

Theories of Cognition and Image Categorization: What Category Labels Reveal About Basic Level Theory

Abebe Rorissa and Hemalata Iyer

Department of Information Studies, University at Albany, State University of New York, Albany, NY 12222.

E-mail: {arorissa, hi651}@albany.edu

Information search and retrieval interactions usually involve information content in the form of document collections, information retrieval systems and interfaces, and the user. To fully understand information search and retrieval interactions between users' cognitive space and the information space, researchers need to turn to cognitive models and theories. In this article, the authors use one of these theories, the basic level theory. Use of the basic level theory to understand human categorization is both appropriate and essential to user-centered design of taxonomies, ontologies, browsing interfaces, and other indexing tools and systems. Analyses of data from two studies involving free sorting by 105 participants of 100 images were conducted. The types of categories formed and category labels were examined. Results of the analyses indicate that image category labels generally belong to superordinate to the basic level, and are generic and interpretive. Implications for research on theories of cognition and categorization, and design of image indexing, retrieval and browsing systems are discussed.

Introduction

Categorization is crucial to human perception and cognition. Without their abilities to categorize, people would not be able to deal with the multitude of things they perceive daily. In terms of the use and role of categorization in information organization and retrieval, the general topics of browsing in information retrieval as well as its merits have been the subject of past research (Bates, 1989, 2007; Börner, 2000; Chang & Rice, 1993; Layne, 1994; B. O'Connor, 1988; B.C. O'Connor, 1993; O'Connor, O'Connor & Abbas, 1999; Vailaya, Jain, & Zhang, 1998). In anticipation of the growing number and size of image collections today and the challenges such growth will pose, Layne (1994) not only identified the merits of browsing, she also called for the

exploration of ways and means for the provision of access to categories of images in addition to access to individual images. Layne (1994) appropriately raised the question of what should be the basis for the categories of images. Like the other researchers such as Rasmussen (1997) and Jørgensen (1995), Layne (1994) posits that not only should image groupings, but also their indexing should be based on the attributes of the individual images.

The question of how to generate useful groupings of documents to support browsing and exploratory search is still an open one (Hearst, 2006). Although a meaningful grouping of information is desired by users, one of the main problems remains to be the difficulty of labeling the groups (Hearst, 2006). Layne's (1994) question about the basis for image category browsing and the fact that the choice of labels of groups should be based on relevant research prompted us to examine the nature of labels assigned by users to groups of images, thereby allowing us to discover the types of attributes people use to categorize images. This was done within the framework of the basic level theory.

Use of the basic level theory to understand human categorization is both appropriate and essential to user-centered design of taxonomies, browsing interfaces, and ontologies as well as other indexing tools and systems (Iyer, 1995). Basic level theory was chosen because existing taxonomies used to study image descriptions and image group labels were constructed inductively (they begin with the descriptions) and they were constructed with a specific purpose in mind. On the other hand, basic level theory provides a neutral, hierarchical, and general model, which could be used deductively. What is more, it is widely applied to study the nature of categories of objects (e.g., Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976), environmental scenes and concepts (e.g., Tversky & Hemenway, 1983), discourse (e.g., Wisniewski & Murphy, 1989). Our goal in using basic level theory was to find out the level of abstraction (*vis-à-vis* the three levels: subordinate, basic, and superordinate) of labels of groups of images supplied by human participants of two studies.

Received July 19, 2007; revised December 21, 2007; accepted December 27, 2007

© 2008 ASIS&T • Published online xx xxxxxx 2008 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/asi.20825

The search for a way to identify various types of attributes of individual images is not new (e.g., Enser & McGregor, 1992; Jörgensen, 1995, 1998). Previous researchers have constructed and used various taxonomies of image attributes. Even though there are studies that looked at descriptions of images and/or image queries with respect to various categories of attributes (Enser & McGregor, 1992; Jörgensen, 1995, 1998), no previous research looked at category labels with respect to the basic level theory and the three levels of categories of subordinate, basic, and superordinate. Jörgensen (1995), although she did not look at category labels specifically, did analyze attributes in terms of the three levels and found that, in a sorting task, groupings were at the superordinate level whereas subgroups were at the basic level. Jörgensen (1995) also suggested further research in this area and, recently, Green (2006) called for such research to be extended to the realm of standard subject access tools and pointed out that “The implications of the universality-of-basic-level-concepts hypothesis for complex or precoordinate classes has received as yet essentially no attention” (p. 12).

To address that, we need to first look into the nature of information organization, search, and retrieval interactions and either develop or adopt ways, theories, and models to understand it. Information search and retrieval interactions usually involve information content in the form of document collections, information retrieval systems and interfaces, and the user. The document collections and retrieval systems and interfaces exist in the information space; user’s needs and activities exist in the cognitive space of these interactions (Ingwersen, 1992, 1994). We adopted a mixture of the definitions of information space and cognitive space given by Newby (2001) and Ingwersen (1992). A cognitive space is the collection of concepts, information need, problem space, work-task/interest, and work domain(s) of a human information user. On the other hand, an information space is the set of various representations of objects and semantic entities (e.g., documents and their surrogates) in a system and may include documents (including image documents), words, and relations among the documents and words.

Categories of objects and concepts associated with them have levels of generality and abstraction. Taxonomies of domains of objects are built with this as the underlying assumption even though it is not always explicitly stated. For instance, the Dewey Decimal Classification System divides the universe of discourse into 10 classes, each of which is further subdivided into 10 sections with lower levels of generality and/or abstraction. Because categorization is central to how people organize the world around them, as well as how libraries and information systems organize information, such cognitive models help researchers structure their investigation and delineation of this central concept. Categorization has a dual purpose of grouping similar objects together and separating dissimilar objects (Tversky & Hemenway, 1983). It is also crucial to human perception, cognition, and the processes of learning, recognition, and judgment (Tversky & Gati, 1978). Furthermore, categorization “enables us to

reduce the numbers of entities in the world to manageable proportions” (Tversky & Hemenway, 1984, p. 170). Some even argue that “an understanding of the categories people use and how they assign concepts to those categories is one of the central clues to the understanding of human behaviour” (Canter, Brown, & Groat, 1985, p. 79).

