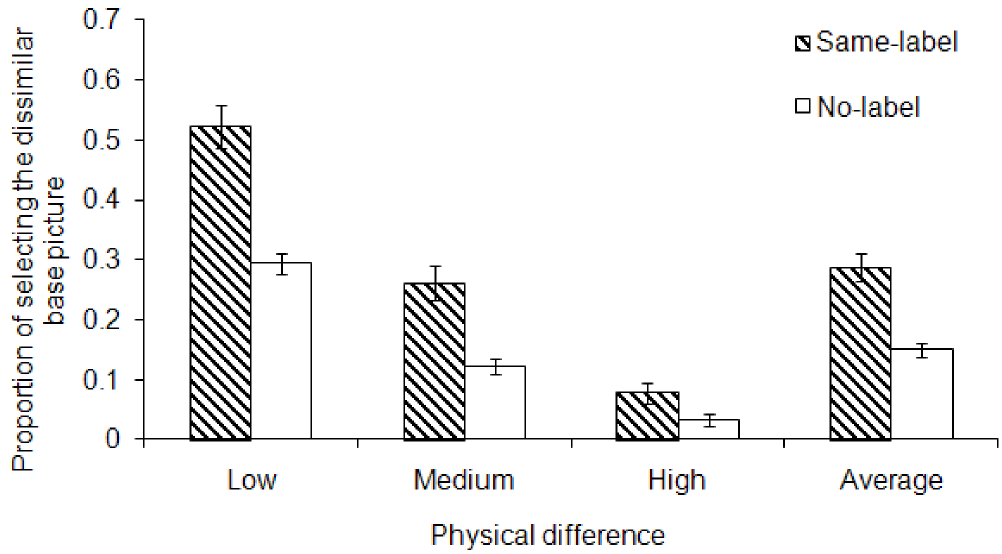


Figure 4. Mean proportions of participants selecting the dissimilar base picture (that carried the same label as the target) over the other base picture in Experiment 1. The error bars represent two standard error units calculated from each condition.

The interaction between the label condition and the physical difference was also significant, F (2, 198) = 15.43, MSE = 0.01, p < .001, η^2 = .14. Having the same label as the target increased the proportion of participants selecting the dissimilar base picture compared with the no-label condition at the low and medium level of physical difference: low-difference, t (99) = 5.73, SE = .04, p < .01, d = 1.14; medium-difference, t (99) = 4.20, SE = .04, p < .01, d = .84. However, at the high level of physical difference, the difference between the same-label condition and no-label condition did not reach the significance level after Bonferroni adjustment; high-difference, t (99) = 2.22, SE = .04, p = .03, d = .44.

The labels influenced inductive judgments when the type of question was controlled. Since all participants carried out the same similarity judgment task, the difference between the same-label condition and the no-label condition must have stemmed from the labels attached to the pictures. Given that labels alone can influence similarity judgment of pictures, Experiment 2 tested the hypothesis that the labeling effect depends on types of categories (living things vs. man-made objects). I hypothesize that category labels influence inductive generalization more for living things than for man-made objects.

EXPERIMENT 2: TYPES OF CATEGORIES

The goal of Experiment 2 is to investigate whether the labeling effect is stronger for living things than man-made objects. Studies suggest that inductive generalizations for living things do not always rely on physical appearances (Gelman, 2003; Rips, 1989), while inductive generalizations for man-made objects depend on physical appearances (Barton & Kmatsu, 1989; Carey, 1985). Thus, I predict that when labels are pitted against physical features for living things, people will use the labels more than when labels are pitted against physical features for man-made objects.

To manipulate the types of categories exclusively with labels, participants in Experiment 2 were presented with the same set of animal tissue pictures with labels representing living things or man-made objects. Since participants are unlikely to have experienced seeing animal tissue pictures (for examples, see Appendix A), and they are presented with the same set of pictures in all conditions, the background knowledge stems mostly from the labels attached to the pictures.

Figure 5 shows sample trials of Experiment 2. All subjects received the same pictures with labels, except that the meanings associated with the labels were manipulated in the instructions. In some cases, the labels represented living things (cells). In the animal cell condition, participants were told that the pictures represented animal cells and the pictures' labels described the animal cell names (e.g., "Pig bone cell" in Figure 5a). In the diseased cell condition, they were told that the pictures represented cells infected by diseases and the pictures' labels described the disease names (e.g., "Epsilersy" in Figure 5b). Thus, in both animal cell and diseased cell conditions, labels

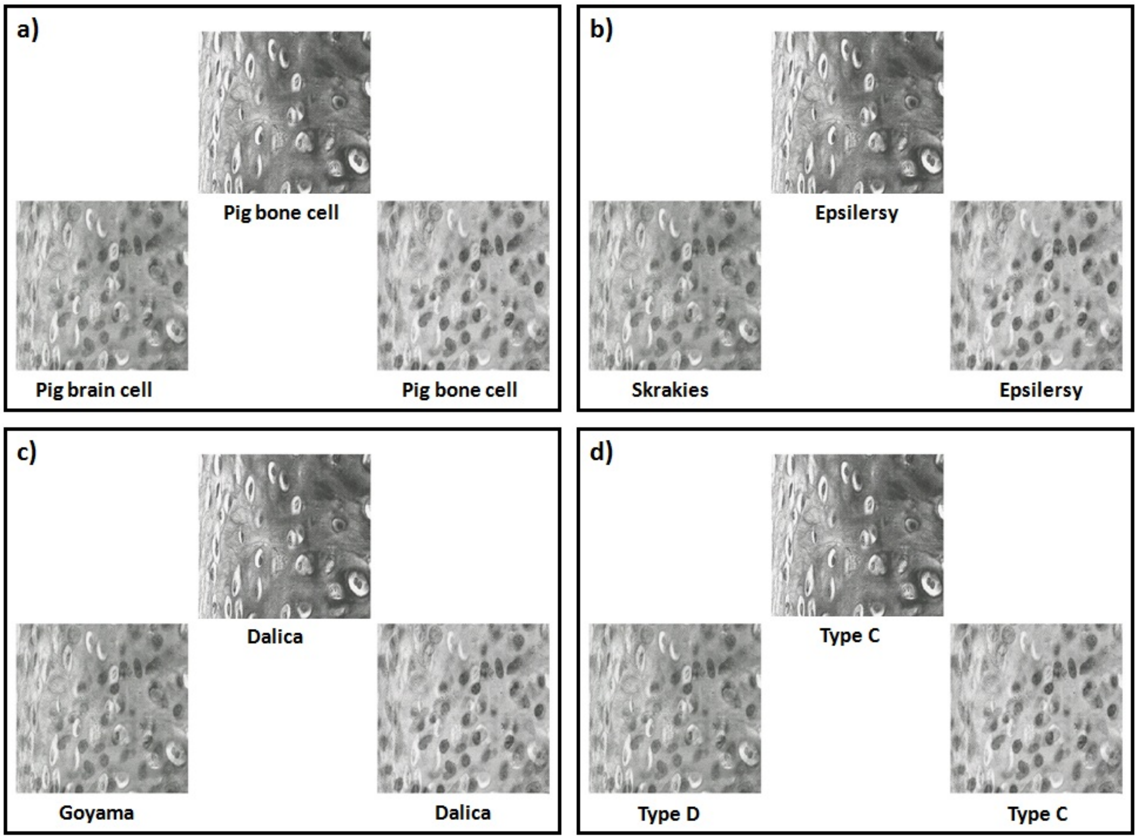


Figure 5. Sample trials shown in Experiment 2. Labels indicated animal cells in the animal cell condition (a), diseases the cells carried in the diseased cell condition (b), artists who painted the abstract paintings in the painter condition (c), or picture types in the picture type condition (d).

represented categories of living things (cells).

In the other cases, the labels represented man-made objects (paintings, pictures). In the painter condition, participants were told that the pictures represented abstract paintings, and pictures' labels described the painters' names (painter condition, e.g., "Dalica" in Figure 5c). In the picture type condition, participants were told that they would be shown pictures and the labels described the picture types (e.g., "Type D" in

Figure 5d). Thus, in both painter and picture type conditions, labels represented categories of man-made objects (paintings/pictures). The proportion of selecting dissimilar base pictures will be greater when the labels indicate animal cells and diseased cells than when the labels indicate paintings and pictures.

Method

Participants. Two hundred fifty two undergraduate students participated for course credit. They were randomly assigned to one of the five conditions: no-label (n = 48), animal cell (n = 48), diseased cell (n = 53), painter (n = 56), and picture type (n = 47).

Materials. Five pairs of original animal tissue pictures were selected from a well-known textbook of veterinary histology (Bacha & Bacha, 2000). The procedure of creating 60 triads of original and morphed pictures was identical to the procedure described in Experiment 1 (see Appendix A for samples of morphed stimuli). As in Experiment 1, the triads of pictures had three levels of physical difference (low-, medium-, and high-difference).

Design. The experiment had a 5 (Label Condition; no-label, animal cell, diseased cell, painter, picture type conditions; between-subjects) × 3 (Physical Difference; low-difference, medium-difference, high-difference; within-subjects) factorial design.

Procedure. The procedure of Experiment 2 was identical to the one described in Experiment 1. Participants were presented with 60 triads of animal tissue pictures and judged which base picture was more similar to the target.

Results and Discussion

Results from Experiment 2 are summarized in Figure 6. As predicted, animal cell labels and diseased cell labels significantly influenced similarity judgment of pictures; however, picture type labels and painter labels did not influence similarity judgment significantly. These results show the influence of types of categories on the labeling effect.

The main effect of the label condition was significant, F (4, 247) = 6.17, MSE = 0.10, $p < .001$, $\eta^2 = .09$. The proportion of participants selecting the dissimilar base pictures was significantly higher in the animal cell condition (M = 0.31) and diseased cell condition (M = 0.29) as compared to the proportion observed in the no-label condition (M = 0.17) after Bonferroni adjustment: animal cell vs. no-label condition, t (94) = 3.54, SE = .04, $p < .01$, d = 0.72, diseased cell vs. no-label condition, t (99) = 3.33, SE = .04, $p < .01$, d = 0.66. However, the proportion of participants selecting the dissimilar base pictures was statistically indistinguishable between the no-label condition (M = 0.17) and the painter condition (M = 0.18), t (102) = .42, SE = .02, p = .67, d = 0.08, as well as between the no-label condition and the picture type condition (M = 0.20), t (93) = 1.23, SE = .02, $p < .01$, d = 0.25 (see also Appendix B). There was no interaction between label condition and physical difference, F (8, 494) < 1.0.

