University at Albany, State University of New York
Scholars Archive

Information Studies Faculty Scholarship

Information Studies

2008

User-generated descriptions of individual images versus labels of groups 3 of images: A comparison using basic level theory

Abebe Rorissa

University at Albany, State University of New York, arorissa@albany.edu

Follow this and additional works at: http://scholarsarchive.library.albany.edu/dis_fac_scholar

Part of the Collection Development and Management Commons

Recommended Citation

User-generated descriptions of individual images versus labels of groups of images: A comparison using basic level theory

Abebe Rorissa *

Department of Information Studies, University at Albany, State University of New York, Draper Hall, Room 113, 135 Western Avenue, Albany, NY 12222, United States

ABSTRACT

Although images are visual information sources with little or no text associated with them, users still tend to use text to describe images and formulate queries. This is because digital libraries and search engines provide mostly text query options and rely on text annotations for representation and retrieval of the semantic content of images. While the main focus of image research is on indexing and retrieval of individual images, the general topic of image browsing and indexing, and retrieval of groups of images has not been adequately investigated. Comparisons of descriptions of individual images as well as labels of groups of images supplied by users using cognitive models are scarce. This work fills this gap. Using the basic level theory as a framework, a comparison of the descriptions of individual images and labels assigned to groups of images by 180 participants in three studies found a marked difference in their level of abstraction. Results confirm assertions by previous researchers in LIS and other fields that groups of images are labeled using more superordinate level terms while individual image descriptions are mainly at the basic level. Implications for design of image browsing interfaces, taxonomies, thesauri, and similar tools are discussed.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

Although images are visual information sources with little or no text associated with them, users of image collections tend to mainly use text to describe images (e.g., *flickr*, www.flickr.com, and other social tagging services) and formulate their queries. The vast majority of web search engines (e.g., Google Image Search) that allow image searches provide search interfaces for text queries through keywords (Tjondronegoro & Spink, 2008). Keywords, subject headings, and text annotations are integral parts of both concept-based and content-based image indexing and retrieval systems, although some of the annotations by content-based image retrieval systems are achieved through automated means, albeit to a limited extent. However, for now, "successful retrieval is based largely on attaching appropriate annotations to each image and collection since automated image content analysis is still limited" (Shneiderman, Bederson, & Drucker, 2006, p. 69). The good news is that users are investing more effort in annotating not only their own images but also those that belong to other users (Shneiderman et al., 2006).

There are apparent drawbacks to concept-based image retrieval, such as the cost (in terms of time, manpower, and money) of manual annotation and indexing of images (Layne, 1994). Also, manual annotations are not always precise. At the same time, content-based image retrieval systems have failed to bridge the semantic gap (Datta, Li, & Wang, 2005; Lew, Sebe, Djeraba, & Jain, 2006; Neumann & Gegenfurtner, 2006; Smeulders, Worring, Santini, Gupta, & Jain,

* Tel.: +1 518 442 5123; fax: +1 518 442 5367.
  E-mail address: arorissa@albany.edu

2000). Words remain the predominant, if not the only means for describing the high level semantic attributes and content of images. This is perhaps partly because "the natural way to describe images is through words and phrases" (Rafferty & Hidderley, 2005, p. 180). Therefore, at least in the foreseeable future, image indexing and retrieval systems will depend on text annotations, especially for image collections with social tagging and photo sharing services such as *flickr* where users add textual tags (descriptions) to images, both their own and others'. This makes the need to understand the nature of attributes and descriptions of individual and groups of images by users even more crucial, and works such as this more relevant and significant.

The astronomical increase in the number of images in digital libraries, both personal and institutional, coupled with increased reliance on text annotations for archiving (by users and institutions), indexing (by image retrieval systems), and searching, calls for a two pronged approach to organization, indexing, and visualization of images for retrieval and browsing purposes. One approach focuses on grouping of images for browsing and exploratory search and providing browsing interfaces (through faceted, hierarchical, and other types of taxonomies). The second approach indexes both individual images and groups of images for directed search through formal queries using terms that represent image content or concepts people associate with that content. The bulk of image research focuses on the latter, indexing and retrieval of individual images, while there is minimal attention given to groupings of images except when there is talk about taxonomies and, recently, folksonomies. Among researchers who emphasize the two pronged approach are Hearst and her colleagues (Hearst, 2006a, 2006b; Yee, Swearingen, Li, & Hearst, 2003), Marchionini (2006), and Shneiderman et al. (2006). Hearst (2006a) emphasized the difficult task of labeling groups of images and, at the same time, recognized users' desire for meaningful groupings of search results. In response to these needs and desires, they have designed, built, and evaluated appropriate image browsing and retrieval systems, some of which use hierarchical and facet based groupings. In order to inform the design of such systems and other systems for indexing and retrieval of image collections (the web, social tagging services, digital libraries, etc.), a continued investigation as to what types of attributes are described through tagging and descriptions of individual and groups of images is crucial. This will ensure that image retrieval system design is user-centered and that these systems would remain effective and efficient for the retrieval of individual images and browsing groupings of those images by human users.

Knowledge of the underlying structure of user-generated tags could also inform and lead to better design of professional indexing tools such as taxonomies and thesauri. In the past, calls have been made for new ways and means for image indexing, constant evaluation and improvement of current image indexing tools and systems (Jörgensen, 1998; Layne, 1994), and for user-centered indexing (Fidel, 1994; Matusiak, 2006). Some of the reasons for not involving users in the indexing process so far, that is in a priori indexing, are that it is difficult (Hearst, 2006a), not practical (Fidel, 1994), and not cost effective. In fact, one of the limitations of concept-based image indexing and retrieval is cost of manual assignment of index terms (Layne, 1994). With social tagging becoming ever so widespread, these are now less of a limitation; at least as far as user-generated image collections such as *flickr* are concerned. However, an understanding of the nature of social tagging and the tags themselves is paramount at this early stage of social tagging in order to incorporate the users' language into indexing tools and systems. This can only be achieved through empirical research that investigates the nature of image attributes, including user-generated descriptions of both individual images and groups of images.

One of the means by which researchers sought to understand the nature of image attributes and their descriptions by users is through the study of the meanings and levels of abstraction of the terms used in those descriptions. Classifications and taxonomies of the various types of attributes of individual images and the terms used to describe them have been the subject of previous research. Most of the researchers are among the highly cited in image research and LIS literature while their taxonomies and classification frameworks are among the widely used (e.g., Enser & McGregor, 1992; Greisdorf & O'Connor, 2002b; Hollink, Schreiber, Wielinga, & Worring, 2004; Jörgensen, 1995, 1998; Laine-Hernandez & Westman, 2006; Panofsky, 1955; Shatford, 1986). There is no doubt that research into types and classes of image attributes as well as terms people use to describe and search for images is crucial to user-centered design of image organization and retrieval systems.

Terms human beings use to describe objects, convey knowledge, and express concepts have levels of generality and abstraction. Construction of hierarchical taxonomies and ontologies of domains of objects and knowledge, especially those that will aid browsing, needs to take this into account. As a hierarchical taxonomy and general framework, the basic level theory had been widely used in fields such as psychology, cognitive science, and philosophy, although its use has been limited in library and information science research and practice (Green, 2006). One of those who applied basic level theory in library and information science is Jörgensen (1995). Jörgensen (1995) analyzed attributes in terms of the three levels of abstraction, namely subordinate, basic, and superordinate, and then, along with Green (2006), called for further research in this area. Iyer (1995) also discusses the characteristics, pros, and cons of applying the basic level theory to information retrieval vocabularies. Not only are comparisons of descriptions of individual images and labels of groups of images generally rare, LIS research in general did not use the basic level theory enough to tackle relevant research problems or to make sense of these comparisons. Therefore, in this work, results of an analysis of text descriptions of individual images and labels of groups of images supplied by participants of three studies are presented. The focus is on the types of terms used as well as their level of abstraction vis-à-vis the basic level theory. The primary question that guided this research was: to what extent do user-generated textual descriptions of individual images differ from labels assigned to groups of images?