In light of this, our work focuses mainly on the topic of image categorization and labels assigned by people to groups of images. Although almost all concept-based (manual) indexing of images is based on one or more attributes of the images (Jörgensen, 1995; Rasmussen, 1997), the question of what types of attributes should be considered for indexing both individual and image collectives is still not fully answered. In both concept- and content-based image indexing and retrieval, the focus is usually on individual images. However, there is a need to examine indexing and representation of image collectives because, among other things, “categories have the potential to reduce the search space and, thus, search time, because there are fewer categories than the total number of individual members” (O’Connor, O’Connor, & Abbas, 1999, p. 681). Doing so within the general framework of the cognitive viewpoint (Belkin, 1990) or, specifically, with the help of established cognitive models and theories, such as the basic level theory, could enhance our understanding of human categorization behavior. How people categorize images and the type of labels they use to name the categories may provide some clues as to what attributes they generally pay attention to when viewing them, and hence provide some answers to the question of what types of attributes should be used to index groups of images.

The study of image categorization by people and incorporating results from such studies into the design of information systems for browsing purposes should be informed by relevant theories and models. To fully understand information search and retrieval interactions between users’ cognitive space and the information space, researchers need to turn to cognitive models and theories. One of these theories is the basic level theory, which states that categories of objects have a hierarchical structure with three levels of abstraction, namely superordinate (e.g., furniture), basic (e.g., chair), and subordinate (e.g., rocking chair; Rosch & Mervis, 1975; Rosch et al., 1976). “At the basic level, a concept can be comprehended most easily as a whole, becoming a gestalt, or more than the sum of its parts” (Iyer, 1995, p. 46).

According to Rosch et al. (1976), human concepts have three levels of categorization matching the above three levels of abstraction in taxonomies of natural categories. Instances of basic level concepts (e.g., chair) are generally more similar than are those of superordinate concepts (e.g., furniture). This is due, in part, to the fact that instances of basic concepts share more common attributes than instances of superordinate concepts do (Rosch et al., 1976). For instance, various types of chairs have more in common than various types of furniture (e.g., chair, table, bed). There is an even greater similarity among instances of subordinate concepts (e.g., rockers) than among instances of basic or superordinate concepts (Rosch et al., 1976). In general, while the

degree of within-category similarity is high for basic level and subordinate categories, the within-category similarity of members of superordinate categories is low (Markman & Wisniewski, 1997). What is more, the degree of between-category similarity is low for both basic level and superordinate categories, though the between-category similarity of subordinate categories is high, especially if they belong to the same or neighboring basic level categories (Markman & Wisniewski, 1997; Murphy & Brownell, 1985). Due to this, superordinates are more distinct compared to subordinates and basic level category members. Hence, "people are less likely to mistakenly classify something at the superordinate level" (Markman & Wisniewski, 1997, p. 68).

Research by psychologists (Horton & Markman, 1980; Mervis & Crisafi, 1982) confirmed that children learn the more concrete terms before abstract ones and nouns are learned before verbs (Gentner, 1982). As a result, not only do children learn basic level concepts before superordinate ones, as adults, they continue to use basic level terms more than they do superordinates in their discourse (Wisniewski & Murphy, 1989). Two of the contributing factors to this high level of basic level terms are the higher frequency of basic level terms compared to superordinate or subordinate terms (Markman & Wisniewski, 1997) and the fact that, when referring to objects, the basic level is the most useful level of abstraction, at least for novices (Rosch et al., 1976; Tanaka & Taylor, 1991). Markman and Wisniewski (1997) argue that perhaps the reason why children acquire basic level terms sooner than superordinates is because in social situations, within which they learn categories, basic level names are short and appear in high frequency. There also seems to be a fundamental difference between children's and adults' view of superordinate categories. Whereas children view superordinate category members as collections with a part-whole structure (e.g., a father is part of a family, or an arm is part of a body), adults view them as classes whose members have an ISA relation to them (e.g. a car is a, ISA, kind of vehicle; Wisniewski & Murphy, 1989).

In summary, the majority of previous research that dealt with the issue of image indexing and cognitive models focused mainly on individual image attributes rather than groups of images and labels assigned to those groups. In this article, we use the basic level theory to investigate what user-supplied image category labels reveal about the underlying nature of human categorization and category naming. It has been suggested that the preferred level of categorization when people are thinking about the world around them is the basic level (Murphy & Wisniewski, 1989; Rosch et al., 1976). The basic level is where linguistic, perceptual, functional, and cognitive concepts converge. Apparently, this is true when the task involves identification of single and isolated objects rather than label collections of objects (Murphy & Wisniewski, 1989). Whether this holds true for groups of images as revealed by labels assigned to them by users is the focus of this article. To this end, we conducted two studies involving free sorting of 100 images and analyzed category labels supplied by 105 participants using content analysis to

see if the basic level theory applies to categorization of groups of images. More specifically, our goal was to identify the underlying structure of image category labels supplied by participants of two studies. Based on an extensive analysis of category names in written text, Wisniewski and Murphy (1989) conclude that superordinate terms are used more to refer to groups of objects. An analysis of human subjects' categorization of pictures of objects isolated or in scenes led to a similar conclusion (Murphy & Wisniewski, 1989), whether this will hold for image categorization and labels supplied by users remains to be seen.

Methodology

Studies 1 and 2

Study 1 was conducted in March 2003 using the free-sorting method (Coxon, 1999) for data collection, where participants were asked to sort a sample of 50 images into a number of groups of similar images, without any constraints on time and number of groups, label and describe each group, as well as provide a prototype image and feature. Study 2 was similar to Study 1, but was conducted between March and November 2006.

Materials

For both studies, a separate random sample each of 50 color digital images was selected from disk number 6 (which contains all photographs in the "people" category) of the *Hemera Photo Objects Volume I*, a stock photo collection of over 50,000 images (<http://www.hemera.com>). Each image was printed on a 4-by-5-inch (10.2-by-12.7-cm) card.