These results suggest that the labeling effect differs based on the type of categories. When labels convey information about living things (e.g., animal cells, diseased cells), people use the labels in similarity judgments considerably. However, when labels represent information about man-made objects (e.g., paintings, pictures)

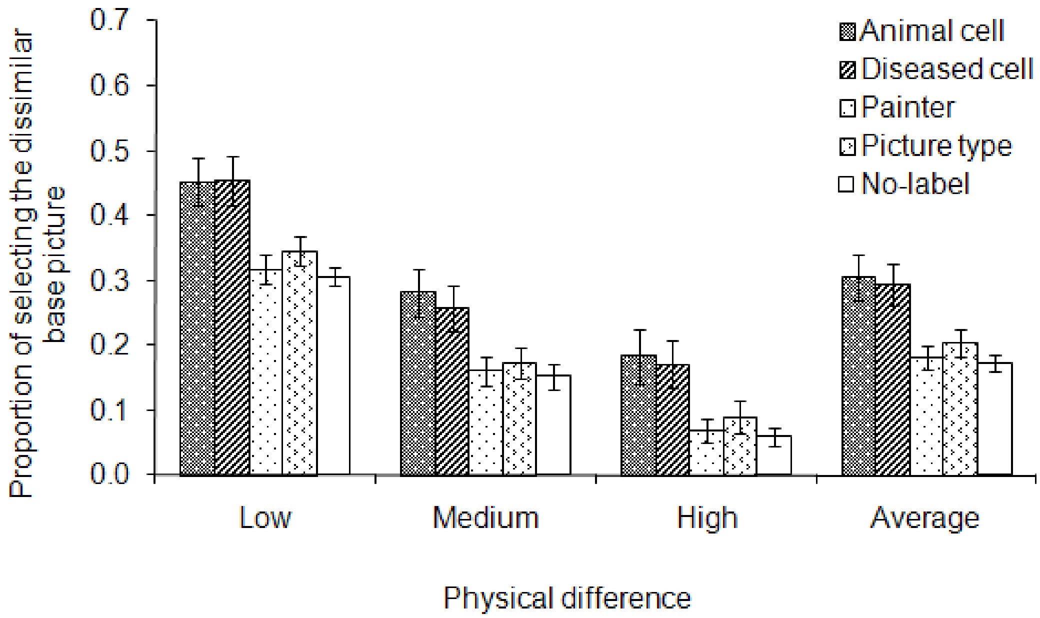


Figure 6. Mean proportions of participants selecting the dissimilar base picture over the other base picture in Experiment 2. The error bars represent two standard error units calculated from each condition.

people do not use the labels in similarity judgments. Since participants received the same set of triads of animal tissue pictures, the difference in the labeling effect observed in animal cell, diseased cell, picture type and painter label conditions originated in the types of categories.

In Experiment 3, I further examined the impact of knowledge that labels convey. Though the type of categories was different between the conditions, other factors such as the amount of knowledge and types of labels could have influenced similarity judgments (Heit, 1998; Waxman & Markow, 1995; Wisniewski & Medin, 1994). In Experiment 3, I tested the labeling effect while manipulating the amount of knowledge and types of labels in addition to the types of categories.

EXPERIMENT 3:

AMOUNT OF BACKGROUND KNOWLEDGE AND TYPES OF LABELS

The goal of Experiment 3 is to investigate the labeling effect in relation to the amount of knowledge and types of labels, in addition to types of categories. The amount of knowledge might have been involved in the labeling effect that was found in Experiment 2. The results of Experiment 2 showed that people used labels more often when they represented animal cells or diseased cells than when they represented picture types or painters. In general, however, people may carry more knowledge regarding cells versus painters. This difference in the amount of background knowledge could have influenced the labeling effect. Experiment 3 directly measured the amount of background knowledge and tested this idea.

Additionally, previous research suggests that when labels represent a group of items (category membership), people tend to use categorical labels in inductive generalization (Gelman & Heyman, 1999; Waxman & Booth, 2001; Waxman & Markow, 1995; Yamauchi et al., 2007; Yamauchi & Markman, 2000). In Experiment 2, animal cell labels might have influenced similarity judgments more than painter labels because animal cell labels are categorical labels that denote a group of cells (e.g., pig bone cell) whereas painter labels are indexical labels that indicate a single painter (e.g., Dalica). Experiment 3 tested this idea by comparing the labeling effect of categorical labels versus indexical labels.

Experiment 3 tested whether the labeling effect stems from these three types of background knowledge: the type of categories that labels represent (living things vs.

man-made objects), the amount of knowledge that labels convey, and the types of labels (categorical labels vs. indexical labels). To test these three factors exclusively through information carried by labels, the same set of animal tissue pictures were presented to participants with the same set of labels (e.g., Dalica, Goyama). The background knowledge associated with the labels was solely manipulated through instructions. Since all participants were presented with the same trials (Figure 7), the labeling effect in this experiment was unambiguously attributed to the effect of the background knowledge manipulated through instructions.

The instructions were manipulated in three ways. The instructions of the diseased cell condition and the painter condition were identical to the instructions used

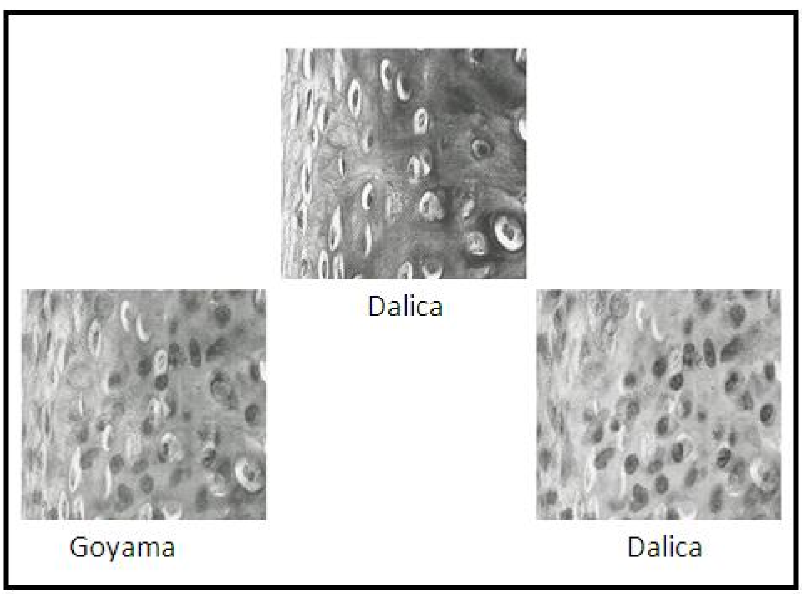


Figure 7. Sample trial shown in Experiment 3. Participants in all conditions received the same set of pictures and labels. The meanings of labels were manipulated mainly via instructions.

in Experiment 2. In the diseased cell condition, participants were told that the pictures were diseased cells and the pictures' labels were names of diseases. In the painter condition, participants were told that the pictures represented paintings and that labels referred to the painters' names. In a third condition (painting style condition), participants were also informed that pictures represented paintings but that the labels indicated the names of painting styles. In the painter condition, labels were indexical in the sense that the label indicated a painter who created the given pictures; however, in the painting style condition, labels were categorical because the label represented a painting style that could include other paintings that were drawn by many painters in that style. After carrying out the similarity judgment task, participants were asked to list as many exemplars as possible associated with the categories (diseases, painters, or painting styles). The amount of knowledge was measured by counting the number of exemplars that each participant generated.

Method

Participants. Three hundred and four undergraduates participated for course credit. They were randomly assigned to one of three conditions: diseased cell (n = 93), painting style (n = 106), and painter (n = 105) conditions.

Materials. Materials were the same as in Experiment 2.

Design. The experiment consisted of a 3 (Label Condition; diseased cell, painting style, and painter conditions; between subjects) × 3 (Physical Difference; low-difference, medium-difference, high-difference; within-subjects) factorial design.

Procedure. Participants in Experiment 3 were asked to carry out a similarity judgment task identical to that described in Experiment 1 and 2. However, after the similarity judgment task, participants in Experiment 3 were asked to list as many exemplars that they could think of in 2 minutes. Participants in the diseased cell condition were asked to list disease names, participants in the painting style condition were asked to list painting style names, and participants in the painter condition were asked to list painter names.

Results and Discussion

As in Experiment 2, when the labels indicated living things (diseased cell condition), participants used the labels significantly more often than when the labels represented man-made things (painter and painter style condition) (Figure 8). The main effect of label condition was significant, $F(2, 301) = 4.22$, $MSE = .08$, $p = .01$, $\eta^2 = .03$. The proportion of participants selecting the dissimilar base pictures (the tendency to use category labels pitted against physical similarity) was significantly higher in the diseased cell condition ($M = 0.26$) than in the painting style condition ($M = 0.21$) or the painter condition ($M = 0.19$): diseased cell vs. painting style, $t\ (197) = 2.02$, $SE = .03$, $p = .05$, $d = 0.29$, diseased cell vs. painter, $t\ (196) = 2.68$, $SE = .02$, $p < .01$, $d = 0.38$. The labeling effect between the painting style and painter conditions was statistically indistinguishable, $t\ (209) = .63$, $SE = .02$, $p = .53$, $d = 0.09$. There was no interaction effect of physical difference and label condition, $F(2,392) < 1.0$.

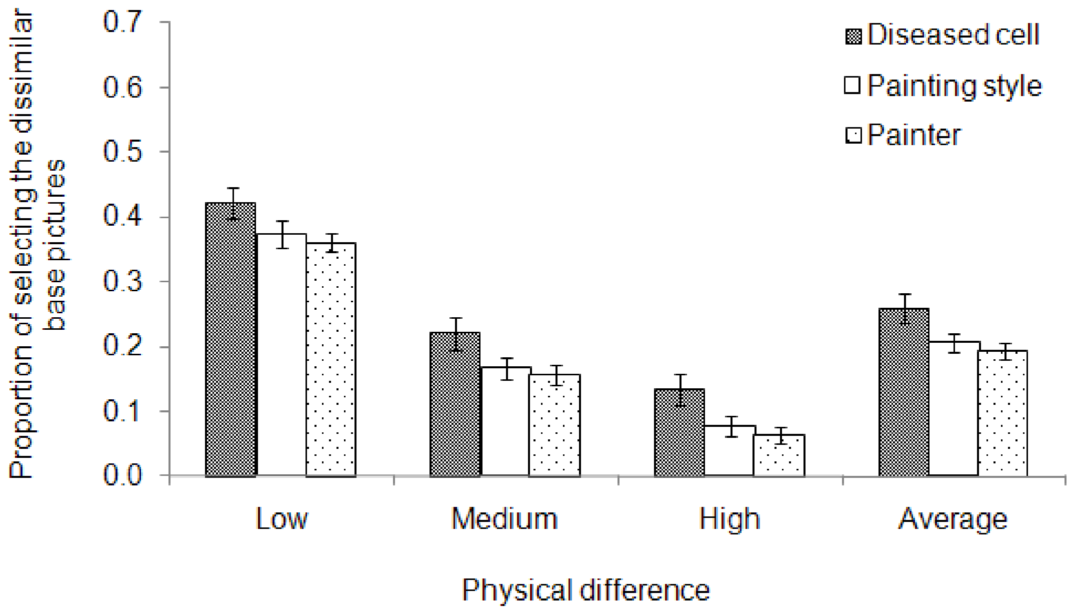


Figure 8. Mean proportions of participants selecting the dissimilar base picture over the other base picture in Experiment 3. The error bars represent two standard error units calculated from each condition.