2. Image attributes and indexing

An attribute of a stimulus is a characteristic of contents of the stimulus. In the case of an image, it is a characteristic of both its visual and non-visual or semantic contents. An image attribute is "what is depicted or represented in the image" (Layne, 1994, p. 586) and it is "not limited to purely visual characteristics, but includes other cognitive, affective, or interpretive responses to the image such as those describing spatial, semantic, or emotional characteristics" (Jörgensen, 2003, p. 3) . Almost all forms of indexing and representation, in both concept-based and content-based image representation and retrieval, are based on one or more attributes of the images (Jörgensen, 1995; Rasmussen, 1997). While the number of individual attributes of a particular image could be innumerable, there are finite numbers of types as well as levels of abstraction of image attributes identified in the literature of both concept-based and content-based image retrieval. Some group image attributes into two (e.g., Jörgensen, Jaimes, Benitez, & Chang, 2001), three (e.g., Jörgensen, 1998; Markkula, Tico, Sepponen, Nirkkonen, & Sormunen, 2001; Panofsky, 1955; Shatford, 1986), and four or more broad categories (e.g., Greisdorf & O'Connor, 2002a). These finite number of types and levels of abstraction of image attributes are strong candidates to serve as frameworks to provide an interface to large image collections for browsing and exploratory search.

The earliest of these classifications and taxonomies of image attributes was proposed by Panofsky (1955) whose work and the three classes (namely pre-iconography, iconography, and iconology) of the meaning of visual arts form "the basis for much of the theoretical work that has been done on the classification of art images" (Jörgensen, 2003, p. 117). Preiconographical attributes are basic or natural characteristics, usually names of objects. Iconographical level attributes have meaning attached to them as a result of interpretation, while iconological attributes involve deeper syntheses and multiple interpretations (Jörgensen, 2003). Shatford's (1986) three groups of attributes, which are "specific of", "generic of", and "about", loosely match the three classes in Panofsky's (1955) classification.

One researcher who studied the nature of attributes of images extensively is Jörgensen (1995, 1996, 1998, 2003). In her dissertation research (Jörgensen, 1995), participants performed a description, a sorting, and a searching task. She identified three general categories of image attributes (which are further subdivided into 12 classes), namely perceptual (P), interpretive (I), and, to a lesser extent, reactive (R). According to Jörgensen (1995), perceptual attributes relate to physical characteristics of images including objects in the images, image color, and other visual elements; whereas interpretive attributes are in the eyes of the viewer and require more than just perceiving. In an attempt to further classify image attributes, Jörgensen and her colleagues (Jörgensen et al., 2001) provide a hierarchical conceptual framework for categories of visual attributes in the form of a "pyramid". They divide image attributes into 10 levels and two general categories. The first category, syntax, consists of four mainly perceptual (or basic) attributes, while the second category, semantics, is made up of six conceptual (higher level or semantic) attributes.

A number of these investigators have developed these and other taxonomies through user reactions to images and image description tasks (Greisdorf & O'Connor, 2002a; Greisdorf & O'Connor, 2002b; Jörgensen, 1995, 1996, 1998; O'Connor, O'Connor, & Abbas, 1999; Turner, 1994, 1995) as well as through solicitation of queries from image users (Chen, 2001a, 2001b; Choi & Rasmussen, 2002, 2003; Goodrum & Spink, 2001; Hastings, 1995). Rosch (1975) laid the foundations for the basic level theory. Rosch and others (Rosch, 1975; Rosch & Mervis, 1975; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976) have also studied categories of attributes of objects vis-à-vis the basic level theory. A recent work by Laine-Hernandez and Westman (2006) reviewed most of the above-mentioned classifications (except basic level theory) and suggested a framework that integrates some of their common classes and added a few more classes.

So far, the discussion in this paper focused on the types, levels of abstraction, and classifications of image attributes as well as the terms used to describe those attributes. One of the major problems in information storage, organization, access, and retrieval is related to representation (including indexing) of the documents. This research, because it is a good match for the type of data analyzed (text descriptions of individual images and labels of groups of images), adopted the concept-based definition of image indexing which is the representation of images through textual descriptions of their contents and what they are *of* or *about* (Shatford, 1986). While an image could be of and about something, its *ofness* is mainly concrete and objective (e.g., a picture of a cowboy) and its *aboutness* is more abstract and subjective (Layne, 1994; Shatford, 1986). Human viewers of a picture attach meaning to what the picture represents above and beyond its *ofness*. The visual nature of images and the fact that an image is not just of something and that an image is truly worth a thousand words compounds the problem of their representation and indexing even more.

Blair (1990) argues that representation is "primarily a problem of language and meaning" (p. 1). It is even more so in the case of concept-based image indexing where the documents being represented are visual images, rather than structured text documents, and the task of representation involves description of image attributes using text while the task of retrieval involves matching those text descriptions with users' query terms. O'Connor et al. (1999) argue that representation of images using terms generated through users' reactions might address this problem. As pointed out earlier, one of the limitations of concept-based image indexing and retrieval is the cost of manual assignment of index terms by human indexers, especially for large image collections. In spite of this, textual descriptions of image attributes remain one of the popular methods of their indexing (Enser, 2000; Jörgensen, 1995, 1996, 1998, 2003). Most image users formulate their queries using words although this is neither always the necessary nor the only way (O'Connor et al., 1999).

In an ideal situation, automatic extraction, annotation, and indexing of images based on low-level as well as high-level and semantic features would be possible. However, most of the current content-based image retrieval (CBIR) methods could

only enable the extraction of low-level features such as color, shape, texture, luminosity, and edge. Even though content-based image retrieval (CBIR) mechanisms offer cost effective alternatives, they still lag far behind in bridging the semantic gap (Datta et al., 2005; Lew et al., 2006; Neumann & Gegenfurtner, 2006; Smeulders et al., 2000). There is still continued dependence on concept-based image indexing and retrieval, mostly with the help of human indexers. Therefore, it is increasingly important for image retrieval systems to take advantage of the recent phenomenon of social tagging (e.g., *flickr*) and for system designers to adopt more user-centered approaches. To address this, continued investigations and full understanding of the nature of image descriptions by users are required and this work is a step in that direction.

In summary, concept-based image indexing methods based on textual descriptions of image content remain the dominant representation mechanisms. Multimedia searching on the World Wide Web is mainly based on keywords. Therefore, any image indexing tool and system should encompass the full range of attributes, including those that are perceived and concrete or otherwise. Most of the various classifications and taxonomies of image attributes discussed above were mainly developed in LIS through grounded theory approaches and some of them are non-hierarchical. This work focuses on a hierarchical and widely used (across several disciplines) framework. The basic level theory and its three levels of abstraction (subordinate, basic, and superordinate), are used here to compare descriptions of individual images and labels of groups of images. The choice of basic level theory was justified because existing taxonomies used by previous LIS researchers for similar investigations were built inductively and were constructed with a specific purpose in mind. On top of its wider application in a number of fields, basic level theory provides a neutral, hierarchical, general, and deductive framework. It also brings a new approach to LIS research on this topic.

3. Basic level theory

According to the basic level theory, objects, human concepts, and natural categories have a hierarchical structure with three levels of abstraction and generality, namely subordinate, basic, and superordinate (Rosch & Mervis, 1975; Rosch et al., 1976). Among these three levels, people, in general, prefer to use the basic level when they are thinking about the world around them (Murphy & Wisniewski, 1989; Rosch et al., 1976). The basic level is preferred especially when a task involves identification of single and isolated objects rather than labeling groups of objects (Murphy & Wisniewski, 1989). For an in depth description of the three levels, please see Appendix A.