Participants

Participants of Study 1 were 30 graduate students (16 women and 14 men) at a major southwestern U.S. university who volunteered to participate in the study. Those who participated in Study 2 were 75 graduate students (59 women and 16 men) at a major northeastern U.S. university who also volunteered to participate in the study. Participants of both studies were between 21 and 60 years old except one of the participants of Study 2 who was over 60. In Study 1, 13 of the participants were in the Masters program, 16 in the PhD program, and one was working on a Certificate of Advanced Studies (CAS). Participants in Study 2 were all Masters students. Only 7 of the participants in Study 1 and 8 in Study 2 had a degree in a visual arts related field. Most (83%) of the participants of Study 1 either worked with or used digital images at least once a week; 72% of the participants of Study 2 did so at least once every 2 weeks.

Procedures

Data collection. Participants of both studies were instructed to inspect the images first and then sort them into as many or as few groups as they wished using their own general

criteria for similarity. They were also told to feel free to rearrange, break, or remake the groups until they reached an arrangement that was satisfactory to them. On a data collection form, participants wrote down labels and longer descriptions, selected a prototype image for each category, and wrote down the corresponding significant feature of the prototype image. In the instructions, participants were instructed to choose an image that they think is a prototype (that best represents the group) for each group. The concept of *prototype image* was explained to participants as an image that is a good representative of the group of images while a significant feature of the prototype image was explained as one feature that made them pick the image as a prototype. The cards were reshuffled before being given to the next participant. Table 1 presents the summary of the two studies.

Table 2 illustrates an example of the images grouped into a category together with the category label, longer description of the category, prototype image, and associated significant feature of the prototype image provided by a participant. In this case, the participant seemed to have paid attention to the setting including, but not limited to, the background in the images, and people's activities and clothing in labeling the category with "nature" and providing "leaves" as the significant feature of the prototype image for the category.

Data analysis. Content analysis requires the creation and testing of a coding scheme that involves the definition of recording units and categories, assessment of the accuracy of coding, revision of coding rules, and coding the entire text

(Weber, 1990). The image category labels (including single words and "word senses") were the recording units (basic unit of text to be coded or categorized). A word sense could be a phrase constituting a semantic unit such as idioms ("taken for granted"), proper nouns (the Empire State Building), as well as other phrases (e.g. "sweet home," "life routine"; Weber, 1990).

A coding scheme/dictionary (Appendix) was developed based on the definition, explanations, and examples provided in the basic level theory literature and several discussions between the two authors. After testing it on a random sample of image category labels (10% from each study), the final version of the coding scheme/dictionary was developed. Using this scheme, the entire sample of category labels was then coded independently by the two authors. Then, the authors compared their coding and resolved the differences through several discussions until a 100% agreement was reached. In addition, a third coder, a graduate student blind to the purpose of the study, coded a random sample (20% from each study) of image category labels. Two popular intercoder agreement measures, percent agreement and Cohen's (1960) kappa, were used to measure the reliability of coding between the authors and the third coder. The computed values of the percent agreement and Cohen's kappa were 0.91 and 0.79 for Study 1 and 0.89 and 0.80 for Study 2, respectively. Given that values over 0.70 are considered satisfactory in most situations (Neuendorf, 2002), the coding procedure was deemed to be reliable and appropriate.

Throughout our data analyses, we treated the two studies as separate due to purely methodological reasons. Specifically, we intended to show that results obtained from the two studies are similar and that they have both internal (which we have shown, partly, through intercoder agreement analysis) and external reliability (consistency). We conducted the chi-square test of independence and *F*-tests to establish this and the homogeneity of the two samples from Studies 1 and 2.

TABLE 1. Summary of the two studies.

Study details	Study	
	1	2
No. of images	50	50
Type of images	People	People
Method of similarity judgment	Free sorting	Free sorting
No. of participants	30	75
Total number of groups formed	240	659
Minimum no. of groups	3	2
Maximum no. of groups	7	24
Mean no. of groups	8	8.79
Median no. of groups	7	8
Mode no. of groups	7	9
<i>SD</i> (No. of groups)	3.3	4.1

Findings

In total, participants of the two studies supplied 899 category labels (240 in Study 1; 659 in Study 2). In relation to the types of terms used to refer to individual and groups of objects, basic level terms are often used to refer to the former while superordinate terms are used to refer to the latter. According to Wisniewski and Murphy (1989, p. 249) "If superordinate representations include information about groups of objects, we would expect to find these terms usually referring to multiple objects. Likewise, if basic category terms are used primarily to refer to single objects, we would expect these terms to refer to single objects most of the time." Because participants of our studies were asked to sort images into groups and label the groups, we set out to determine to what extent they use singular, plural, and generic terms. As shown in Figure 1, participants of both studies used less singular terms (referring to a single object, e.g., gunman; 9.6% and 9% in Studies 1 and 2, and 9.1% over all, respectively) than plural terms (referring to more than one object,

TABLE 2. Example of a category of images.

	Category label	Nature
	Longer description	Images of people in natural environment
	Prototype image	
	Significant feature	Leaves

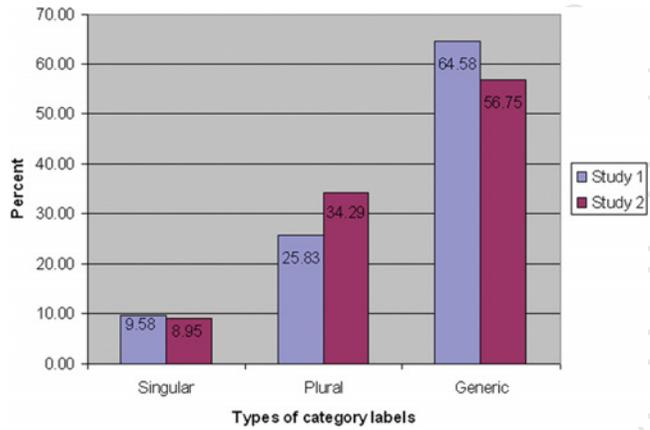


FIG. 1. Percentage of singular, plural, and generic category labels.