Participants had more knowledge about diseases than about painters/painting styles. As Figure 9 shows, participants listed more disease names (M = 9.47, SD = 3.72) than painting style names (M = 4.64, SD = 2.36) and painter names (M = 3.22, SD = 2.35), The number of exemplars participants listed for disease names, painting style names, and painter names was significantly different from each other, $F(2, 301)$ = 129.42, MSE = 8.07, $p < .01$, $\eta^2 = .46$. More disease names were generated than names of painting styles or painter names, diseases vs. painting styles, t (197) = 11.09, SE = .44, p = .001, d = 1.58, diseases vs. painters, t (196) = 14.32, SE = .44, $p < .001$, d = 2.04.

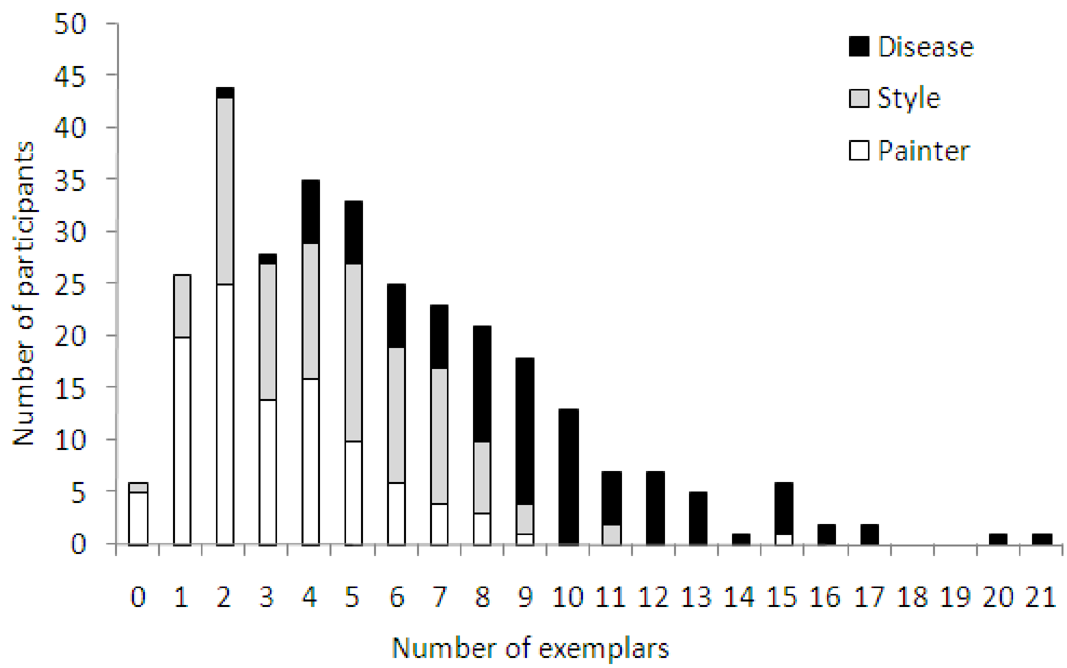


Figure 9. The number of exemplars participants listed for diseases, painting styles, and painters. X axis represents the number of exemplars; Y axis represents the number of participants.

Also, more painting styles were generated than painter names, t (209) = 4.39, SE = .33, p < .001, d = 0.60.

People might have used disease labels more than painter or painting style labels because they knew more diseases than painters and painting styles. To test this idea, I reanalyzed the data with ANCOVA with the number of exemplars as a covariate. This analysis showed that when the amount of background knowledge (i.e., the number of exemplars associated with labels) was controlled, the difference between the diseased cell, the painting style and painter conditions was reduced: $F(2, 300) = 2.84$, $MSE = .08$, $p = .06$, $\eta^2 = .02$. This analysis indicates that the amount of background knowledge influences the effect of types of categories in inductive generalizations.

To test the relationship between the types of categories and the amount of knowledge further, I analyzed the correlation between the number of exemplars and the average proportion of selecting dissimilar base pictures separately for living things (diseased cell condition) and for man-made objects (painting style condition and painter condition). As Figure 10 shows, there was no significant relationship between the number of exemplars generated and the proportion of selecting dissimilar base pictures for living things (diseased cell condition) (Figure 10a), $r(91) = .04$, $p = .72$. To analyze the relationship between the number of exemplars and the proportion of selecting dissimilar base picture for man-made objects, painter and painting style conditions were combined and analyzed. The correlation between the number of exemplars and the proportion of selecting dissimilar base pictures was not significant, $r(209) = -.07$, $p = .33$, though there was a negative (but very weak) relationship between the number of exemplars and the proportion of selecting dissimilar base pictures (Figure 10b).

Though participants in the diseased cell condition generated more exemplars of diseases than participants in the painter and painting style conditions, there was no correlation between the number of exemplars and the proportion of selecting dissimilar pictures (the tendency to use category labels pitted against physical similarity) in the diseased cell condition. Although there was an interesting negative (but very weak) relationship between the number of exemplars and the proportion of selecting dissimilar pictures, this correlation was not significant. This result suggests that for living things and man-made objects, participants may use labels relatively independently from the

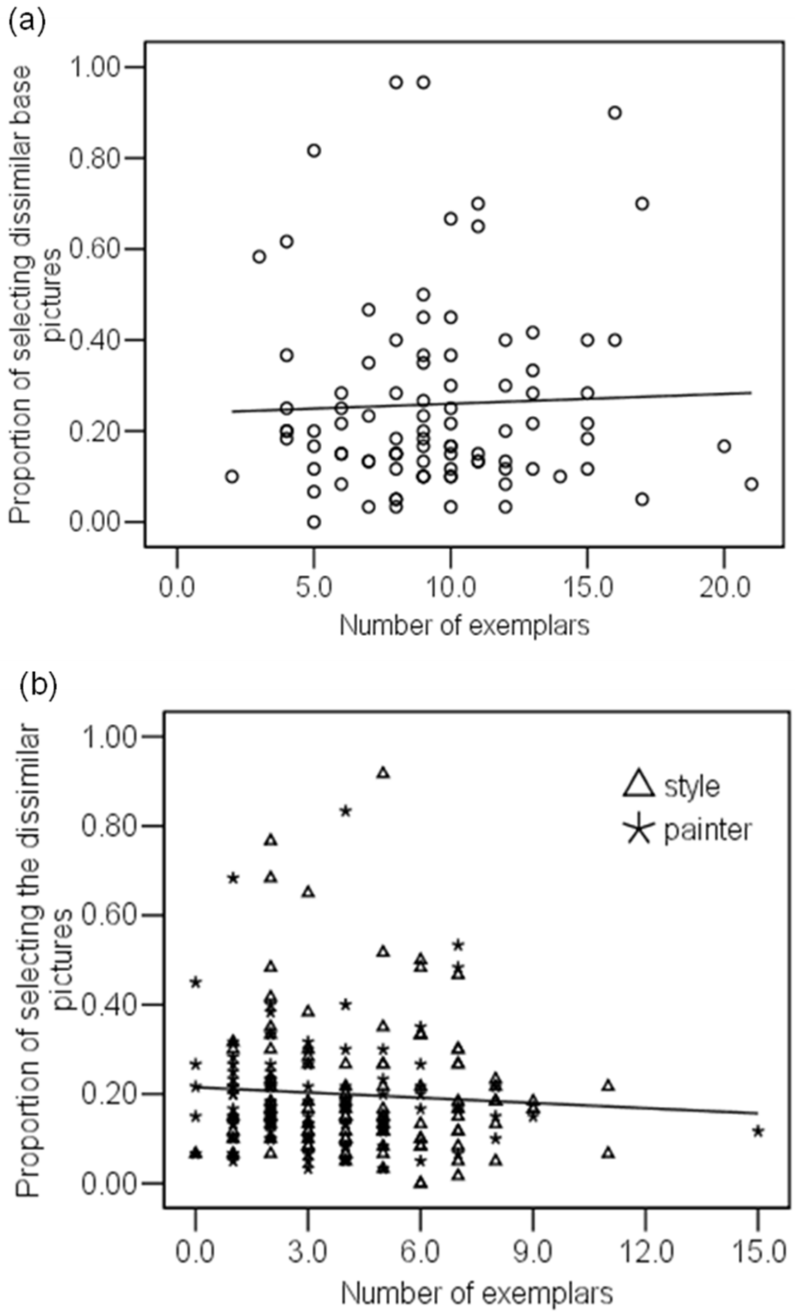


Figure 10. Relationship between the number of exemplars and the proportion of participants selecting the dissimilar pictures in the diseased cell condition (a), and the painter and painting style conditions (b).

amount of knowledge.

As in Experiment 2, the labeling effect was higher when the labels indicated living things (diseased cell condition) than when the labels represented man-made objects (painter and painting style conditions). Within man-made objects, the labeling effect did not differ between the painting style condition and the painter condition. The results suggest that, at least for the man-made objects, categorical labels (painting style condition) are not considered more important than indexical labels (painter condition). The results also suggest that the impact of the types of categories on inductive generalization does not necessarily depend on the amount of knowledge. Though participants listed more painting style names than painter names, the labeling effect did not differ between these two conditions.

In Experiment 4, I tested the influence of these three types of knowledge further using a different set of stimuli. In Experiment 3, the stimuli in the painting style and painter conditions were cell pictures and did not look like paintings; thus it is possible that participants did not use the painting labels. Therefore, Experiment 4 tested the effect of labels attached to real paintings.

EXPERIMENT 4: REAL PAINTINGS

The main goal of Experiment 4 is to replicate the results of Experiment 3 using a different set of pictures (Figure 11). Though participants in the painter and painting style conditions of Experiment 3 were instructed that the presented pictures were paintings, they might have thought that the pictures were in fact pictures of tissues. Thus, another possible explanation for this differential effect of labels is that participants might have thought diseased cell labels were more important than painter labels because the pictures looked more like cells than paintings. Experiment 4 tested this idea. The design, task, and materials in Experiment 4 were identical to Experiment 3, except that real abstract paintings were used instead of pictures of animal tissues.