There are a number of characteristics that set the three levels of abstraction of concepts apart. Whereas basic level concepts are generally more similar to each other, superordinate concepts are not. Basic level concepts are more similar because their instances share more common attributes than instances of superordinate concepts do (Rosch et al., 1976). It has been confirmed that the more common features two stimuli have, the more similar they are (Tversky, 1977). For instance, various types of chairs have more in common than various types of furniture (e.g., chair, table, and bed). Subordinate concepts have an even greater similarity among their instances than among instances of basic or superordinate concepts (Rosch et al., 1976). For instance, several types of rockers (subordinate term) are more similar than various types of tables (basic term) and furniture. Stated differently, basic and subordinate level categories have a higher degree of within-category similarity than members of superordinate categories (Markman & Wisniewski, 1997). What is more, while both basic level and superordinate categories have a low degree of between-category similarity, subordinate categories have a high degree of between-category similarity, especially if they belong to the same or neighboring basic level categories (Markman & Wisniewski, 1997; Murphy & Brownell, 1985). This makes members of a superordinate category less confusable, and more likely to be correctly classified at their respective level (Markman & Wisniewski, 1997), compared to members of a subordinate or a basic level category.

Children have been found to learn the more concrete terms before abstract ones (Horton & Markman, 1980; Mervis & Crisafi, 1982) and nouns before verbs (Gentner, 1982). In other words, they learn basic level concepts before superordinates (Horton & Markman, 1980; Mervis & Crisafi, 1982) and, perhaps this is due to the fact that in social situations, within which children learn categories, basic level names are short and appear in high frequency (Markman & Wisniewski, 1997). As adults, in their discourses, they continue to use more basic level terms than they do superordinates (Wisniewski & Murphy, 1989). There are two main reasons for this. First, in a natural language, basic level terms appear with higher frequency compared to superordinate or subordinate terms (Markman & Wisniewski, 1997). Second, for novices, the basic level is the most useful level of abstraction when referring to objects (Rosch et al., 1976; Tanaka & Taylor, 1991). On top of this, children and adults differ in their view of superordinate categories. To children, superordinate categories are collections with a part-whole structure (e.g., a father is part of a family, or an arm is part of a body) whereas to adults, they are classes whose members have an ISA relation (e.g., a car is a, ISA, kind of vehicle) (Wisniewski & Murphy, 1989).

In summary, both children and adults, particularly novices to a domain of knowledge, prefer basic level terms to refer to single and isolated objects. In written text, superordinate terms are used more to refer to groups of objects (Wisniewski & Murphy, 1989). This was confirmed through analysis of human subjects' categorization of pictures of objects isolated or in scenes (Murphy & Wisniewski, 1989). Clearly, there is a distinction in the level of abstraction, at least with respect to the basic level theory, between terms used to describe individual and groups of objects. This is the distinction this work sought to ascertain. More specifically, its goal is to see if there is a fundamental difference in the level of abstraction of descriptions of individual images and labels of groups of images vis-à-vis the basic level theory. Specifically, this work sought to determine if people use more basic level terms to describe individual images and use superordinate terms to label groupings of

234 them than terms at the other two levels. In order to achieve this, data from 180 participants who either provided text
235 descriptions or group labels for 130 individual and groups of images were analyzed through content analysis.

4. Methods

4.1. Studies

238   Three separate studies were conducted between March 2003 and November 2006. For Studies 1 and 2, data were col-
239 lected using the free sorting method (Coxon, 1999) where participants were asked to sort two separate samples of 50 images
240 without any constraints on time and number of categories and assign labels to groups of images. Study 3 was different from
241 Studies 1 and 2 in that data were collected, as part of a larger study, through an image description task involving 30 images
242 performed by 75 participants.

4.2. Materials

244   For Studies 1 and 2, two separate random samples of 50 color digital images were selected from disc number 6 (which
245 contains all photographs in the "people" category) of the *Hemera Photo Objects Volume I*, a stock photo collection of over
246 50,000 images (http://www.hemera.com), printed on a 4×5 in. (10.2×12.7 cm) card. For Study 3, a random sample of
247 30 images, 375×250 pixels in dimension, was selected from a collection of color photographs included in O'Connor and
248 Wyatt (2004).

4.3. Participants

250   Participants in the three studies were 180 volunteer graduate students at two major US universities (one in the Southwest
251 and the other in the Northeast). Thirty of those participated in Study 1 (16 female and 14 male), 75 in Study 2 (59 female and
252 16 male), and the remaining 75 in Study 3 (56 female and 19 male). All participants were between the ages of 21 and 60
253 years old.

4.4. Procedures

4.4.1. Data collection

256   Participants of Studies 1 and 2 were instructed to: (a) inspect the images first, (b) sort them into as many or as few groups
257 (or categories) as they wished using their own general criteria for similarity, and (c) label each group. The cards were reshuf-
258 fled before being given to the next participant. A different data collection procedure was followed in Study 3. An e-mail mes-
259 sage introducing the study and containing the URL for an image description task was sent to each participant during the
260 summer and fall 2004 semesters. Two weeks later, a follow-up e-mail message was sent to those participants who did
261 not complete the task. Each participant read the instructions and was presented with a web-based form with a random im-
262 age displayed at the top. No two participants were shown the 30 images in the same order. To ensure internal validity, a time
263 constraint of 90 seconds was enforced within which participants were asked to type as many terms describing the images
264 and their attributes as possible. To familiarize participants with the task, two images (not in the sample) were included at the
265 beginning of the task. Participants were not constrained in terms of the number of words they could use as group labels or
266 descriptions.

4.4.2. Data analysis

268   The main data analysis technique used was content analysis while the χ^2 test of independence and one-way analysis of
269 variance (one-way ANOVA) were used mainly to test the difference between the number of labels of groups of images and
270 terms used to describe individual image features at the three levels of abstraction (subordinate, basic, and superordinate). A
271 four-step content analysis procedure was followed: (a) the creation and testing of a coding scheme that involves the defini-
272 tion of recording units and categories, (b) assessment of the accuracy of coding, (c) revision of coding rules, and (d) coding
273 the entire text (Weber, 1990). Terms used by participants to describe individual image features and labels of groups of
274 images (including single words and "word senses") were the recording units (basic unit of text coded or categorized). Here,
275 the word "term" is used to mean a concept that is referenced using one or more words. Phrases constituting a semantic unit
276 such as idioms ("taken for granted"), proper nouns ("the Empire State Building"), as well as other phrases (e.g., "sweet
277 home", "life routine") (Weber, 1990) are examples of "word senses". These "word senses" and terms were coded as single
278 terms.
279   Due to lack of an established coding dictionary, a coding scheme (see Appendix A) was specifically developed for the work
280 by Rorissa and Iyer (2008) based on the definition, explanations, and examples given for the three levels of abstraction
281 (superordinate, basic, and subordinate) in previous literature. It was modified and enhanced mostly through consultation
282 of the book by Iyer (1995) and through several discussions between the author, a colleague of the author, and a graduate
283 student. Labels of groups of images from Studies 1 and 2 were already coded (please see Rorissa & Iyer, 2008) while the

284 author and the graduate assistant coded all the descriptions of individual images supplied by the participants of Study 3
285 using the coding dictionary until they reached 100% agreement. A third coder, a colleague of the author who was blind to
286 the purpose of the study, coded a random sample (20%) of the descriptions of individual images (Study 3). Percent agreement
287 and Cohen's (1960) κ, two popular measures of intercoder agreement, were used to measure the reliability of coding be-
288 tween the author and the third coder. The computed values of the percent agreement and Cohen's (1960) κ were 0.91
289 and 0.79 for Study 1, 0.89 and 0.80 for Study 2, and 0.92 and 0.80 for Study 3, respectively. These values are within the rec-
290 ommended range for good levels of intercoder agreement (Neuendorf, 2002).