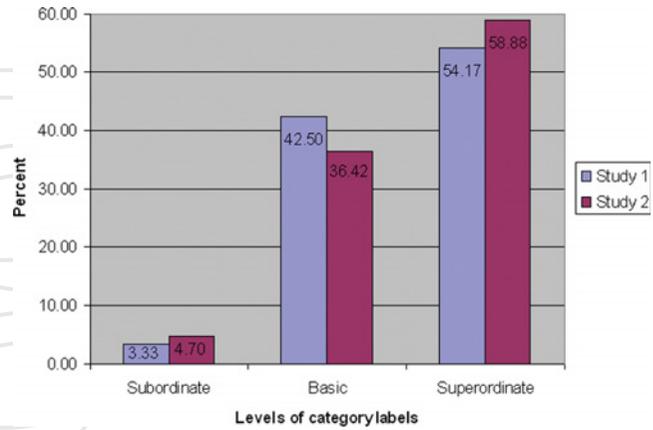


FIG. 2. Percentage of subordinate, basic, and superordinate category labels.

e.g., clowns; 25.8% and 34.3% in Studies 1 and 2, and 32.0% overall) and generic terms (not referring to a quantity, e.g., relaxing terms; 64.6% and 56.7% in Studies 1 and 2, and 58.9% overall) as category labels. Generic terms were the most often supplied category labels. Generic refers to the whole class of objects or concepts rather than to specific ones. For example, consider a picture of a child with a jump rope. The label “play” or “playing” will be coded as a generic term, whereas the label “girl” will be coded as a “singular” term. Another example is that of a man sitting on the couch doing nothing. The caption “relaxing” will be considered to be a generic term as it is not about the man or objects in the picture, but about a feeling or the interpretation of the picture. In the case of terms coded as plural, sometimes a picture had one child with a backpack and the label assigned was “school girls.”

We conducted a chi-square test of independence to see if the number of singular, plural, and generic category labels supplied by participants of Studies 1 and 2 are independent of the samples of images and participants. The test produced a nonsignificant chi-square ($\kappa^2 = 5.841, df = 2, p > .05$), which led us to assume that the types of category labels (Figure 1) are independent of the samples of images and participants; hence, the two samples (and corresponding populations) are not significantly different in terms of the number of singular, plural, and generic category labels.

Out of the 899 category labels supplied by participants of both studies (Table 3), as shown in Figure 2, there were

fewer subordinate level terms (3.33% and 4.7% in Studies 1 and 2, and 4.34% overall, respectively) than basic level terms (42.5% and 36.42% in Studies 1 and 2, and 38.04% overall) and superordinate level terms (54.17% and 58.88% in Studies 1 and 2, and 57.62% overall). Superordinate level terms were the most often supplied category labels. These numbers and percentages match the number of singular, plural, and generic category labels (Figure 1).

Once again, a chi-square test of independence was conducted to see if the number of subordinate, basic, and superordinate category labels is independent of the samples of images and participants. We obtained a nonsignificant chi-square ($\kappa^2 = 3.150, df = 2, p > .05$) and conclude that the levels of category labels (vis-à-vis the basic level theory; Figure 2) are independent of the samples of images and participants and the two samples (and corresponding populations) from Studies 1 and 2 are not significantly different in terms of the number of category labels that are subordinate, basic, and superordinate.

To test whether or not the two samples (from Studies 1 and 2) have different variances in terms of the number of category labels at the three levels, we used the *F*-test. Statistically nonsignificant *F*-values, subordinate: $F(74, 29) = 0.75, p > .05$; basic: $F(74, 29) = 1.13, p > .05$; superordinate: $F(74, 29) = 0.75, p > .05$, led us to conclude that the corresponding number of subordinate, basic, and superordinate category labels from the two studies satisfy the homogeneity of variance criterion and have equal variances. An independent groups *t* test (with equal variances) was then performed to compare the corresponding mean number of subordinate, basic, and superordinate level category labels from Studies 1 and 2 (Table 4 and Figure 3). For all the three levels, the *t* test confirmed a nonsignificant difference, subordinate: $t(103) = 0.96, p > .05$; basic: $t(103) = 0.35, p > .05$; superordinate: $t(103) = 1.65, p > .05$. Therefore, we can conclude that there were no significant differences between the mean number of subordinate, basic, and superordinate level category labels from Study 1 and the corresponding mean number of subordinate, basic, and superordinate level category labels from Study 2, respectively.

TABLE 3. Examples of category labels supplied by participants.

Level of category label	Sample category labels
Subordinate	Attack Dog, Dress Hats, Halloween Costume, Heterosexual Couples, Satchel, School Girls, Smiling Face
Basic	Bag, Costume, Face, Girls, Happiness, Hunting, Party
Superordinate	Animal, Clothes, Emotion, Most Body, Outdoors Activity, People, Sports

AQ1

AQ3

AQ1

TABLE 4. The mean and standard deviation of number of subordinate, basic, and superordinate category labels supplied by participants of Studies 1 ($n = 30$) and 2 ($n = 75$).

Study	Subordinate		Basic		Superordinate	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
1	0.27	0.69	3.40	2.37	4.33	2.29
2	0.41	0.72	3.20	2.76	5.17	2.38

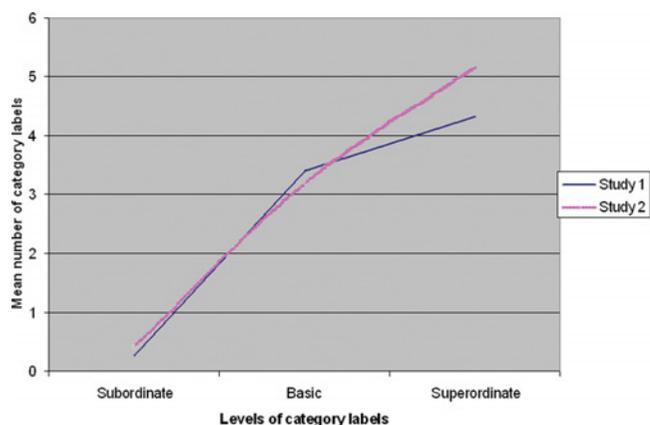


FIG. 3. Mean number of subordinate, basic, and superordinate category labels.