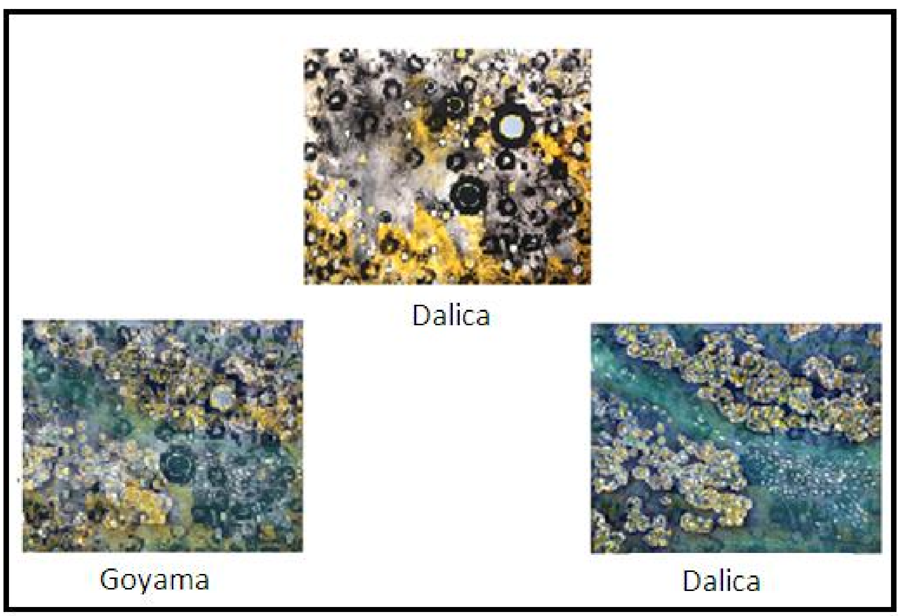


Figure 11. Sample trials shown in Experiment 4. Participants in all conditions received the same set of pictures and labels. The meanings of labels were manipulated via instructions.

Method

Participants. Three hundred eighty-seven undergraduates participated for course credit. They were randomly assigned to one of three conditions: diseased cell (n = 130), painting style (n = 121), and painter (n = 136) conditions.

Materials. Five pairs of abstract paintings were selected from various websites. The procedure of creating 60 triads of original and morphed pictures was identical to the procedure described in Experiment 1 (see Appendix A for samples of morphed stimuli). As in previous experiments, the triads of pictures had three levels of physical difference (low-, medium-, and high-difference).

Design and procedure. The design and procedure in Experiment 4 were identical to those used in Experiment 3. Experiment 4 used a 3 (Label Condition; diseased cell, painting style, and painter conditions; between subjects) × 3 (Physical Difference; low-difference, medium-difference, high-difference; within-subjects) factorial design. Participants were presented with the same 60 triads of paintings with labels and were asked to judge which base picture was more similar to the target. However, participants in the diseased cell condition were instructed that the stimuli were cell pictures infected by diseases and that the labels specified the names of the diseases that infected the cells. Participants in the painting style condition were informed that the stimuli were paintings and that the labels indicated the names of painting styles. Participants in the painter condition were also told that the stimuli were paintings but that the labels indicated the painters' names. As in Experiment 3, participants were asked to judge similarity of pictures and then to list as many exemplars of the category as possible.

Results and Discussion

Results from Experiment 4 replicated the results from Experiment 3. As in Experiment 2 and Experiment 3, when labels represented living things (diseased cell condition), participants used the labels considerably more often than when the labels indicated man-made objects (painter and painting style conditions) (Figure 12). The main effect of label condition was significant, F (2, 384) = 5.24, MSE = 0.05, p = .006, η^2 = .03. The proportion of participants selecting the dissimilar base pictures (the tendency to use category labels) was significantly higher in the diseased cell condition (M = 0.21) than in the painting style condition (M = 0.16) or in the painter

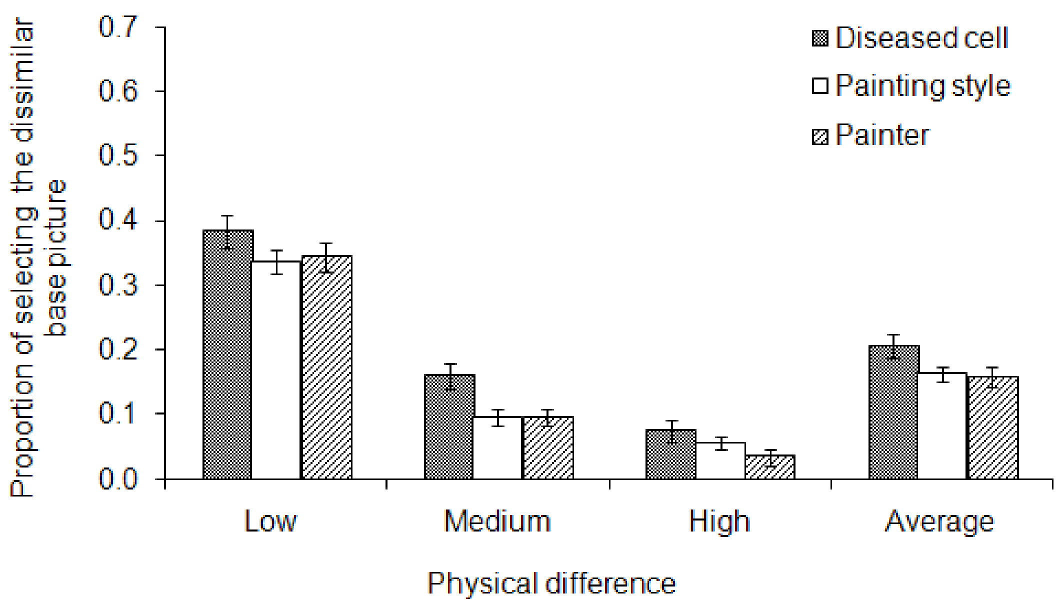


Figure 12. Mean proportions of participants selecting the dissimilar base picture over the other base picture in Experiment 4. The error bars represent two standard error units calculated from each condition.

condition (*M* = 0.16): diseased cell vs. painting style, *t* (249) = 2.30, *SE* = .02, *p* = .02, *d* = .29; diseased cell vs. painter condition, *t* (264) = 2.80, *SE* = .02, *p* = .006, *d* = .34. However, the proportion of selecting dissimilar base pictures was not different in the painting style condition compared to the painter condition, *t* (255) = 0.38, *SE* = .01, *p* = .71, *d* = .05.

As in Experiment 3, people had more knowledge of diseases than paintings (Figure 13). Participants listed more disease names (*M* = 9.69, *SD* = 3.65) than style names (*M* = 4.74, *SD* = 2.41) and painter names (*M* = 3.54, *SD* = 2.67). The number of exemplars participants listed for the three conditions were significantly

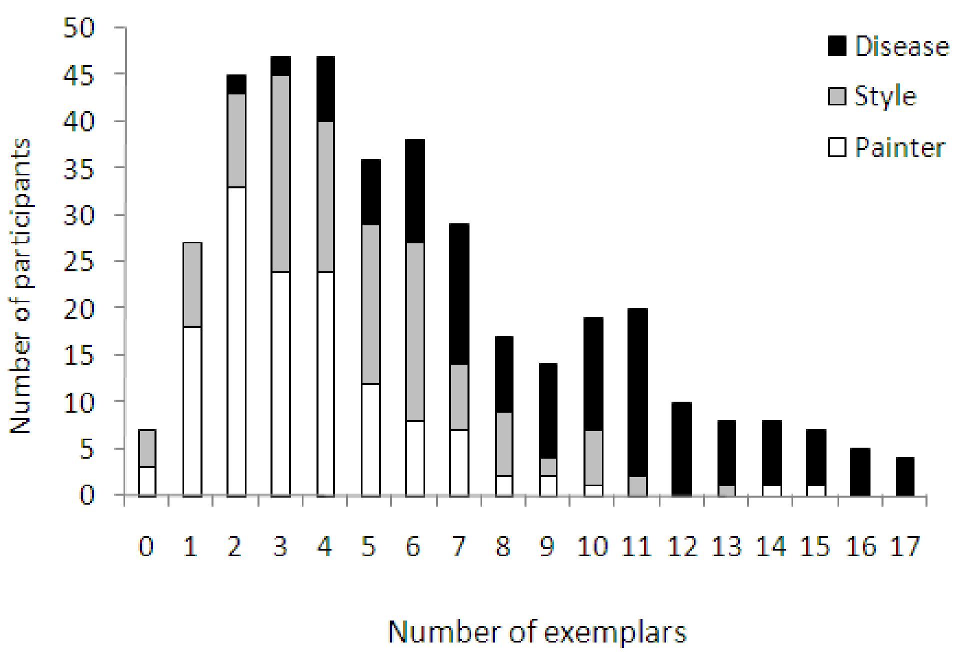
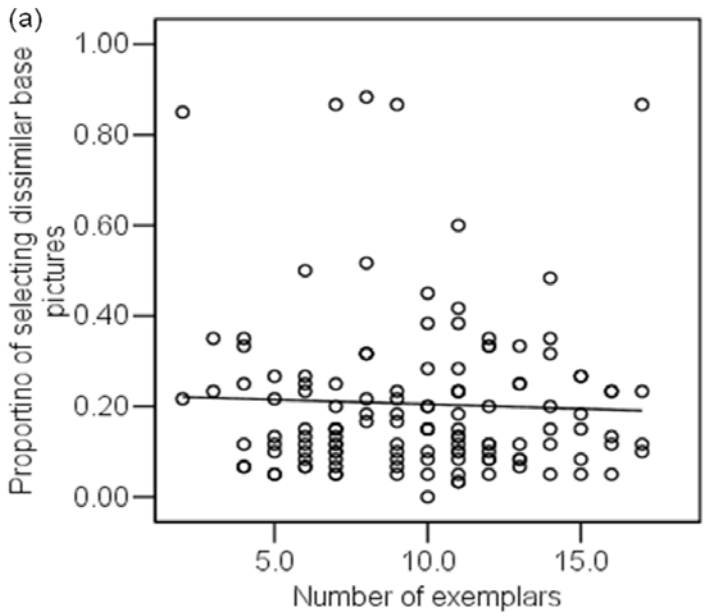


Figure 13. The number of exemplars participants listed for diseases, painting styles, and painters in Experiment 4. X axis represents the number of exemplars; Y axis represents the number of participants.

different from each other, $F(2, 384) = 159.09$, $MSE = 8.73$, $p < .001$, $\eta^2 = .45$. Participants listed significantly more disease names than painting style names, t (249) = 12.20, $SE = .41$, $p < .001$, $d = 1.54$; they generated significantly more disease names than painter names, t (264) = 16.27, $SE = .38$, $p < .001$, $d = 2.00$. They also listed significantly more painting style names than painter names, t (255) = 3.76, $SE = .32$, $p < .001$, $d = .47$. [1]

These results suggest that it is possible that people used disease labels more than painter and painting style labels in similarity judgments because they knew more about diseases than painters and painting styles. To test this possibility, I reanalyzed the data with ANCOVA by controlling the effect of the number of exemplars as a covariate. The analysis showed a different result from Experiment 2: when the number of exemplars

[1] After the experiment, participants in the diseased cell and painter conditions viewed 10 original painting stimuli one by one and were asked to estimate the extent to which the pictures look like diseased cells or paintings. Participants who received the diseased cell condition were asked to estimate the likelihood of each original painting stimulus to be a diseased cell. Participants who received the painter condition were asked to estimate the likelihood of each original painting stimulus to be a painting. Participants estimated that the painting stimuli are likely to be paintings ($M = 62.37$) more than diseased cells ($M = 48.82$), $t(221) = 7.84$, $p < .001$.