291 **5. Results**

292 *5.1. Description of the three studies*

293   In total, participants of Studies 1 and 2 supplied 899 group labels (240 in Study 1; 659 in Study 2). The total number of
294 terms supplied by participants of Study 3 is 15,301, including duplicates. Out of these, 4110 (26.86%) were unique terms. The
295 minimum, maximum, median, mode, and mean number of terms supplied by all participants of Study 3 per image were 352,
296 **Q1** 652, 514, 550, and 510, respectively. Table 1 is a summary of the three studies (see Table 2).

297 *5.2. Level of abstraction of labels of groups of images and descriptions of individual images*

298   Out of the 899 labels of groups of images supplied by 105 participants of Studies 1 and 2, 39 (4.34%) were coded at the
299 subordinate level (3.33% and 4.7% for Studies 1 and 2, respectively), 342 (38.04%) were coded at the basic level (42.5% and
300 36.42% for Studies 1 and 2, respectively), and 518 (57.62%) were coded at the superordinate level (54.17% and 58.88% for
301 Studies 1 and 2, respectively). Unlike group labels, descriptions of individual images required some cleaning. Words and
302 phrases used in the descriptions of individual images by 75 participants of Study 3 were screened first in order to ensure
303 that only proper words and phrases, not stop words, were considered for coding. After the screening, a total of 15,301 proper
304 terms were coded where terms supplied by more than one participant were considered separate and coded accordingly. Of

Table 1
Summary of the three studies

Study details	Study		
	1	2	3
Number of images	50	50	30
Type of images	People	People	Miscellaneous
Method of data collection (task)	Free sorting (group labeling)	Free sorting (group labeling)	Image description
Number of participants	30	75	75
Total number of groups or terms supplied	240	659	15 301[a]
Min. number of groups or terms supplied	3	2	1,[b] 81,[c] 352[d]
Max. number of groups or terms supplied	7	24	24,[b] 347,[c] 652[d]
Mean number of groups or terms supplied	8	8.79	6.8,[b] 204,[c] 510[d]
Median number of groups or terms supplied	7	8	6,[b] 199,[c] 514[d]
Mode number of groups or terms supplied	7	9	5,[b] 165,[c] 550[d]
SD of number of groups or terms supplied	3.3	4.1	3.07,[b] 66.34,[c] 81.89[d]

[a] Terms were counted as separate if supplied by two or more participants.
[b] Per image per participant.
[c] Per participant.
[d] Per image.

Table 2
Examples of labels of groups of images and descriptions of individual images supplied by participants

Subordinate	Basic	Superordinate
Apartment building	Building	Structure
Attack dog	Dog	Animal
Dallas	City	Urban area
Halloween Costume	Costume	Clothes
JFK	Man	Person
Right hand	Hand	Body
Satchel	Bag	Container
Smiling face	Happiness	Emotion
Tree trunk	Tree	Plant
Video camera	Camera	Equipment
Volkswagen	Cars	Vehicle

Table 3
The mean and standard deviation of number of subordinate, basic, and superordinate group labels (Studies 1 and 2) and descriptions (Study 3)

Study	n	Subordinate		Basic		Superordinate	
		Mean	SD	Mean	SD	Mean	SD
1	30	0.27	0.69	3.40	2.37	4.33	2.29
2	75	0.41	0.72	3.20	2.76	5.17	2.38
3	75	56.67	30.04	129.12	51.89	18.23	13.43

the 15,301 words and phrases, 4250 (27.78%), 9684 (63.29%), 1367 (8.93%) were at the subordinate, basic, and superordinate levels, respectively (see Table 4).

On the face of it, based on the above frequencies and the mean number of subordinate, basic, and superordinate group labels and descriptions (Table 3), there seems to be an apparent difference in the level of abstraction (vis-à-vis the three levels of subordinate, basic, and superordinate) between the types of terms people use to label groups of images and those they use to describe individual images. For instance, participants of the group labeling tasks (Studies 1 and 2) supplied about one and half times as many superordinate labels as the number of basic and subordinate labels combined. On the other hand, participants of Study 3 used more than seven times as many basic level terms than superordinate terms to describe individual images. However, further analysis must be done to determine whether or not these apparent differences are also statistically significant. The next Section 5.3 is devoted to this. Because two studies (Studies 1 and 2) were conducted to solicit labels of groups of images, before conducting further analysis to answer the research question this work is attempting to answer, combining the data from the two studies is essential. This will create two distinct levels of the type of task (labeling groups of images, Studies 1 and 2, and describing individual images, Studies 3). In order to justify this, three separate tests were conducted.

First, the homogeneity of their variances, that is, whether or not samples from the two studies (Studies 1 and 2) have the same variances in terms of the number of labels of groups of images supplied by participants coded at the three levels (subordinate, basic, and superordinate) was tested using the *F*-test. At all three levels, *F*-values ($F(74, 29) = 0.75$, 1.13, and 0.75, with $p > 0.05$ for subordinate, basic, and superordinate levels, respectively) were statistically non-significant. This is proof that the number of corresponding subordinate, basic, and superordinate group labels from the two studies satisfy the homogeneity of variance criterion and assumption of equal variances is justified.

Second, to complement the homogeneity of variance test, an independent groups *t*-test (with equal variances) was conducted in order to see if the mean number of subordinate, basic, and superordinate level group labels from the two studies are comparable. The *t*-test indicated a non-significant difference between the corresponding mean number of group labels at the subordinate ($t(103) = 0.96$, $p > 0.05$), basic ($t(103) = 0.35$, $p > 0.05$), and superordinate ($t(103) = 1.65$, $p > 0.05$) levels for the two studies. Hence, not only that the two samples have equal variances, the mean number of subordinate, basic, and superordinate level group labels from the two studies have very little or no difference (pairwise).

Third, a chi-square (χ^2) test of independence was conducted to determine whether the numbers of subordinate, basic, and superordinate labels of groups of images from the two studies are independent of the samples of images and participants. A non-significant chi-square ($\chi^2 (2, N = 899) = 3.150$, $p > 0.05$) confirmed the fact that the levels of abstraction of group labels are independent of the samples of images and participants. Hence, the two samples from Studies 1 and 2 were considered similar in their structure vis-à-vis the three levels. Based on results from the homogeneity of variances test, the independent groups *t*-test, and the χ^2 test, the two samples from Studies 1 and 2 (labels of groups of images) were combined and will be treated as a single sample (group labels) in subsequent analyses (except in one-way analysis of variance). Table 4 presents the frequencies and percentages of the group labels (Studies 1 and 2 combined) and descriptions of individual images (Study 3).