In a final analysis, we examined the significance of the difference in the number of image category labels at the three levels of abstraction (Table 4 and Figure 3). Due to unequal variances, especially between the number of superordinate and subordinate as well as basic and subordinate level category labels, data for both studies were transformed (by taking the square root of the number of groups of images) to satisfy the homogeneity of variance criterion. Using Levene's test for homogeneity of group variances, the transformation yielded statistically nonsignificant differences between all possible pairs of variances for both studies. A one-way (single factor—levels of category labels) analysis of variance (ANOVA) revealed that there was a significant overall difference between the number of subordinate, basic, and superordinate level category labels for each of the two studies, Study 1: $F(2,87) = 59.3, p < .001$; Study 2: $F(2,222) = 128.3, p < .001$. Tukey's honestly significant difference (HSD) post hoc test revealed that for Study 2, all three pairs of means were significantly different from each other, whereas the difference between the mean number of superordinate and basic level category labels were not significant for Study 1. This finding may be attributed to the small sample size. For Study 2, there were significantly higher mean numbers of superordinate category labels ($M = 5.17$) than the mean number of basic category labels ($M = 3.20$). For both studies, the mean number of superordinate (Study 1: $M = 4.33$; Study 2: $M = 5.17$) and basic (Study 1: $M = 3.40$; Study 2: $M = 3.20$) category labels was significantly higher than the mean number of subordinate category labels (Study 1: $M = 0.27$; Study 2: $M = 0.41$).

Discussion

Our main objective in this study was to identify the underlying structure of image category labels supplied by participants of two studies vis-à-vis the basic level theory and to determine the types of labels supplied (i.e., to what extent labels at the subordinate, basic, and superordinate level of abstraction were supplied). That is, we sought to determine whether people often use superordinate, basic, or subordinate level category labels to name groups of images they sorted. People may predominantly use basic level terms in free-naming tasks (Rosch et al., 1976) and adult discourse contains more basic level than superordinate level terms (Wisniewski & Murphy, 1989). Contrary to the fact that the preferred level of categorization when people are thinking about the world around them is the basic level (Murphy & Wisniewski, 1989; Rosch et al., 1976), when it comes to labeling image groups, our findings above showed that people generally tended to use more superordinate level terms than basic and subordinate level terms; subordinate level terms were hardly used. This is more in line with Wisniewski and Murphy's (1989) assertion that superordinate terms are used more to refer to groups of objects.

Our findings show that not only do the proportion of the number of category labels that are subordinate, basic, and superordinate remain the same for both studies (an indication of the reliability or consistency of our results), a significantly higher number of superordinate category labels than both basic and subordinate category labels were supplied by participants while they supplied hardly any subordinate labels. In a previous study, Jörgensen (1995) found that although 73% and 22% of the attributes listed by subjects in an image description task were perceptual and interpretive, respectively, an image sorting task yielded more interpretive attributes (68%) than perceptual attributes (20%). Perceptual attributes are those that depict the physical characteristics of the image such as objects, color, visual elements, and others. Interpretive attributes are those that need more than just perception. Features that fall under this category are abstract concepts, external relationships, content, etc. (Jörgensen, 1995). Jörgensen (1995) also analyzed attributes in terms of the three levels and found that, in a sorting task, groupings were at the superordinate level while subgroups were at the basic level. She (Jörgensen, 1996) also found that attributes of images used by subjects during the sorting task (which yielded more interpretive attributes) match those of a searching task. Because interpretive features involve higher levels of abstraction, they are more akin to superordinate level features. According to Fidel (1997), values of interpretive attributes need "personal reflection and abstraction, and therefore might be in the eyes of the viewer" (p. 184).

Recent studies that dealt with similar topics conclude that people use more interpretive features to categorize images. For instance, Greisdorf and O'Connor (2002) confirmed the fact that people tend to use more interpretive than perceptual attributes to label groups of images, unless they were told to use perceptual attributes as well. Laine-Hernandez and

Westman (2006) found that high-level interpretive concepts were used in an image categorization exercise by journalists. This is perhaps further proof that groups of images need to be indexed using interpretive attributes and/or superordinate terms. However, perceptual attributes may also be useful, especially for users with visual arts background, in indexing and searching for images. We should note that some concepts that are superordinates for experts could be considered basic concepts for novices (Wisniewski & Murphy, 1989). The difference in experts' and novices' knowledge of subordinate category attributes could also be reflected in their feature listing or category names (Tanaka & Taylor, 1991). Tanaka and Taylor (1991) have also shown that domain-specific knowledge affects whether people categorize objects at the basic, superordinate, or subordinate levels. One of their findings is that experts named objects using subordinate level names with the same frequency as they named using basic level names. However, because our participants in both studies were novices in image creation and processing and had similar levels of domain knowledge, our results are free of this confounding factor.

Two other possible factors are the types and sample sizes of images used in the two studies. We believe that the types of images do not have a bearing on the kind of labels the participants would use because they were asked to group together images that they perceived as similar and label those groups. They were free to label them using terms at all three levels. However, 4.34% of the labels were at the subordinate level, 38.04% were at the basic level, and 57.62% were at the superordinate level. We also believe that the homogeneity or lack thereof in the samples of images may not have had any effect on the type of labels used. Both our samples of images were from a relatively homogeneous category (people) consisting of images with a lot in common. The above results were obtained despite that. This is because, based on previous literature that applied basic level theory to study categorization of objects, concepts, and discourses, people, in general, used superordinate level terms to label top level groups. There is also a clear distinction between the level of terms people use to label groups of images and terms they use to describe individual images. One of the authors of this work (Rorissa, 2007), using heterogeneous samples of images as materials, found that while descriptions of individual images are predominantly at the basic level, the majority of group labels were at the superordinate level. We believe part of the reason is that an image is not just of a single object or about a single subject and any group of images will require the use of a higher level term as a label to represent the *ofness* and *aboutness* of the images in the group.