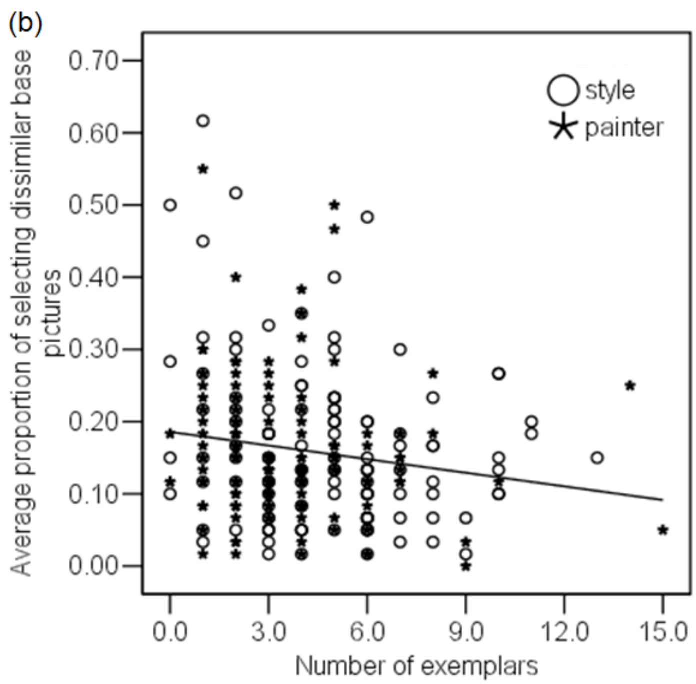


Figure 14. Relationship between the number of exemplars and the proportion of participants selecting the dissimilar pictures in diseased cell condition (a), and the painter and painting style conditions (b).

associated with labels was controlled, the difference between the diseased cell and the painter and painting style conditions increased, $F(2, 383) = 6.88$, $MSE = 0.05$, $p = .001$, $\eta^2 = .04$. This analysis indicates that the effect of types of category in inductive generalizations is independent from the amount of background knowledge.

It is possible that the number of exemplars has a different influence on the labeling effect with man-made objects (the painter/painting style conditions) compared to living things (diseased cell condition) (Figure 14). To test this possibility, I analyzed the correlation between the number of exemplars and the average proportion of selecting dissimilar base pictures. The correlation was not significant for living things (Figure 14a), $r(128) = -.04$, $p = .64$. However, for man-made objects (painters, painting styles), there was a significant, negative relationship between the number of exemplars and the proportion of selecting the dissimilar pictures (Figure 14b), $r(255) = -.16$, $p = .01$. For man-made objects, participants who could list many exemplars used labels in their similarity judgment less often; participants who could list fewer exemplars used labels more often.

The results suggest that the impact of types of categories on the labeling effect does not depend on the amount of knowledge. As in Experiment 3, though participants in the diseased cell condition generated more exemplars of diseases than participants in the painter and painting style conditions, there was no correlation between the number of exemplars and the proportion of selecting dissimilar pictures for living things. For man-made objects, there was a negative correlation between the number of exemplars and the proportion of selecting dissimilar pictures. Also within man-made objects, participants generated more painting style names than painter names but the labeling effect between these two conditions did not differ.

The influence of labels on similarity judgments was determined by the interaction between types of categories and the amount of knowledge. Participants used labels more for living things than man-made objects; however, within the living things, the relationship between the amount of knowledge (the number of exemplars) and the proportion of selecting dissimilar base pictures was not significant. This result suggests that the tendency to use category labels may be relatively independent from the amount of knowledge. For man-made objects, however, people who had less knowledge (listed fewer exemplars) used the labels more than people who had more knowledge (listed many exemplars). This result suggests that for man-made things (paintings or the stimuli that were introduced as paintings), the labeling effect depends on the amount of knowledge. The types of categories (living things vs. man-made things) and the amount of knowledge (the number of exemplars) jointly interact with the tendency to use category labels in inductive generalization.

GENERAL DISCUSSION

How are labels used in inductive generalization? In this paper, I suggest that people do not always form naïve assumptions or evaluate overall similarity when using labels. I propose that the labeling effect is more complicated and is based on the interaction between the types of categories and the amount of knowledge.

Summary of Results

Altogether, the results suggest that the labeling effect is stronger for living things than for man-made objects and depends on the amount of knowledge. In Experiment 1, when dissimilar looking animal faces were presented with the same category labels (animal names) these pictures were judged to be more similar than when dissimilar looking animal faces were presented without labels. Given that the labeling effect was replicated in a similarity judgment task in Experiment 1, Experiment 2 investigated the influence of the types of categories in the similarity judgment task. Animal tissue pictures were presented with category labels and these labels specified that the pictures were either living things (animal cells, diseased cells) or man-made objects (abstract paintings, pictures). The results demonstrated that labels significantly affected similarity judgment more for living things than for man-made objects. Experiment 3 demonstrated that, in addition to the types of categories, the amount of knowledge (listed exemplars) interacted with the labeling effect. The labeling effect was stronger for living things (e.g., diseased cells) than for man-made objects (e.g., paintings); and people generated more exemplars for living things than for man-made objects. Using real paintings, Experiment 4 demonstrated that types of knowledge and the amount of knowledge jointly interacted with the labeling effect. For man-made objects, the amount of knowledge negatively correlated with the tendency to use category labels in similarity judgments; while for living things, there was no correlation between the amount of knowledge and the tendency to use category labels.

The results demonstrate that people rely on labels in judgments of living things more than in judgments of man-made objects and this tendency is not entirely due to the amount of knowledge. For man-made objects, the tendency to use category labels is related to the amount of knowledge, while for living things, the tendency to use category labels is relatively independent from the amount of knowledge. The results suggest that the strong tendency to rely on category labels in inductive generalization is a result of the interaction between types of categories and the amount of knowledge.

Limitations and Future Studies

The interaction between the amount of knowledge and the types of categories is rather limited in two ways. In this study, only diseased cells and animal cells were examined as examples of living things, and only paintings and pictures were used as examples of man-made objects. A variety of stimuli should be used in future studies that test the interaction between the amount of knowledge and types of categories. While it is necessary to examine the labeling effect in living things versus man-made objects using a variety of stimuli, this will not entirely address the criticism that the type of category is not a valid factor in cognitive judgments. This criticism stems from the fact that many researchers are skeptical about grouping categories in terms of their contents (Ahn, 1998; Medin, Lynch, & Solomon, 2000; Rehder, 2003; Rehder, Hastie, 2004; Roberts, 2007; Sloman & Malt, 2003). Even though the distinction between living things and man-made objects was found and replicated in different aspects of cognitive psychology (Barton & Komatsu, 1989; Carey, 1985; Hillis & Caramazza, 1991; Hong, Chao, & No, 2009; Keil, 1989; Medin & Ortony, 1989; Warrington & Shallic, 1984), it is difficult to have a precise definition of ontological categories, such as living things and man-made objects (Hirschfeld & Gelman, 1994). Madin, Lynch and Solomon (2000) correctly pointed out that “a fear is that domain-specificity theorists simply define kinds into existence by stating a priori that certain kinds of content are important” (p 137). As a result, recent research has focused on finding a universal principle that distinguishes living things and man-made objects (Ahn, 1998, 2000; Rehder, 2003; Rehder, Hastie, 2004). However, finding a universal principle is not the main purpose of this current study and future studies are needed to address this question further in terms of the labeling effect.

The interaction between the amount of knowledge and types of categories is also limited because the amount of knowledge was measured only for the labels. Given that the task used in this study involves pictures as well as labels, the amount of knowledge of pictures might have contributed to the tendency to use labels in inductive generalization. Experiment 4 indirectly addressed this idea by replicating the result of Experiment 3 with different stimuli. Experiment 3 showed that when stimuli were cell pictures, people used disease labels attached to pictures of cells more than painter and painting style labels. In Experiment 4, though the stimuli were real paintings, participants still tended to use disease labels more than painter and painting style labels.

However, it is still possible that, in Experiment 3, the people who were familiar with pictures of cells might have used the cell pictures more than the disease labels to make similarity judgments. Future studies should investigate the labeling effect using two factors: the amount of knowledge for labels as well as the amount of knowledge for pictures.

Previous studies on the labeling effect demonstrated that when labels indicate category membership (labels that represent a group of items) the categorical labels have a significant impact on inductive generalization (Gelman & Heyman, 1999; Rhemtulla & Hall, 2008; Waxman & Booth, 2001; Waxman & Markow, 1995; Yamauchi et al., 2007; Yamauchi & Markman, 2000; Yamauchi & Yu, 2008). However, in this research, there was no significant difference between the effect of categorical labels and the effect of indexical labels (labels that denote a single exemplar). Given that the effect of categorical labels and indexical labels were not compared for living things, this result should only be generalized for man-made objects. The results of this study suggest that types of labels appeared to have little impact on the labeling effect for man-made things, but the role of types of labels for living things needs to be investigated further.