5.3. Difference between labels of groups of images and descriptions of individual images

To reiterate, the research question guiding this work is: to what extent do user-generated textual descriptions of individual images differ from labels assigned to groups of images? Within the framework of the basic level theory and the

Table 4
Total frequency and percentage of labels of groups of images and descriptions of individual images that are at the subordinate, basic, and superordinate levels

Level of abstraction	Group labels		Descriptions		Total	
	Frequency	%	Frequency	%	Frequency	%
Subordinate	39	4.34	4250	27.78	4289	26.48
Basic	342	38.04	9684	63.29	10026	61.89
Superordinate	518	57.62	1367	8.93	1885	11.64
Total	899	100.00	15,301	100.00	16,200	100.00

343 three levels of abstraction (subordinate, basic, and superordinate), it could be restated as: do labels of groups of images
344 and descriptions of individual images supplied by people have different underlying structures with respect to the three
345 levels? If so, which level is used more for each task (labeling groups of images and describing individual images)? To
346 answer this, a χ^2 test of independence and a one-way (single factor – level of abstraction) analysis of variance (one-
347 way ANOVA) were used. The χ^2 test of independence was based on the frequencies in Table 4 and it was used mainly
348 to see if the group labels and descriptions of individual images have different underlying structures with respect
349 to the three levels (subordinate, basic, and superordinate). A statistically significant chi-square value (χ^2 (2,
350 $N = 16,200) = 1993.33$, $p = 0.000$) was obtained. This was interpreted as evidence that the type of task (labeling groups
351 of images and describing individual images) is not independent of the level of abstraction. That is, the underlying struc-
352 ture of labels of groups of images and descriptions of individual images are different vis-à-vis the basic level theory and
353 the three levels.

354 These data show that there is a difference in the underlying structure of labels of groups of images and descriptions of
355 individual images and, as a result, addressed the first part of the question. However, the χ^2 test of independence does not
356 help answer the question: which level is used more for each task (labeling groups of images and describing individual
357 images)? To answer this second part of the question, a one-way (single factor – level of abstraction) analysis of variance
358 was necessary. First, data from each of the three studies were subjected to an initial screening to determine whether they
359 satisfy the homogeneity of group variances criterion. For each study, the groups were the three levels of abstraction (subor-
360 dinate, basic, and superordinate). The first screening resulted in unequal variances for data from all three studies. A correc-
361 tive measure, transformation of the data (taking the square root of the number of subordinate, basic, and superordinate level
362 terms), was taken in order to satisfy the homogeneity of group variances criterion. Levene's test for homogeneity of group
363 variances confirmed that the transformation yielded statistically non-significant differences between all possible pairs of
364 variances of the number of subordinate, basic, and superordinate level group labels and descriptions of individual images
365 for all three studies.

366 To determine if terms at one of the three levels of abstraction were used more for each of the two tasks (labeling groups of
367 images and describing individual images) than at the other two levels, One-way ANOVA was conducted on the single factor
368 of level of abstraction with the three levels of subordinate, basic, and superordinate. Tukey's honestly significant difference
369 (HSD) post hoc test was used for multiple comparisons. Both tests were conducted on the transformed data from all three
370 studies and results from analyses of the transformed data are presented below.

371 For all three studies, results from the one-way ANOVA revealed that there was a significant overall difference between
372 the levels of abstraction. That is, the number of subordinate, basic, and superordinate level terms used as labels of groups of
373 images and descriptions of individual images were significantly different for both the group labeling task (Study 1: $F(2,
374 87) = 59.3$, $p = 0.000$; Study 2: $F(2, 222) = 128.3$, $p = 0.000$) and description task (Study 3: $F(2, 222) = 225.6$, $p = 0.000$). Tu-
375 key's HSD post hoc test conducted to compare all three pairs of means yielded mixed results for the group labeling tasks
376 (Studies 1 and 2). For both studies, even though all other pairs of the mean number of subordinate, basic, and superordinate
377 level terms were significantly different from each other, the difference between the mean number of superordinate and ba-
378 sic level group labels were not significant for Study 1. This could be attributed to the small sample size in Study 1. Partic-
379 ipants of Study 2 supplied more group labels at the superordinate level ($M = 5.17$) than at the basic ($M = 3.20$) and
380 subordinate ($M = 0.41$) levels. Participants of Study 1 also supplied more basic level group labels ($M = 3.40$) than subordinate
381 ones ($M = 0.27$). On the other hand, comparisons of the number of subordinate, basic, and superordinate terms supplied by
382 participants of the other task, description of individual images (Study 3), indicated that individual images are described
383 more at the basic level ($M = 129.12$) than at the subordinate ($M = 56.67$) and superordinate ($M = 18.23$) levels. This is a clear
384 indication that labels of groups of images and descriptions of individual images have markedly different underlying struc-
385 tures vis-à-vis the three levels. What is more, superordinate level terms are preferred to those at the subordinate and basic
386 levels for labeling groups of images while basic level terms are used more than those at the other two levels to describe
387 individual images.

6. Discussion and concluding remarks

6.1. Summary of results

390 Participants of three studies involving 130 images completed two sorting (and labeling of groups of images) tasks and a
391 description of individual images task. The 180 participants produced a total of 899 group labels and 15,301 description
392 terms. Using the basic level theory and three levels of abstraction (subordinate, basic, and superordinate) as framework, this
393 work sought to determine the extent of the difference between user-generated labels of groups of images and descriptions of
394 individual images. In addition, it attempted to test previous researchers' assertions that people use more superordinate level
395 terms to label groups of things whereas they mainly use basic level terms to refer to individual objects and concepts. It has
396 been argued that people predominantly use basic level terms in free-naming tasks (Rosch et al., 1976). In their discourses,
397 when adults are thinking about the world around them, their preferred level of categorization is the basic level and they use
398 more basic level terms than superordinate ones (Murphy & Wisniewski, 1989; Rosch et al., 1976; Wisniewski & Murphy,
399 1989). However, Wisniewski and Murphy (1989) assert that superordinate terms are used more to refer to groups of objects.

400 With the help of a combination of the χ^2 test of independence, one way (single factor - level of abstraction) ANOVA, and Tu-
401 key's HSD post hoc test, this work provided support for these assertions. They hold true also when people label groups of
402 images and describe individual images.
403   Through analyses of data from two tasks, sorting and labeling of groups of images (Studies 1 and 2) and description of
404 individual images (Study 3), this work has shown that:

405 • User-supplied labels of groups of images and descriptions of individual images have different underlying structures. Also,
406   the type of task (labeling groups of images and describing individual images) is not independent of the level of abstraction
407   (vis-à-vis the three levels of subordinate, basic, and superordinate).
408 • For all three studies, the number of subordinate, basic, and superordinate level terms have statistically significant
409   differences.
410 • While, for the most part, more superordinate than basic and subordinate terms were supplied as group labels, descriptions
411   of individual images were mainly at the basic level as opposed to the other two levels.

413   These results are in agreement with previous research in fields other than LIS, and they also complement the work of
414 Jörgensen (1995) who, after conducting similar studies that involved image sorting tasks, arrived at the conclusion that
415 image groupings by people were at the superordinate level while subgroups were at the basic level. Rosch et al. (1976)
416 have also observed that sorting tasks that involved subjects to group things together lead to categorization at the super-
417 ordinate level. This could be due to the fact that grouping entities establishes a global order and provides an overall struc-
418 ture while description of the entities concentrates on the individual entity (Soergel, 1985). In conclusion, this research is
419 consistent with the views of many researchers (Greisdorf & O'Connor, 2002a; Hollink et al., 2004; Jörgensen, 1998; Laine-
420 Hernandez & Westman, 2006) concerning labeling groups of images and description of individual images. When describing
421 images, people made references to objects (mainly named at the basic level) in images much more than other types of
422 attributes.