Concerning sample size (or the number of images being sorted), both of our studies reported in this work were based on two separate samples of 50 images each. It is plausible that if people are provided with a very large collection/sample of images and asked to sort them into groups of similar images and label the groups, they may use more subordinate or basic level terms than were used by participants of our two studies. It would also be interesting to see the

difference in the results if the users are asked to form subgroups, in addition to the top level groups, of a very large sample of images. To ascertain this, we suggest that this study be replicated with a very large sample of images that is also heterogeneous in terms of the types of contents of the images. Our study provides an understanding and insight into the use of the basic level theory and a possible way to apply this theory to image organization. However, because this is an exploratory research, and the sample size and the types of images can have a significant impact on the results, there is a need for further research that considers these factors.

Implications for Indexing

Given the above results and discussion, what are their implications for image indexing? Generally, grouping and presenting thumbnails of images in a large collection together with index terms and captions facilitates browsing. Users can select an image, especially in a query by example system, or can choose to browse through the collections. Unlike directed search, browsing, although goal directed, is appropriate when a user has a fuzzy idea of a search topic or a not well-defined information need and query, or has difficulty formulating effective queries. It involves looking through the images to view and retrieve the ones that have likelihood of matching the searcher's interests. Hence, presenting aggregates of image thumbnails are a helpful means of browsing the image collections. While indexing image collectives, it is recommended that superordinate terms be used, at least for the top level groups. They tend to be names of classes, such as, fruits, people, nature, etc., and serve as excellent grouping mechanisms. This will also help improve interoperability between image indexing in the various databases and systems. Although our discussions mainly focused on groupings of images and how appropriate the superordinate level is for labeling and indexing such groupings, any hierarchical browsing interface and menu should also provide access through all three levels. In addition, it may be helpful to provide index terms representing the most significant feature (prototypical feature) symbolizing that class. For example, in one of our studies, the category label used to name a group of images consisting of people displaying some sort of emotional behavior was "emotions" and the significant feature of the prototype image suggested was "facial expression." The latter might serve as a cue to the recognition of the items within the collective. Another interesting example of a significant feature of the prototype image was "outdoor activity," which alluded to the background instead of the people in the picture. This could serve as distinguishing feature because all pictures had people in them. Table 5 provides two examples of a category label ("emotions") and the corresponding prototype images.

Another example from the study furthers the argument. Pictures of yoga were grouped along with pictures of people playing football, camping, hunting, etc., with "sports" as a category label and the significant feature of the prototype

TABLE 5. Two prototype images with category label “emotions.”

		
Detailed description	Expression of emotion	Facial and body expression of emotion
Significant feature	Head down	Happy smiling woman

image was “football.” Obviously, here the focus has been on the physical aspect of yoga, as a means for physical fitness. Hence, these pictures were grouped into a “sports” category. Another category label under which yoga pictures were grouped was “spirituality” with the significant features of the prototype image being “meditation” and “classic poses.” Here, the participants seem to have focused on the esoteric and spiritual aspects of yoga. This points to the fact that because categorization is individualistic, it will be helpful to include the significant features of the prototype images as index terms together with the category labels. This will provide an inclusive representation of the needs of various types of users.

The question then is which level of superordinate classes do we pick? Can a particular level be recommended as being more appropriate for indexing image collectives? Would it be the same level regardless of the image content? These are questions that need to be addressed. However, what is clear is that the idea of basic level can be very helpful in determining it. We may not yet have a robust automated methodology for identifying the basic level category, concepts, and terms to describe/name them. However, there are studies that have addressed this issue and suggested methodologies. The use of WordNet is one such methodology (Green, 2006). The basic level can be intuitively and manually determined using a set of criteria as suggested and used in this article (Appendix). Once this is accomplished for a given domain of images, the hierarchic structure (superordinate and subordinate) can be developed around the basic level. In conclusion, “Recognition of basic level categories is also relevant to indexing and information retrieval (IR) vocabularies. If indeed information clusters around a certain level in a hierarchy it will be useful to incorporate those basic level terms in the tools used for organizing and retrieving information and to include them in indexing” (Iyer, 1995, p. 47).

Concluding Remarks and Future Research

Results of the above two studies reveal that user supplied image category labels are predominantly at the superordinate level. Subordinate level category labels were hardly used by participants. Similar previous studies on the nature of categorization of images strongly support these results (Jørgensen, 1995, 1996; Greisdorf & O’Connor, 2002; Laine-Hernandez & Westman, 2006). Rosch et al. (1976) have also observed that sorting tasks that involved subjects to group things together lead to categorization at the superordinate level. It is

important to view the findings of these studies in the context of user-centered design of indexing tools and systems and social tagging. Social tagging is a fairly recent phenomenon and has generated a lot of interest and has tremendous potential. It may prove to be an alternative to professional indexing and/or may serve as a complement to professional indexing. More research in this area is needed regarding its role and its effectiveness in providing access to images. Understanding the nature of tags users assign to either individual or groups of images will shed some light on their cognitive space and pave the way to fully user-centered indexing. The findings of this research could be effectively used to examine and evaluate the indexing in image databases and digital library collections. It could also be applied to the study of image collections in the existing social tagging services such as flickr because these services are becoming more and more popular and extensively used.

With social tagging gaining popularity (and we should add that most of the people who do the tagging are indexing novices who use free-naming like participants of our studies) and the ongoing debate about user-centered indexing and how to incorporate the users’ language into the indexing process (Fidel, 1994; Matusiak, 2006; Soergel, 1985), user-supplied index terms are becoming not just viable alternatives to indexing by professionals, they could also inform the design of indexing tools and systems. One of the functions of these tools is to provide classificatory structures and taxonomies for index terms and the documents they are used to represent or index. For these tools to have a solid grounding in research and theory, studies such as ours (that used a model with an underlying hierarchical taxonomy) and other similar research should use appropriate theories and models to understand the nature of tags and labels people use to label individual and groups of images, including the level of abstraction of the tags. This will aid in the effort to design mechanisms such as taxonomies to facilitate browsing and exploratory search (Hearst, 2006; Marchionini, 2006). For instance, the construction of a thesaurus, more specifically, the determination of the relationships between broader and narrower terms, could be informed by whether user tags are at the superordinate or the basic and subordinate levels. Furthermore, an effective browsing and exploratory search interface and system could be enhanced through users’ annotations of the documents (Shneiderman, Bederson, & Drucker, 2006).