By definition, a common noun refers to a group of objects, while a proper noun designates a particular individual (Kripke, 1980). In this sense, the distinction between categorical labels and indexical labels can be viewed such that common nouns and proper nouns will generate different labeling effects. The results of this study showed, however, that common nouns did not generate a greater labeling effect compared to proper nouns. This could be because proper nouns also denote some categorical information. Proper nouns typically originated as expressions that describe individuals (Hall, Waxman, Brédart, & Nicolay, 2003) such as one's job (e.g., Carpenter, Baker), places (e.g., Hills, Fields) and desirable qualities (e.g., Faith, Hope). Most proper names have lost their major descriptive roles today (e.g., Mr. Carpenter is most certainly not a carpenter); however, many proper names serve some descriptive roles such as gender (e.g., Diana would not likely be a boy) and family relations (e.g., Edward Jr. would be a son of Edward). Additionally, adults are likely to know that a proper noun can be a homophone of an adjective (e.g., Lefty as a nickname for left-handed people). Even infants as young as 12 months use both proper nouns (e.g., Sib) as well as common nouns (e.g., horse) to designate multiple things (Graham, Kilbreath, & Welder, 2004; Hall et al., 2003; Jaswal, 2007; Jaswal & Markman, 2007; Waxman & Booth, 2001; Waxman & Markow, 1995). These studies suggest that, though the distinction between common nouns and proper nouns (i.e., categorical labels and indexical labels) may exist, people are still able to use both common and proper nouns to indicate groups of things and use them equally in inductive generalization.

Theoretical Implications

The goals of this research are to explain the labeling effect by looking at the role of background knowledge and to complement the naïve theory and the similarity-based approach. Both views explain the labeling effect: people are willing to generalize a characteristic of one item to the other item, even when the items look dissimilar but have the same category label. However, these two views differ on the role of labels. The naïve theory approach suggests that matching labels act as an assumption during the course of induction, whereas the similarity-based approach suggests that matching labels serve as a type of shared feature and contribute to overall similarity.

The naïve theory approach suggests that people treat the shared labels as a type of rule. Developmental studies show that young children and infants who have less knowledge about the items treat the shared labels as more important than the shared similarity (Brown, 1957; Gelman & Coley, 1990; Hall et al., 2003; Jaswal & Markman, 2007; Rhemtulla & Hall, 2008; Waxman & Booth, 2001; Waxman & Markow, 1995), suggesting that young children make the assumption that shared labels indicate shared category membership. Later adult studies support this claim by demonstrating that those shared labels that indicate the shared category membership are qualitatively different from physical features (Yamauchi & Markman, 2000) and are considered in earlier stages of inductive generalization (Yamauchi, Kohn & Yu, 2009).

The similarity-based approach argues that shared labels are just a type of feature. In the computational model by Sloutsky and Fisher (2004a), the shared labels and physical features were both treated as types of features. By assigning larger attention weights for labels than for physical features, the computational model could explain children's tendency to use category labels when they are pitted against physical similarity. Recent evidence showed that, for newly learned categories, shared category labels did not influence inductive generalization. Children learned to assign category labels to drawings of bugs based on one rule (whether the number of legs is greater than the number of buttons in each bug). Later, children were asked to perform an inductive generalization task (e.g., whether one characteristic of a cartoon bug generalizes to the other cartoon bug). They found that the newly learned category labels were not used more than physical features (Sloutsky & Fisher, 2004a, 2004b; Robinson & Sloutsky, 2004; Fisher & Sloutsky, 2005; Sloutsky, Kloos & Fisher, 2007).

The results of this study suggest that the labeling effect is complicated and affected by the interaction between types of categories and amount of knowledge. Perhaps matching category labels are sometimes used as a type of rule, depending on the type of category. Since labels for living things were consistently used more than labels for man-made objects, participants might have assigned larger attention weights for labels for living things than for man-made objects. Additionally, since the labeling effect for living things was relatively independent from the amount of knowledge, participants might have been reluctant to change these weights based on the amount of knowledge. This suggests that for living things people may instill rules for category labels independently from the amount of knowledge.

For man-made objects, results suggest that people may use category labels as a type of feature, for which they can assign different attention weights. When participants had more knowledge about paintings, they weighted labels less than physical features in their similarity judgments; when they had less knowledge about paintings, they weighted labels more than physical features in their judgments. This result suggests that for man-made objects, people are willing to evaluate the weight of labels and physical features based on the amount of knowledge they have.

Additional Considerations

Why do types of categories and the amount of knowledge jointly interact with the labeling effect? One possible reason is that the way we experience living things and man-made objects may differ. Johnson and Mervis (1997) found that, when compared with novices, experts learn to detect more subtle physical features of the actual items for living things (e.g., birds). Novices may encounter living things and man-made objects in different ways. For example, novices may understand types of diseases and cells through naming and describing their characteristics and symptoms, rather than looking at actual pictures of cells. On the other hand, novices may understand types of paintings by looking at actual paintings rather than by naming and describing characteristics of the paintings. The way people experience items in real situations may determine the salience of labels in inductive generalization.

The interaction between the amount of knowledge and types of categories can also be due to limited cognitive resources. The category structure of living things is more fuzzy and complicated than man-made objects, so that people seek readily available information—the shared category labels. Categories of living things have family resemblance structure, while categories of man-made objects are often defined by a single feature (e.g., function; Coley et al., 2004; Malt, 1990, 1993; Rosch & Mervis, 1975). Since the category structure of man-made objects is relatively simple and clear, participants may be more willing to rely on the amount of knowledge and physical features instead of relying on easily accessible information such as the shared category label. However, for living things, the category structure is fuzzy so participants may choose to use the shared category label as a default heuristic in an attempt to reduce the cognitive load to handle inductive generalization (Lagnado & Shanks, 2003; Ross & Murphy, 1996; Yamauchi, 2005).

CONCLUSION

Previous research on the labeling effect suggests that people use shared category labels as a type of rule or as a type of feature. The present study indicates that background knowledge has a critical role in the labeling effect: types of categories and the amount of knowledge jointly interact with the labeling effect. Four experiments demonstrated that for living things, the labeling effect was strong and relatively independent from the amount of knowledge; for man-made objects however, the labeling effect was weak and depended on the amount of knowledge. The results suggest that people are willing to evaluate shared category labels as a type of feature for man-made objects more than for living things.

REFERENCES

Ahn, W. (1998). Why are different features central for natural kinds and artifacts?: The role of causal status in determining feature centrality. *Cognition, 69*, 135-178.

Ahn, W. (2000). Causal status as a determinant of feature centrality. *Cognitive Psychology, 41*, 361-416.

Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Erlbaum.

Bacha, W. J., Jr., & Bacha, L. M. (2000). *Color atlas of veterinary histology* (4th ed.). Philadelhia, PA: Lippincott Williams & Wilkins.

Barsalou, L. W. (1985). Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*, 629-654.

Barton, M. E., & Komatsu, L. K. (1989). Defining features of natural kinds and artifacts *Journal of Psycholinguistic Research, 18*, 433-447.

Brown, R. W. (1957). Linguistic determinism and the part of speech. *Journal of Abnormal and Social Psychology, 55*, 1-5.

Caramazza, A., & Shelton, J. R. (1998). Domain-specific knowledge systems in the brain: The animate–inanimate distinction. *Journal of Cognitive Neuroscience, 10*, 1-34.

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

Chao, M. M., Chen, J., Roisman, G., & Hong, Y.-y. (2007). Essentializing race: Implications for bicultural individuals' cognition and physiological reactivity. *Psychological Science, 18*, 341-348.

Coley, J. D., Hayes, B. K., Lawson, C., & Moloney, M. (2004). Knowledge, expectations, and inductive reasoning within conceptual hierarchies. *Cognition, 90*, 217-253.

Damasio, H., Grabowki, T. j., Tranel, D., Hichwa, R. D., & Damasio, A. R. (1996). A neural basis for lexical retrieval. *Nature, 27*, 499-505.

DeLoache, J. S. (2004). Becoming symbol-minded. *TRENDS in Cognitive Sciences, 8*(2), 66-70.

Fisher, A. V., & Sloutsky, V. M. (2005). When induction meets memory: Evidence for gradual transition from similarity-based to category-based induction. *Child Development, 76*, 583-597.

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

Gelman, S. A., & Coley, J. D. (1990). The importance of knowing a dodo is a bird: Categories and inferences in 2-year-old children. *Developmental Psychology, 26*, 796-804.

Gelman, S. A., & Heyman, G. D. (1999). Carrot-eaters and creature-believers: The effects of Lexicalization on children's inferences about social categories. *Psychological Science, 10*(6), 489-493.

Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition, 23*, 183-209.

Gelman, S. A., & Waxman, S. R. (2007). Looking beyond looks: Comments on Sloutsky, Kloos, and Fisher (2007). *Psychological Science, 18*(6), 554-555.

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

Hall, G. D., Waxman, S. R., Brédart, S., & Nicolay, A.-C. (2003). Preschoolers' use of form class cues to learn descriptive proper names. *Child Development, 74*, 1547-1560.

Heit, E. (1998). A Bayesian analysis of some forms of inductive reasoning. In M. Oaksford & N. Chater (Eds.), *Rational models of cognition* (pp. 212-217). Oxford: Oxford University Press.

Heit, E. (2000). Properties of inductive reasoning. *Psychological Bulletin & Review, 7*(4), 569-592.

Heit, E., & Hayes, B. K. (2005). Relations among categorization, induction, recognition, and similarity: Comment on Sloutsky and Fisher (2004). *Journal of Experimental Psychology: General, 134*(4), 606-611.

Heit, E., & Rubinstein, J. (1994). Similarity and property effects in inductive reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 411-422.

Heyes, B. K., & Thompson, S. P. (2007). Causal relations and feature similarity in children's inductive reasoning. *Journal of Experimental Psychology: General, 136*(3), 470-484.

Hillis, A. E., & Caramazza, A. (1991). Category-specific naming and comprehension impairment: A double dissociation. *Brain, 114*, 2081-2094.

Hirschfeld, L. A., & Gelman, S. A. (1994). Toward a topography of mind: An introduction to domain specificity. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the Mind: Domain Speciticity in Cognition and Culture*. Cambridge, UK: Cambridge University Press.

Hong, Y.-y., Chao, M. M., & No, S. (2009). Dynamic interracial/intercultural processes: the role of lay theories of race. *Journal of Personality, 77*, 1283-1309.

Jaswal, V. K. (2007). The effect of vocabulary size on toddlers' receptiveness to unexpected testimony about category membership. *Infancy, 12*, 169-187.

Jaswal, V. K., & Markman, E. M. (2007). Looks aren't everything: 24-month-olds' willingness to accept unexpected labels. *Journal of Cognition and Development, 8*, 93-111.

Johnson, K. E., & Mervis, C. B. (1997). Effects of varying levels of expertise on the basic level of categorization. *Journal of Experimental Psychology: General, 126*, 248-277.

Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (2008). The advantage of abstract examples in learning math. *Science, 320*, 454-455.

Keates, J., & Graham, S. A. (2009). Category markers or attributes: Why do labels guide infants' inductive inferences? *Psychological Science, 19*, 1287-1293.

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

Kripke, S. A. (1980). *Naming and necessity*. Oxford, England: Willy-Blackwell.

Lagnado, D. A., & Shanks, D. R. (2003). The influence of hierarchy on probability judgment. *Cognition, 89*, 157-178.

Lupyan, G. (2008). From chair to "chair": A representational shift account of object labeling effects on memory. *Journal of Experimental Psychology: General, 137*, 348-369.

Lupyan, G., Rakison, D. H., & McClelland, J. L. (2007). Language is not just for talking: Redundant labels facilitate learning of novel categories. *Psychological Science, 18*, 1077-1083.

López, A., Atran, S., Coley, J. D., Medin, D. L., & Smith, E. E. (1997). The tree of life: Universal and cultural features of folkbiological taxonomies and inductions. *Cognitive Psychology, 32*, 251-295.

Malt, B. C. (1990). Features and beliefs in the mental representation of categories. *Journal of Memory and Language, 29*, 289-315.

Malt, B. C. (1993). Concept structure and category boundaries. In G. V. Nakamura, D. L. Medin & R. Taraban (Eds.), *Categorization by humans and machines. The psychology of learning and motivation: Advances in research and theory* (Vol. 29, pp. 363-390). San Diego: Academic Press.

Malt, B. C., Ross, B. H., & Murphy, G. L. (1995). Predicting features for members of natural categories when categorization is uncertain. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 646-661.

Markman, A. B., & Ross, B. H. (2003). Category use and category learning. *Psychological Bulletin, 129*, 592-613.

McDonald, J., Samuels, M., & Rispoli, J. (1996). A hypothesis-assessment model of categorical argument strength. *Cognition, 59*, 199-217.

Medin, D., & Ortony, A. (1989). Psychological essentialism. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 179 - 195). New York: Cambridge University Press.

Medin, D. L., & Atran, S. (2004). The native mind: Biological categorization and reasoning in development and across cultures. *Psychological Review, 111*(4), 960-983.

Medin, D. L., Coley, J. D., Storms, G., & Hayes, B. K. (2003). A relevance theory of induction. *Psychonomic Bulletin & Review, 10*(3), 517-532.

Medin, D. L., Lynch, E. B., Coley, J. D., & Atran, S. (1997). Categorization and reasoning among tree experts: Do all roads lead to Rome? *Cognitive Psychology, 32*, 49-96.

Medin, D. L., Lynch, E. B., & Solomon, K. O. (2000). Are there kinds of concepts? *Annual Review of Psychology, 51*, 121-147.

Moss, H. E., Tyler, L. K., Durrant-Peatfield, M., & Bunn, E. M. (1998). 'Two eyes of a see-through': Impaired and intact semantic knowledge in a case of selective deficit for living things. *Neurocase, 4*, 291-310.

Murphy, G. L. (2003). The downside of categories. *TRENDS in Cognitive Sciences, 7*(12), 513-514.

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

Murphy, G. L., & Ross, B. H. (1994). Predictions from uncertain categorizations. *Cognitive Psychology, 27*, 128-193.

Napolitano, A. C., & Sloutsky, V. M. (2004). Is a picture worth a thousand words? The flexible nature of modality dominance in young children. *Child Development, 75*, 1850-1870.

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

Özçalişkan, Ş., Goldin-Meadow, S., Gentner, D., & Mylander, C. (2009). Does language about similarity play a role in fostering similarity comparison in children? *Cognition, 112*, 217-228.

Rehder, B. (2003). A causal-model theory of conceptual representation and categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(6), 1141-1159.

Rehder, B. (2006). When similarity and causality compete in category-based property generalization. *Memory & Cognition, 34*(1), 3-16.

Rehder, B., & Hastie, R. (2004). Category coherence and category-based property Induction. *Cognition, 91*, 113-153.

Rhemtulla, M., & Hall, G. D. (2008). Monkey business: Children's use of character identity to infer shared properties? *Cognition, 113*, 127-176.

Rips, L. J. (1975). Inductive judgments about natural categories. *Journal of Verbal Learning and Verbal Behavior, 14*, 665-685.

Rips, L. J. (1989). Similarity, typicality, and categorization. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 21-59). Cambridge, England: Cambridge University Press.

Rips, L. J. (2001). Necessity and natural categories. *Psychological Bulletin, 127*, 827-852.

Roberson, D., Davies, I., & Davidoff, J. (2000). Color categories are not universal: Replications and new evidence from a stone-age culture. *Journal of Experimental Psychology: General, 129*(3), 369-398.

Roberts, M. J. (2007). *Integrating the mind: Domain general vs. domain specific processes in higher cognition*. New York: Psychology Press.

Robinson, C. W., & Sloutsky, V. M. (2004). Auditory dominance and its change in the course of development. *Child Development, 75*, 1387-1401.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7*, 573-605.

Ross, B. H., & Murphy, G. L. (1996). Category-based predictions: Influence of uncertainty and feature associations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 736-753.

Saalbach, H., & Imai, M. (2007). Scope of linguistic influence: Does a classifier system alter object concepts? *Journal of Experimental Psychology: General, 136*(3), 485-501.

Sacchett, C., & Humphreys, G. W. (1992). Calling a squirrel a squirrel but a canoe a wigwam: A category-specific deficit for artifactual objects and body parts. *Cognitive Neuropsychology, 9*, 73-86.

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

Sloman, S. A. (1994). When explanations compete: the role of explanatory coherence on judgments of likelihood. *Cognition, 52*, 1-21.

Sloman, S. A., & Malt, B. C. (2003). Artifacts are not ascribed essences, nor are they treated as belonging to kinds. *Language and Cognitive Processes, 18*, 563-582.

Sloutsky, V. M., & Fisher, A. V. (2004a). Induction and categorization in young children: A similarity-based model. *Journal of Experimental Psychology: General, 133*(2), 166-188.

Sloutsky, V. M., & Fisher, A. V. (2004b). When development and learning decrease memory: Evidence against category-based induction in children. *Psychological Science, 15*(8), 553-558.

Sloutsky, V. M., Kloos, H., & Fisher, A. V. (2007). When looks are everything: Appearance similarity versus kind information in early induction. *Psychological Science, 18*, 179-185.

Sloutsky, V. M., & Spino, M. A. (2004). Naive theory and transfer of learning: When less is more and more is less. *Psychonomic Bulletin & Review, 11*, 528-535.

Smith, E. E., Shafir, E., & Osherson, D. (1993). Similarity, plausibility, and judgments of probability. *Cognition, 49*, 67-96.

Smith, S. M., Ward, T. B., Tindell, D. R., Sifonis, C. M., & Wilkenfeld, M. J. (2000). Category structure and created memories. *Memory & Cognition, 28*, 386-395.

Tenenbaum, J. B., & Griffiths, T. L. (2001). Generalization, similarity, and Bayesian inference. *Behavioral and Brain Sciences, 24*, 629-640.

Tversky, A. (1977). Features of similarity. *Psychological Review, 84*(4), 327-352.

Ward, T. B., Patterson, M. J., Sifonis, C. M., Dodds, R. A., & Saunders, K. N. (2002). The role of graded category structure in imaginative thought. *Memory & Cognition, 30*, 199-216.

Warrington, E. K., & McCarthy, R. (1983). Category specific access dysphasia. *Brain, 106*, 859-878.

Warrington, E. K., & Shallice, T. (1984). Category specific semantic impairments. *Brain, 107*, 829-253.

Wattenmaker, W. D. (1995). Knowledge structures and linear seperability: Integrating information in object and social categorizatoin. *Cognitive Psychology, 28*, 274-328.

Wattenmaker, W. D., Dewey, G. I., Murphy, T. D., & Medin, D. L. (1986). Linear separability and concept learning: Context, relational properties, and concept naturalness. *18*, 158-194.

Waxman, S. R., & Booth, A. E. (2001). Seeing pink elephants: Fourteen-month-olds' interpretations of novel nouns and adjectives. *Cognitive Psychology, 43*(217-242).

Waxman, S. R., & Braun, I. (2005). Consistent (but not variable) names as invitations to form object categories: new evidence from 12-month-old infants. *Cognition, 95*, B59-B68.

Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12 to 13-month-old infants. *Cognitive Psychology, 29*.

Waxman, S. R., & Markow, D. B. (1998). Object properties and object kind: Twenty-one-month-old infants' extension of novel adjectives. *Child Development, 69*(5), 1313-1329.

Wisniewski, E. J., & Medin, D. L. (1994). On the interaction of theory and data in concept learning. *Cognitive Psychology, 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. (2009). Finding abstract commonalities of category members. *Journal of Experimental & Theoretical Artificial Intelligence, 21*, 155-180.

Yamauchi, T., Kohn, N., & Yu, N. Y. (2007). Tracking mouse movement in feature inference: Category labels are different from feature labels. *Memory & Cognition, 5*(852-863).

Yamauchi, T., & Markman, A. B. (2000a). Inference using categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(3), 776-795.

Yamauchi, T., & Markman, A. B. (2000b). Learning categories composed of varying instances: The effect of classification, inference, and structural alignment. *Memory & Cognition, 28*, 64-78.

Yamauchi, T., & Yu, N. Y. (2008). Category labels versus feature labels: Category labels polarize inferential predictions. *Memory & Cognition, 36*(3), 544-553.