6.2. General implications of results

424   The existence of a marked difference, in terms of level of abstraction, between user-supplied labels of groups of images
425 and descriptions of individual images shown in this work within the framework of the basic level theory and the three levels
426 of abstraction will have a number of implications. Chief among these is that basic level theory could provide a possible solu-
427 tion to the problem and concern raised by some (e.g., Hollink et al., 2004) regarding lack of a common classification frame-
428 work and taxonomy to study the nature of image attributes and compare results across several studies. The basic level theory
429 provides the theoretical grounding and explanation to some of the views and assertions of the above researchers. What is
430 more, basic level theory is closest to a universal hierarchical taxonomy of concepts and objects because it was found to
431 be reliable across cultures (Lakoff, 1985). For instance, basic level theory explains, partially, why people make more refer-
432 ences to objects in images than to other attributes and why, as results of the three studies above showed, groups of objects
433 are labeled using superordinate level terms. Basic level terms are used by people more when the task involves identification
434 of single and isolated objects rather than labeling groups of them (Murphy & Wisniewski, 1989). Moreover, children learn
435 basic level concepts first (Horton & Markman, 1980; Mervis & Crisafi, 1982). Because basic level names are short and appear
436 in high frequency in natural languages (Markman & Wisniewski, 1997), adults continue to use more basic level terms (Wis-
437 niewski & Murphy, 1989). Besides, for novices, the basic level is the most useful level of abstraction when referring to objects
438 (Rosch et al., 1976; Tanaka & Taylor, 1991).
439   The significance of the basic level theory is not limited to just naming of objects. Search behaviors of image users exhibit
440 some of the hierarchical nature of the basic level theory. For instance, Goodrum and Spink (2001) showed that users use, for
441 the most part, search terms such as "girls" (a basic level term) and modify them to "pretty girls" (a subordinate level term).
442 Moreover, most natural categories and some of the classificatory structures (e.g., taxonomies and classification systems) are
443 hierarchical and they are organized from general to specific. For instance, some indexing tools such as the Art & Architecture
444 Thesaurus (AAT) contain facets that are conceptually organized hierarchically from abstract to concrete. However, this orga-
445 nization may not be the same as the organization of natural categories. Natural categories are organized in such a way that
446 the cognitively basic level is almost always 'in the middle' of the hierarchy (Lakoff, 1985). This makes the basic level theory
447 unique because "the classical theory of categories gives no special importance to categories in the middle of a taxonomic
448 hierarchy" (Lakoff, 1985, p. 46).

6.3. Implications of results for browsing and indexing of image collections

450   The results in this work and the use of basic level theory as a framework have other implications, too, especially for the
451 design of browsing interfaces for large image collections and design of taxonomies, classificatory tools, thesauri, and other
452 similar indexing tools. The current practice of design of these interfaces and tools is uncoordinated and no common frame-
453 works exist. A common framework would help the design efforts tremendously because some groups of users, usually pro-
454 fessionals (journalists, architectural designers, art directors, etc.), want to use browsing and textual category labels to search

for images (Armitage & Enser, 1997; Markkula & Sormunen, 2000) and an appropriately designed browsing interface could be crucial to their success in finding the needed image.

For these and other reasons, it is essential to find a common framework. To provide access to large image collections and successfully support such users, a category based approach for browsing and exploratory search could be the way to go (Hearst, 2006a; Shatford, 1986; Yee et al., 2003). One of these approaches is an appropriate hierarchical browsing interface. Several reasons and scenarios could be cited for using the basic level theory as a framework for designing these interfaces and other image indexing tools. For instance, the three levels of abstraction of concepts (subordinate, basic, and superordinate) provide a hierarchical taxonomy in order to build hierarchical classificatory structures which are often appropriate for building browsing interfaces and systems. Users have been shown to prefer hierarchies rather than disorderly groupings because a hierarchical grouping allows flexible ways to access a large document collection (Hearst, 2006a). With social tagging becoming popular and users willing to provide annotations and tags for images, imposing an automatic hierarchical (or faceted) structure onto image collections such as those in *flickr* is possible because user tags are readily available a priori (Hearst, 2006a). Availability of tags also solves one of the problems associated with clustering and groupings of documents, the difficulty of manually labeling groups of documents. Whether the browsing interface is hierarchical or facet based, appropriate labels for groupings of documents are crucial.

The need for better tools to support browsing is real. For instance, half of the searches conducted by users of a video retrieval system (Open Video Digital Library) begin with browsing (Marchionini, 2006). However, because there are at least two types of users of an image collection, or any information system for that matter (those that conduct directed searches and those that prefer browsing), an image retrieval system should provide suitable interfaces and a combination of both searching and browsing services. There are some who view search and browsing as two ends of the exploratory search continuum (Cutrell & Dumais, 2006). We concur with Marchionini (2006) that retrieval systems should support both exploratory search or browsing and lookup because there are perhaps as many users who have no clear idea of what they are looking for as users with specific search requests. Users without concrete requests are likely to need an appropriate browsing interface while those who have a good idea of what they want could use advanced search features of retrieval systems. Both types of users could benefit from browsing interfaces, especially ones with hierarchical taxonomies, at the searching (query formulation) and browsing (search results review) stages

An added advantage of using the three levels of abstraction (subordinate, basic, and superordinate) as a framework for designing either a hierarchical or faceted taxonomy and browsing interface such as Flamenco (Yee et al., 2003) is that it could reduce the number of "drill-down" levels (unlike tools such as WordNet®) while still providing the user with the basic classificatory structure for browsing and exploratory search. An alternative approach in terms of the "drill-down" for image collections where keyword searching (or directed searching) using object names is predominant could start at the basic (middle) level and provide the user with a view where both the superordinate and subordinate levels are visible allowing the user to broaden or narrow his/her search, respectively. This could solve the problem with some browsing interfaces where "the user is usually not given an overview of the category space other than a top-level view of the most general labels" (Hearst & Karadi, 1997, p. 301). Even if a user prefers or starts with a directed search, using his/her query terms, the browsing interface or system could be designed such that it would determine the level of abstraction of individual terms or a combination of the query terms and start at the appropriate level for browsing. In other words, a taxonomy or interface based on the three levels of abstraction (subordinate, basic, and superordinate) should not be limited to providing access to image collections through browsing. Results of a directed search could also be displayed using the taxonomy in order to assist the user in either narrowing down and/or broadening his/her search based on the category within which the bulk of relevant items fall.

Another area where the basic level theory could serve as a framework in the design of browsing interfaces, taxonomies, and indexing tools is interoperability. The design of a general, yet simple and less complex, taxonomy or classificatory structure, with three levels in depth (corresponding to the levels of subordinate, basic, and superordinate) could play a role in polling results together using a single query from several collections and/or systems. Such a taxonomy or classificatory structure would facilitate the interoperability of indexing tools and controlled vocabularies such as thesauri (e.g., AAT) and subject headings (e.g., LCSH, MeSH) as well as cross-searching and browsing across distributed digital libraries, including image collections. Entries in an indexing tool and/or controlled vocabulary could be mapped to the three levels of subordinate, basic, and superordinate before they are matched with entries in the indexing tools and controlled vocabularies of other collections and/or systems. This approach to interoperability between indexing tools and controlled vocabularies, thereby facilitating the interoperability and standardization of browsing interfaces and taxonomies, complements efforts being made by others (e.g., see Nicholson, Dawson, & Shiri, 2006, and the High-Level Thesaurus project at: http://hilt.cdlr.strath.ac.uk/). Such a framework also has the potential to partially solve another problem that arises from matching individual entries and index terms in these indexing tools, controlled vocabularies, and image collections across various domains. Because most of these tools and collections are specialized and were built to serve a specific area/domain or specific types of users, term level matching is rendered ineffective, if not inappropriate.

An additional feature of a taxonomy based on the basic level theory and the three levels of subordinate, basic, and superordinate is that it fits the general trend in the number of levels in depth of existing taxonomies for browsing and the study of image attributes. For instance, the Picture Australia network (www.pictureaustralia.org) has a two-level sub-

ject-based taxonomy, the taxonomy for the British Library's Collect Britain network of historical resources (www.collectbritain.org) and the Heritage Colorado network (www.aclin.org) have three levels, while that of the Online Archive of California (www.oac.cdlib.org) has up to four levels of depth (Chaudhry & Jiun, 2005). Limiting the number of levels of depth in a taxonomy and browsing interface to three has previous research support as well. For instance, most researchers have been using taxonomies with three levels to study the nature of image attributes. The three levels of subordinate, basic, and superordinate make the basic level theory comparable (though not equivalent) to three levels of previous frameworks and taxonomies of the meaning of images and art such as Jörgensen's (1995) (perceptual, interpretive, and reactive), Shatford's (1986) ("specific of", "generic of", and "about"), and Panofsky's (1955) (pre-iconography, iconography, and iconology), making it a logical candidate for a general framework to study the attributes, meaning, and subjects of images.