Even though categorization of images for browsing purposes has its advantages and will improve recall, it does so at the expense of precision (Layne, 1994). However, it should be noted that the main goal of browsing systems is not to maximize precision. It is rather to maximize recall and help users find information serendipitously. To this end, superordinate level terms are ideal candidates for indexing and labeling groups of images because they are often used by people to refer to groups of objects (Wisniewski & Murphy, 1989).

Our studies are an attempt at solving the bigger problem of indexing images for browsing purposes. We hope that results of this study will contribute to better understanding

the nature of image categorization by users that could feed into refining the cognitive models and theories, thereby leading to an improved design of image indexing, retrieval, and browsing systems. To our knowledge, our work is among the first that attempted to understand the nature of image categorization and image category labels using the basic level theory. Because this is one of the pioneering works in this area, future research into this will further clarify other aspects of labeling groups of images. This is a rich area, which has not been adequately explored and has a lot of potential for future research.

However, it is not without limitations. For instance, our use of images with people from a stock photo collection could be a possible limitation of our studies and future studies need to check if the type of images has any bearing on our results. Moreover, these results need to be verified using a much larger sample of images from a variety of sources and collections and subjects for them to be generalized. To assess the types of index terms in traditional indexing tools and systems vis-à-vis the basic level theory, future research should compare these results to findings from analyses of index terms in traditional indexing tools such as thesauri and subject heading lists. Future research should also look at the difference between category labels and descriptions of individual images vis-à-vis the basic level theory.

Acknowledgments

The first author is grateful to the School of Library and Information Sciences, University of North Texas, for financial and academic support. We are also indebted to Dr. Carol Doll, Stephen Maher, and the 105 participants of the two studies. The three anonymous reviewers deserve our gratitude for their extremely thoughtful and thorough comments that improved this article immensely.

References

- Bates, M.J. (1989). The design of browsing and berrypicking techniques for the online search interface. *Online Review*, 13(5), 407–424.
- Bates, M.J. (2007). What is browsing—really? A model drawing from behavioural science research. *Information Research*, 12(4), paper 330. Retrieved October 15, 2007, from <http://InformationR.net/ir/12-4/paper330.html>
- Belkin, N.J. (1990). The cognitive viewpoint in information science. *Journal of Information Science*, 16(1), 11–15.
- Börner, K. (2000, July). Searching for the perfect match: A comparison of free sorting results for images by human subjects and by latent semantic analysis. Paper presented at Information Visualisation 2000, Symposium on Digital Libraries, London, England.
- Canter, D., Brown, J., & Groat, L. (1985). A multiple sorting procedure for studying conceptual systems. In M. Brenner, J. Brown, & D. Canter (Eds.), *The research interview: Uses and approaches* (pp. 79–114). London/Orlando: Academic Press.
- Chang, S., & Rice, R.E. (1993). Browsing: A multidimensional framework. In M.E. Williams (Ed.), *Annual review of information science and technology* (pp. 231–276). Medford, NJ: Learned Information.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37–46.
- Coxon, A.P.M. (1999). *Sorting data: Collection and analysis*. Thousand Oaks, CA.: Sage.
- Enser, P.G.B., & McGregor, C.G. (1992). Analysis of visual information retrieval queries. Report on Project G16412 to the British Library Research and Development Department. London: British Library.
- Fidel, R. (1994). User-centered indexing. *Journal of the American Society for Information Science*, 45, 572–576.
- Fidel, R. (1997). The image retrieval task: Implications for the design and evaluation of image databases. *The New Review of Hypermedia and Multimedia*, 3, 181–199.
- Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S.A. Kuczaj (Ed.), *Language development: Vol. 2. Language, thought, and culture* (pp. 301–334). Hillsdale, NJ: Erlbaum.
- Green, R. (2006). Vocabulary alignment via basic level concepts. OCLC/ALISE research grant report published electronically by OCLC Research. Retrieved January 24, 2007, from <http://www.oclc.org/research/grants/reports/green/rg2005.pdf>
- Greisdorf, H., & O'Connor, B.C. (2002). What do users see? Exploring the cognitive nature of functional image retrieval. In E.G. Toms (Ed.), *Proceedings of the 65th Annual Meeting of the American Society for Information Science and Technology* (pp. 383–390). Medford, NJ: Information Today.
- Hearst, M.A. (2006). Clustering versus faceted categories for information exploration. *Communications of the ACM*, 49(4), 59–61.
- Horton, M., & Markman, E. (1980). Developmental differences in the acquisition of basic and superordinate categories. *Child Development*, 51, 708–719.
- Ingwersen, P. (1992). *Information retrieval interaction*. London: Taylor Graham.
- Ingwersen, P. (1994). Polyrepresentation of information needs and semantic entities: Elements of a cognitive theory for information retrieval interaction. In W.B. Croft & C.J. van Rijsbergen (Eds.), *Proceedings of the 17th Annual international ACM SIGIR Conference on Research and Development in information Retrieval* (pp. 101–110). New York: Springer-Verlag.
- Iyer, H. (1995). *Classificatory structures: Concepts, relations and representation*. Frankfurt/Main: Indeks Verlag.
- Jørgensen, C. (1995). Image attributes: An investigation. Unpublished doctoral dissertation, Syracuse University, Syracuse, New York.
- Jørgensen, C. (1996). Indexing images: Testing an image description template. In S. Hardin (Ed.), *Proceedings of the 59th Annual Meeting of the American Society for Information Science* (pp. 209–213). Medford, NJ: Information Today.
- Jørgensen, C. (1998). Attributes of images in describing tasks. *Information Processing & Management*, 34(2/3), 161–174.
- Laine-Hernandez, M., & Westman, S. (2006). Image semantics in the description and categorization of journalistic photographs. In A. Grove & J. Steff-Mabry (Eds.), *Proceedings of the 69th Annual Meeting of the American Society for Information Science and Technology*. Retrieved June 11, 2007, from http://www.asis.org/Conferences/AM06/proceedings/papers/48/48_paper.html
- Layne, S.S. (1994). Some issues in the indexing of images. *Journal of the American Society for Information Science*, 45, 583–588.
- Marchionini, G. (2006). Exploratory search: From finding to understanding. *Communications of the ACM*, 49(4), 41–46.
- Markman, A.B., & Wisniewski, E. (1997). Similar and different: The differentiation of basic-level categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23(1), 54–70.
- Matusiak, K.K. (2006). Towards user-centered indexing in digital image collections. *OCLC Systems & Services*, 22, 283–298.
- Mervis, C.B., & Crisafi, M.A. (1982). Order of acquisition of subordinate, basic, and superordinate categories. *Child Development*, 53, 258–266.
- Murphy, G.L., & Brownell, H.H. (1985). Category differentiation in object recognition: Typicality constraints on the basic category advantage. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11(1), 70–84.
- Murphy, G.L., & Wisniewski, E.J. (1989). Categorizing objects in isolation and in scenes: What a superordinate is good for. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(4), 572–586.