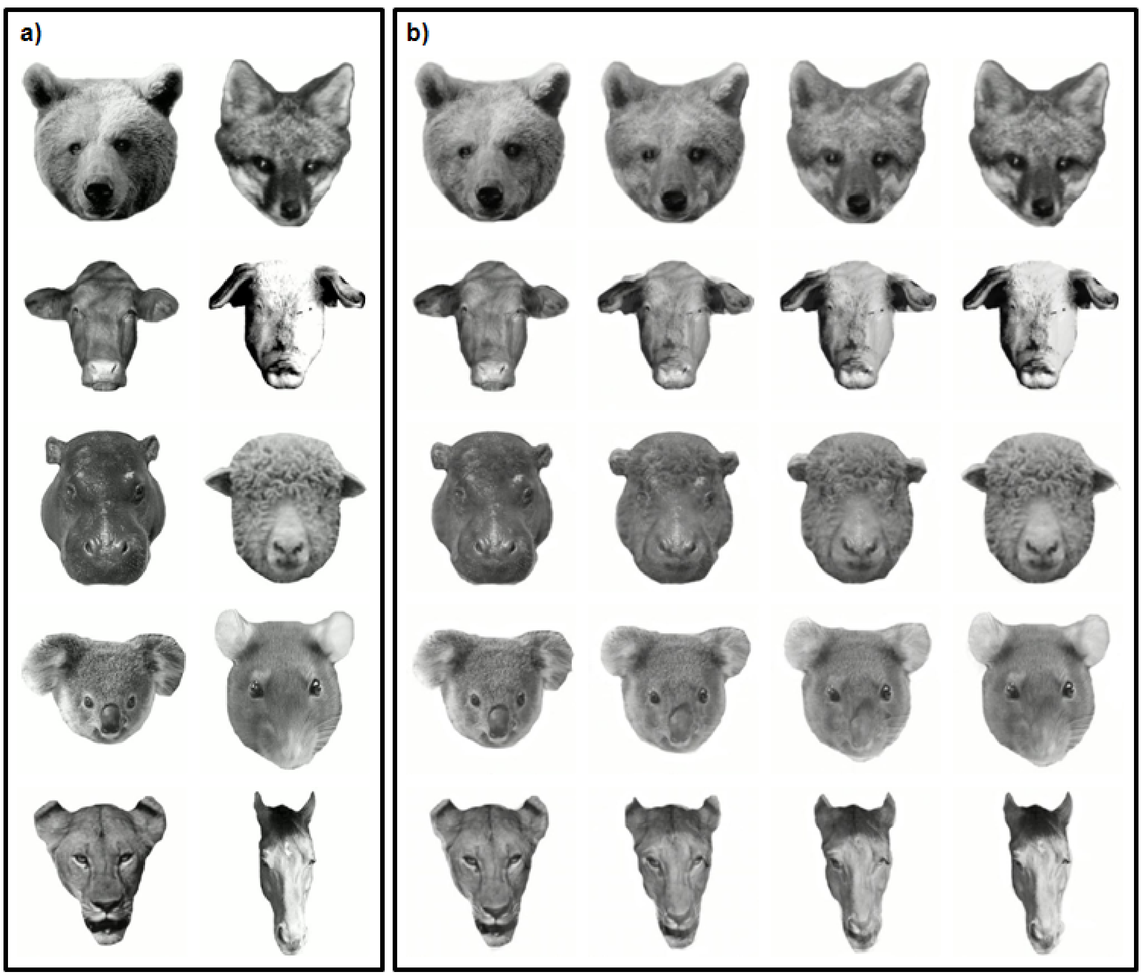
Yu, N. Y., Yamauchi, T., & Schumacher, J. (2008). Rediscovering symbols: The role of category labels in similarity judgment. *Journal of Cognitive Science, 9*, 89-109.

APPENDIX A

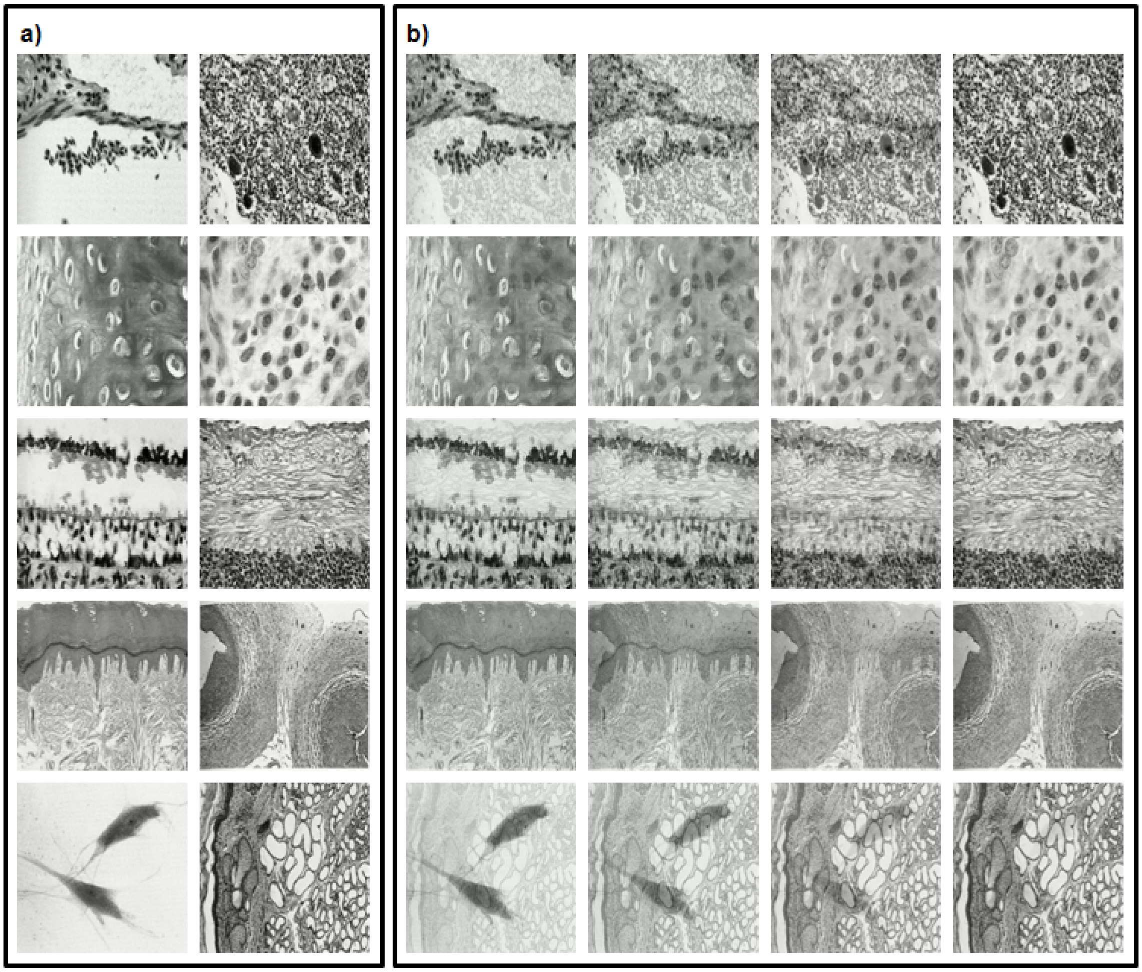
SAMPLE STIMULI

Appendix A shows sample pictures of stimuli I used in the experiments—animal faces, animal tissues, and real paintings. For each pair, one original picture was merged with the other original picture in 18 degrees, producing 18 morphed pictures. For each pair, two original pictures (a) and four examples of 18 morphed pictures (b) are shown.

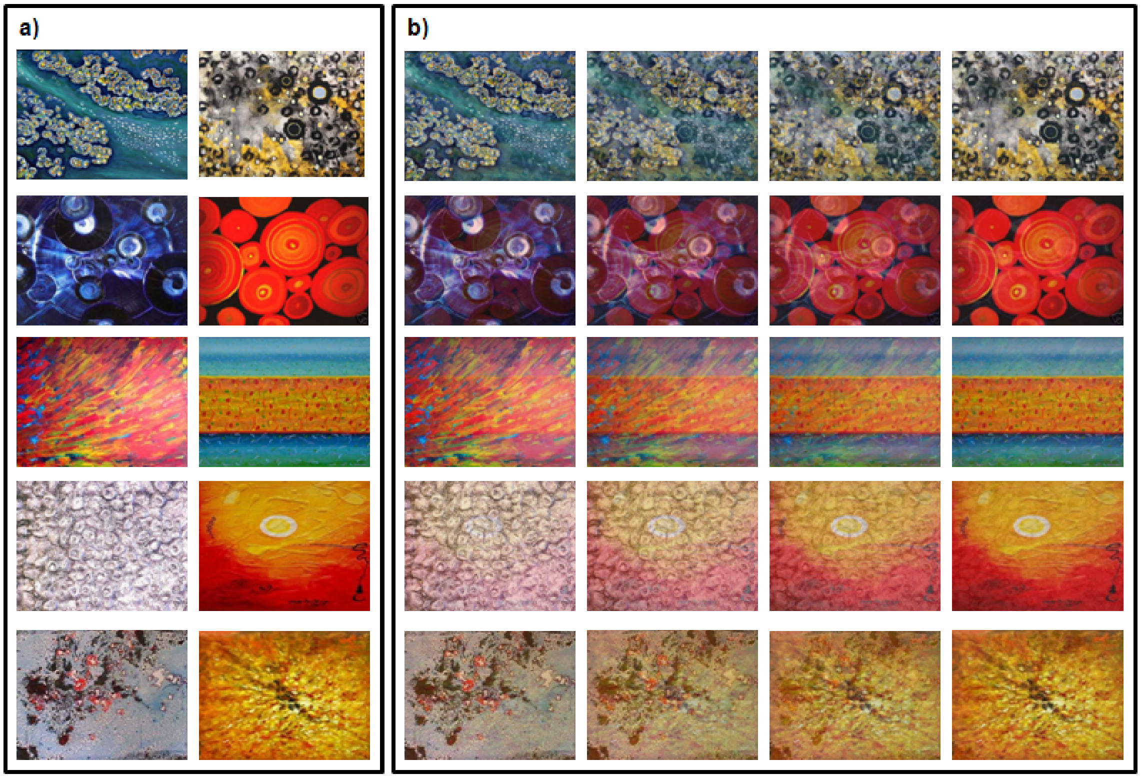
Animal face pictures



Animal tissue pictures



Real paintings



APPENDIX B

ADDITIONAL TABLES

The mean proportion of participants selecting the dissimilar picture was lower in the different-label condition (when the dissimilar picture had a different label from the target) compared to the no-label condition replicating the labeling effect observed in Experiment 1 (Table B-1) and 2 (Table B-2).

Table B-1

Mean proportion of participants selecting the dissimilar picture in Experiment 1.

	No-label	Different-label
Mean	.15	.05 *

Note. Significance level indicated by * ($p < .05$) between the no-label condition and the different-label condition.

Table B-2

Mean proportion of participants selecting the dissimilar picture in Experiment 2.

	No-label	Different-label			
		Animal cell	Diseased cell	Painter	Picture type
Mean	0.17	.12 *	.08 *	.11 *	.16

Note. Significance level indicated by * ($p < .05$) between the no-label condition and the different-label conditions.

VITA

Name	Na Yung Yu
Address	Dept. Psychology, Texas A&M University, 4235 TAMU College Station, Texas, USA
Email address	nayungyu@gmail.com
Education:	B.A., Psychology, English Language and Literature, Pusan National University, South Korea, 2002
	M.A., Psychology, Pusan National University, South Korea, 2004
	Ph.D., Psychology, Texas A&M University, USA, 2010