6.4. Summary and recommendations for future work

In summary, results and the subsequent discussions presented here as well as previous research in LIS and other fields suggest that the basic level theory and the three levels of subordinate, basic, and superordinate have a number of implications and possible applications. They may provide possible solutions to a number of problems ranging from browsing interfaces to standardization of taxonomies and labeling of categories. While a taxonomy based on the three levels would fit most current taxonomies in terms of the number of levels of depth, results of the three studies reported in this work have shown that for browsing purposes labels for individual images and objects in them could be at the basic level while labels for groups of images could be at the superordinate level. Furthermore, taxonomy and thesaurus developers could use the three levels to define categories and for sorting or grouping concepts and terms as part of their procedures. For instance, Step 5 of the taxonomy development procedure used by Chaudhry and Jiun (2005) could be augmented by the use of the basic level theory. A thesaurus developer could use the three levels to determine relationships between terms and concepts, more so for the hierarchical relationships – broader and narrower term relationships – and the associative relationships, albeit to a lesser degree. The basic level theory conforms to the general model of thesaurus construction, at least as far as hierarchical (broader and narrower term) and, to some extent, associative relationships between concepts are concerned. It also conforms to the "class inclusion" type of hierarchical relationship where a "narrower concept has all the characteristics of the broader concept and, in addition, at least one further characteristic. Thus, we can always say, 'B (the narrower concept) is an A (the broader concept) that has the characteristic C' " (Soergel, 1985, p. 282).

This work is among very few studies that deal with the topic of the extent of the difference between labels of groups and descriptions of individual images. It attempted to not only use the basic level theory to determine this difference but also show the implications and applications of the basic level theory to image indexing and retrieval. It should be noted that even though the basic level theory was used as a framework in this work because of its strengths, it has some weaknesses. For instance, the three levels may not hold in all cases across languages. What is more, domain knowledge may affect the level of concepts people use to either group or name objects and concepts that are superordinates for experts could be considered basic for novices (Tanaka & Taylor, 1991; Wisniewski & Murphy, 1989).

It goes without saying that results reported here need to be verified using a number of different samples of images and participants in order for them to be used as a basis for either design of image indexing and retrieval tools or the conduct of future research. Some of the studies reported here were conducted before *flickr* data were made widely and publicly available to researchers. This is an exploratory research and the impact the types of images selected for the three studies could have had on the results needs to be checked by future research. The samples of images of people from a stock photo collection could be a possible limitation. The author plans to use *flickr* as a source of data and materials (images, tags, descriptions, etc.) in future studies regarding user-generated descriptions of images and recommends that other researchers also do the same in order to have comparable results and to arrive at useful and general conclusions. A recent work by Stvilia and Jörgensen (2007) is an example of a step in this direction. What is more, this will also help determine whether the choice of images in the three studies influenced the results.

7. Uncited reference

Q2   Fidel (1997).

Acknowledgements

12   *A. Rorissa / Information Processing and Management xxx (2008) xxx–xxx*

Appendix A. Coding scheme used

Level	Definition
Superordinate	A term that is more generic than a given basic level term. Superordinate level terms • 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 • e.g., vehicle, furniture, food, work, style, personality, crime, transportation, housework
Basic	A term that is more specific than a given superordinate level term. Basic level terms • 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, e.g., car, chair
Subordinate	A term that is more specific than a given basic level term. Subordinate level terms • 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 • e.g., sedan, rocking chair, sports car, dress shirt

Notes:

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

References

Armitage, L. H., & Enser, P. G. B. (1997). Analysis of user need in image archives. *Journal of Information Science, 23*, 287–299.
Blair, D. C. (1990). *Language and representation in information retrieval*. Amsterdam: Elsevier.
Chaudhry, A. S., & Jiun, T. P. (2005). Enhancing access to digital information resources on heritage: A case of development of a taxonomy at the Integrated Museum and Archives System in Singapore. *Journal of Documentation, 61*(6), 751–776.
Chen, H. (2001a). An analysis of image queries in the field of art history. *Journal of the American Society for Information Science and Technology, 52*(3), 260–273.
Chen, H. (2001b). An analysis of image retrieval tasks in the field of are history. *Information Processing & Management, 37*(5), 701–720.
Choi, Y., & Rasmussen, E. M. (2002). Users' relevance criteria in image retrieval in American history. *Information Processing & Management, 38*(5), 695–726.
Choi, Y., & Rasmussen, E. M. (2003). Searching for images: The analysis of users' queries for image retrieval in American history. *Journal of the American Society for Information Science and Technology, 54*(6), 498–511.
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.
Cutrell, E., & Dumais, S. T. (2006). Exploring personal information. *Communications of the ACM, 49*(4), 50–51.
Datta, R., Li, J., & Wang, J. Z. (2005). Content-based image retrieval: Approaches and trends of the new age. In Hongjiang Zhang, John Smith, & Qi Tian (Eds.), *Proceedings of the 7th ACM SIGMM international workshop on Multimedia information retrieval* (pp. 253–262). New York: ACM Press.
Enser, P. G. B. (2000). Visual image retrieval: Seeking the alliance of concept-based and content-based paradigms. *Journal of Information Science, 26*(4), 199–210.
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 *Language development*. In S. A. Kuczaj (Ed.). *Language, thought, and culture* (Vol. 2, pp. 301–334). Hillsdale, NJ: Lawrence Erlbaum.
Goodrum, A., & Spink, A. (2001). Image searching on the Excite Web search engine. *Information Processing & Management, 37*(2), 295–311.