Neuendorf, K.A. (2002). *The content analysis guidebook*. Thousand Oaks, CA: Sage.

Newby, G.B. (2001). Cognitive space and information space. *Journal of the American Society for Information Science and Technology*, 52(12), 1026–1048.

O'Connor, B. (1988). Fostering creativity: Enhancing the browsing environment. *International Journal of Information Management*, 8(3), 203–210.

O'Connor, B.C. (1993). Browsing: A framework for seeking functional information. *Knowledge: Creation, Diffusion, Utilization*, 15(2), 211–232.

O'Connor, B.C., O'Connor, M.K., & Abbas, J.M. (1999). User reactions as access mechanism: An exploration based on captions for images. *Journal of the American Society for Information Science and Technology*, 50(8), 681–697.

Rasmussen, E.M. (1997). Indexing images. In M.E. Williams (Ed.), *Annual review of information science and technology* (pp. 169–196). Medford, NJ: Learned Information.

Rorissa, A. (2007). User generated descriptions of individual versus labels of groups of images: A comparison using basic level theory. Manuscript submitted for publication.

Rosch, E., & Mervis, C. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7(4), 573–605.

Rosch, E., Mervis, C.B., Gray, W.D., Johnson, D.M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, 8(3), 382–439.

Shneiderman, B., Bederson, B.B., & Drucker, S.M. (2006). Find that photo!: Interface strategies to annotate, browse, and share. *Communications of the ACM*, 49(4), 69–71.

Soergel, D. (1985). *Organizing information: Principles of data base and retrieval systems*. Orlando, FL: Academic Press.

Tanaka, J.M., & Taylor, M. (1991). Object categories and expertise: Is the basic level in the eye of the beholder? *Cognitive Psychology*, 23, 457–482.

Tversky, A., & Gati, I. (1978). Studies of similarity. In E. Rosch & B.B. Lloyd (Eds.), *Cognition and categorization* (pp. 79–98). Hillsdale, NJ: Erlbaum.

Tversky, B., & Hemenway, K. (1983). Categories of environmental scenes. *Cognitive Psychology*, 15, 121–149.

Tversky, B., & Hemenway, K. (1984). Objects, parts, and categories. *Journal of Experimental Psychology: General*, 113(2), 169–193.

Vailaya, A., Jain, A., & Zhang, H.J. (1998). On image classification: City images vs. landscapes. *Pattern Recognition*, 31(12), 1921–1935.

Weber, R.P. (1990). *Basic content analysis* (2nd ed.). Newbury Park, CA.: Sage.

Wisniewski, E.J., & Murphy, G.L. (1989). Superordinate and basic category names in discourse: A textual analysis. *Discourse Processes*, 12, 245–261.

AQ2

Appendix

Coding Scheme/Dictionary Used on Category Labels

Level	Definition
Superordinate	<p>A term that is more generic than a given basic level term. Superordinate level terms:</p> <ul style="list-style-type: none"> • Are more generic and plural than specific and singular • Are terms for which one can think of several manifestations (examples) • Describe objects that share few common features and have more distinctive features • Describe more abstract or functional features • Describe groups, collections, or classes of objects • Are terms with at least four levels beneath them (in WordNet) • Convey more relational information (e.g., relative location) • Convey both functional information and general knowledge about the objects they refer to • For example, vehicle, furniture, food, work, style, personality, crime, transportation, housework
Basic	<p>A term that is more specific than a given superordinate level term. Basic level terms:</p> <ul style="list-style-type: none"> • Are singular rather than plural • Are shorter (15 or less characters long) and usually consist of one or two words • Are not phrases • Are included within the names of more specific terms (e.g., car in “sports car,” shirt in “dress shirt”) • Are more informative and they identify/reference single objects/features • Describe objects that share relatively more common features and are relatively similar • Describe concrete and perceptual entities (entities that can be perceived through the senses and mentally represented in a single image) • Describe salient features • Make distinctions or distinguish one object from another • Convey information on the parts and components of objects • For example, car, chair
Subordinate	<p>A term that is more specific than a given basic level term. Subordinate level terms:</p> <ul style="list-style-type: none"> • Are formed from a basic level term and a modifier • Are compound and/or phrases (e.g., department store, dress shirt, sports car) • Have no children (in WordNet) • Refer to specific examples and/or single objects • Describes a particular kind of a basic level object • Convey information concerning the superficial properties of the objects they refer to, such as texture and color • For example, sedan, rocking chair, sports car, dress shirt

Notes. Some of the statements in the above dictionary that refer to WordNet were adapted from Green (2006). For adjectival phrases (e.g., fall leaves), we coded based on the noun (e.g., leaves). If a modifier does not make the noun or verb more concrete/specific, we coded the term based on the verb/noun. If a compound term has two or more basic level terms, we coded it as basic. If a compound term has two or more superordinate or one superordinate and one basic level term, we coded it as superordinate. If the term represents age (e.g., child, elderly), we coded it as a basic level term. If a term is an adjective and it applies to people, we coded it as a superordinate (e.g., normal).

Author Queries

AQ1: Is chi-square correctly noted here? Aren't the number of participants to be included in the notation? Please check?

AQ2: Please provide updated publication information.

AQ3: Have I cited Table 3 in the right place?