650 Green, R. (2006). Vocabulary alignment via basic level concepts. OCLC/ALISE research grant report published electronically by OCLC Research. Retrieved
651   January 24, 2007, from http://www.oclc.org/research/grants/reports/green/rg2005.pdf.
652 Greisdorf, H., & O'Connor, B. C. (2002a). Modeling what users see when they look at images: A cognitive viewpoint. *Journal of Documentation, 58*(1), 6–29.
653 Greisdorf, H., & O'Connor, B. C. (2002b). What do users see? Exploring the cognitive nature of functional image retrieval. In E. G. Toms (Ed.), *Proceedings of
654   the 65th annual meeting of the American Society for information science and technology* (pp. 383–390). Medford, NJ: Information Today.
655 Hastings, S. K. (1995). Query categories in a study of intellectual access to digitized art images. In T. Kinney (Ed.), *Proceedings of the 58th annual meeting of the
656   American Society for information science* (pp. 3–8). Medford, NJ: Information Today.
657 Hearst, M. A., & Karadi, C. (1997). Searching and browsing text collections with large category hierarchies. In *CHI'97 extended abstracts on human factors in
658   computing systems. Looking to the future (Atlanta, Georgia, March 22–27, 1997)* (pp. 301–302). New York, NY: ACM.
659 Hearst, M. (2006a). Clustering versus faceted categories for information exploration. *Communications of the ACM, 49*(4), 59–61.
660 Hearst, M. (2006b). Design recommendations for hierarchical faceted search interfaces. ACM SIGIR Workshop on Faceted Search. Retrieved October 15,
661   2007, from http://flamenco.berkeley.edu/papers/faceted-workshop06.pdf.
662 Hollink, L., Schreiber, A. Th., Wielinga, B., & Worring, M. (2004). Classification of user image descriptions. *International Journal of Human Computer Studies,
663   61*, 601–626.
664 Horton, M., & Markman, E. (1980). Developmental differences in the acquisition of basic and superordinate categories. *Child Development, 51*, 708–719.
665 Iyer, H. (1995). *Classificatory structures: Concepts, relations and representation*. Frankfurt/Main: Indeks Verlag.
666 Jörgensen, C. (1995). Image attributes: An investigation. Unpublished doctoral dissertation, Syracuse University, Syracuse, New York.
667 Jörgensen, C. (1996). Indexing images: Testing an image description template. In S. Hardin (Ed.), *Proceedings of the 59th annual meeting of the American
668   Society for information science* (pp. 209–213). Medford, NJ: Information Today.
669 Jörgensen, C. (1998). Attributes of images in describing tasks. *Information Processing & Management, 34*(2/3), 161–174.
670 Jörgensen, C. (2003). *Image retrieval: Theory and research*. Lanham, MD; Oxford: Scarecrow Press.
671 Jörgensen, C., Jaimes, A., Benitez, A. B., & Chang, S. F. (2001). A conceptual framework and empirical research for classifying visual descriptors. *Journal of the
672   American Society for Information Science and Technology, 52*(11), 933–947.
673 Laine-Hernandez, M., & Westman, S. (2006). Image semantics in the description and categorization of journalistic photographs. In A. Grove & J. Stefl-Mabry
674   (Eds.), Proceedings of the 69th annual meeting of the American Society for information science and technology. Retrieved June 11, 2007, from http://
675   www.asis.org/Conferences/AM06/proceedings/papers/48/48_paper.html.
676 Lakoff, G. (1985). *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: The University of Chicago Press.
677 Layne, S. S. (1994). Some issues in the indexing of images. *Journal of the American Society for Information Science, 45*, 583–588.
678 Lew, M. S., Sebe, N., Djeraba, C., & Jain, R. (2006). Content-based multimedia information retrieval: State of the art and challenges. *ACM Transactions on
679   Multimedia Computing, Communications, and Applications, 2*, 1–19.
680 Marchionini, G. (2006). Exploratory search: From finding to understanding. *Communications of the ACM, 49*(4), 41–46.
681 Markkula, M., & Sormunen, E. (2000). End-user searching challenges indexing practices in the digital photograph archive. *Information Retrieval, 1*(4),
682   259–285.
683 Markkula, M., Tico, M., Sepponen, B., Nirkkonen, K., & Sormunen, E. (2001). A test collection for the evaluation of content-based image retrieval algorithms –
684   A user and task-based approach. *Information Retrieval, 4*(3–4), 275–293.
685 Markman, A. B., & Wisniewski, E. (1997). Similar and different: The differentiation of basic-level categories. *Journal of Experimental Psychology: Learning,
686   Memory, and Cognition, 23*(1), 54–70.
687 Matusiak, K. K. (2006). Towards user-centered indexing in digital image collections. *OCLC Systems & Services, 22*, 283–298.
688 Mervis, C. B., & Crisafi, M. A. (1982). Order of acquisition of subordinate, basic, and superordinate categories. *Child Development, 53*, 258–266.
689 Murphy, G. L., & Brownell, H. H. (1985). Category differentiation in object recognition: Typicality constraints on the basic category advantage. *Journal of
690   Experimental Psychology: Learning, Memory, and Cognition, 11*(1), 70–84.
691 Murphy, G. L., & Wisniewski, E. J. (1989). Categorizing objects in isolation and in scenes: What a superordinate is good for. *Journal of Experimental
692   Psychology: Learning, Memory, and Cognition, 15*(4), 572–586.
693 Neuendorf, K. A. (2002). *The content analysis guidebook*. Thousand Oaks, CA: Sage.
694 Neumann, D., & Gegenfurtner, K. R. (2006). Image retrieval and perceptual similarity. *ACM Transactions on Applied Perception, 3*, 31–47.
695 Nicholson, D., Dawson, A., & Shiri, A. (2006). HILT: A terminology mapping service with a DDC spine. *Cataloging & Classification Quarterly, 42*(3/4), 187–200.
696 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
697   American Society for Information Science and Technologya, 50*(8), 681–697.
698 O'Connor, B. C., & Wyatt, R. B. (2004). *Photo provocations: Thinking in, with, and about photographs*. Lanham, MD; Oxford: Scarecrow Press.
699 Panofsky, E. (1955). Meaning in the visual arts: Papers in and on art history. Garden City, NY: Doubleday.
700 Rafferty, P., & Hidderley, R. (2005). *Indexing multimedia and creative works: The problems of meaning and interpretation*. Aldershot: Ashgate.
701 Rasmussen, E. M. (1997). Indexing images. In M. E. Williams (Ed.), *Annual review of information science and technology* (pp. 169–196). Medford, NJ: Learned
702   Information.
703 Rorissa, A., & Iyer, H. (2008). Theories of cognition and image categorization: What category labels reveal about basic level theory. *Journal of the American
704   Society for Information Science and Technology, 59*(7).
705 Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General, 104*(3), 192–233.
706 Rosch, E., & Mervis, C. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7*(4), 573–605.
707 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.
708 Shatford, S. (1986). Analyzing the subject of a picture: A theoretical approach. *Cataloguing and Classification Quarterly, 6*(3), 39–62.
709 Shneiderman, B., Bederson, B. B., & Drucker, S. M. (2006). Find that photo!: Interface strategies to annotate, browse, and share. *Communications of the ACM,
710   49*(4), 69–71.
711 Smeulders, A. W. M., Worring, M., Santini, S., Gupta, A., & Jain, R. (2000). Content-based image retrieval at the end of the early years. *IEEE Transactions on
712   Pattern Analysis and Machine Intelligence, 22*(12), 1349–1380.
713 Soergel, D. (1985). *Organizing information: Principles of data base and retrieval systems*. Orlando, FL: Academic Press.
714 Stvilia, B., & Jörgensen, C. (2007). End-user collection building behavior in Flickr. In A. Grove & A. Rorissa (Eds.), Proceedings of the 70th annual meeting of
715   the American Society for information science and technology. Retrieved February 29, 2008, from http://www.asis.org/Conferences/AM07/proceedings/
716   papers/27/27_Paper.html.
717 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.
718 Tjondronegoro, D., & Spink, A. (2008). Web search engine multimedia functionality. *Information Processing and Management, 44*(1), 340–357.
719 Turner, J. M. (1994). Determining the subject content of still and moving image documents for storage and retrieval: An experimental investigation.
720   Unpublished doctoral dissertation, University of Toronto, Toronto, Canada.
721 Turner, J. M. (1995). Comparing user-assigned terms with indexer-assigned terms for storage and retrieval of moving images: Research results. In T. Kinney
722   (Ed.), *Proceedings of the 58th annual meeting of the American Society for information science* (pp. 9–12). Medford, NJ: ASIS.
723 Tversky, A. (1977). Features of similarity. *Psychological Review, 84*(4), 327–352.
724 Weber, R. P. (1990). *Basic content analysis* (2nd ed.). Thousand Oaks, CA: Sage.
725 Wisniewski, E. J., & Murphy, G. L. (1989). Superordinate and basic category names in discourse: A textual analysis. *Discourse Processes, 12*, 245–261.
726 Yee, K., Swearingen, K., Li, K., & Hearst, M. (2003). Faceted metadata for image search and browsing. In *Proceedings of the SIGCHI conference on human factors
727   in computing systems. CHI'03 (Ft. Lauderdale, Florida, USA, April 5–10, 2003)* (pp. 401–408). New York: ACM Press.
728