Classification and Categorization: A Difference that Makes a Difference

Elin K. Jacob

Abstract

Examination of the systemic properties and forms of interaction that characterize classification and categorization reveals fundamental syntactic differences between the structure of classification systems and the structure of categorization systems. These distinctions lead to meaningful differences in the contexts within which information can be apprehended and influence the semantic information available to the individual. Structural and semantic differences between classification and categorization are differences that make a difference in the information environment by influencing the functional activities of an information system and by contributing to its constitution as an information environment.

Introduction

Many different and sometimes conflicting responses can be made to the question "What is information?" Floridi (in press) identifies three broad categories intended to elucidate the predominant approaches to understanding the ambiguous phenomenon called information: information as reality (or ecological information), information for reality (or instructional information), and information about reality (or semantic information). The approach adopted here is that information is "differences that make a difference" (Bateson, 1979, p. 99). It is an emergent property—the result of meaningful differences—inherently semantic and therefore about reality.

Analysis of the syntactic differences that distinguish systems of classification from systems of categorization can contribute to a philosophy of information (PI) because these distinctions portend significant consequences for the processes that contribute to what Floridi (2002) describes

Elin K. Jacob, School of Library and Information Science, Indiana University–Bloomington, 1320 E. 10th St., Bloomington, IN 47405–1801
LIBRARY TRENDS, Vol. 52, No. 3, Winter 2004, pp. 515–540

as the "dynamics of information": "(i) *the constitution and modelling of information environments,* including their systemic properties, forms of interaction, internal developments etc.; (ii) *information life cycles,* i.e., the series of various stages in form and functional activity through which information can pass . . . and (iii) *computation,* both in the Turing-machine sense of *algorithmic processing* and in the wider sense of *information processing*" (p. 15. emphasis in original). Examination of the systemic properties and forms of interaction that characterize classification and categorization reveals fundamental differences in their respective organizational structures—differences that influence the functional activities of an information system and contribute to its constitution as an information environment.

The argument elaborated here is that fundamental syntactic distinctions exist between the structure of classification systems and the structure of categorization systems; that these distinctions lead to meaningful differences in the contexts within which information can be apprehended; and that these differences, in turn, influence the semantic information—the information about reality—that is available to the individual.

Information Systems

Shera (1960/1965) has observed that retrieval must be the focus of a theory of library and information science (LIS) and thus "the end toward which all our efforts are directed" (p. 136). Unfortunately, retrieval is too often viewed not as one component in an information system but as a self-contained and independent process. This emphasis on the end product—the retrieval of resources—tends to obscure the fact that effective retrieval depends on both the representation and the organization of a collection of information resources.

Soergel (1985) points out that, because information is used for problem-solving, information systems are developed and extended in response to the problems that confront society. Although this definition of information is not universally accepted, it is useful in understanding the complex set of processes that contribute to the ultimate effectiveness of an information system. Such a system identifies information resources that may be of use in addressing a particular problem; represents the attributes of resources that are relevant to the problem area; organizes these resource representations or the resources themselves for efficient access; and ultimately retrieves a set of resources in response to queries presented to the system by the individual. It would appear, then, that a more productive approach to the problem of retrieval would be to view an information system as a multidimensional whole comprised of several interrelated processes, including, at a minimum, collection development, representation, organization, and retrieval.

Retrieval is the final and therefore the most obvious of the processes that contribute to an information system. Because it is the only process in

which an individual actively participates, it is frequently the only process to which she gives serious consideration. When the individual is seeking information on a particular topic, her attention is focused on the set of resources retrieved by the information system. If these resources appear to be pertinent to the immediate problem, she may not give a second thought to the appropriateness of the terms used to query the information system. Nonetheless, it is the processes of selection, representation, and organization that provide the foundation without which information retrieval (IR) is less than effective, if not impossible. How resources are represented constrains the organizational structure(s) that can be imposed on a collection of information resources; the organizational structure of the collection dictates the search strategies that can be used for retrieval; and the representations themselves determine the set of resources that will be retrieved by the system.

Shera (1956/1965) affirmed the critical roles of representation and organization when he observed that effective retrieval requires an accord between the cognitive organization imposed on information by the individual and the formal organization imposed upon representations by the system. Shera's argument for accord between the individual and the retrieval system rests on three basic assumptions: that there are certain cognitive structures that can be identified and described; that it can be demonstrated that these structures are shared across individuals; and that identification of these shared structures will provide the basis for a theory of organization.

That cognitive accord can be achieved across individuals is a fundamental assumption of the shareability constraint proposed by Freyd (1983). She argued that the intent to communicate without loss of information causes the individual to modify her internal conceptual representations to reflect the cognitive organization assumed to be held by the other participant(s) in the communicative process. If participation in an intentional act of communication does promote normalization of conceptual representations across individuals, as Freyd (1983) argues, it follows that an intentional act of communication between the individual as natural intelligence and the information system would be subject to a similar shareability constraint. Assuming that the processes of representation, organization, and retrieval are necessarily interdependent, failure to address communication between the individual and the information system from the perspective of the system is a significant omission. Thus, an accounting of the dynamics of information should address the role of representation and organization in the creation and communication of meaningful information. More importantly, it should account for the semantic implications occasioned by differences in the forms of organization that can be used to structure an information system.

The need for effective communication between the information system and the individual points to five areas of research: (i) Is communication

between the information system and the individual influenced by the representation of resources? (ii) Does the organizational structure of the information system cause the individual to adjust her internal cognitive structures? (iii) Does the organization of resources contribute to the creation of a meaningful context for information? (iv) Is the meaning of information influenced by the organizational structure of the information system? and (v) What consequences follow from the different organizational structures that can be applied to a collection of information resources?

An understanding of the different forms of organizational structure and the implications that each holds for creating a meaningful context for information is foundational and must therefore precede any discussion of the role that representation and organization play in the dynamics of information. Accordingly, the focus here is on the ramifications of organizational structure for communication between the information system and the individual as natural intelligence. More specifically, the argument presented here addresses the fundamental structural and semantic differences between classification and categorization and how these differences make a difference in the information environment.

Categorization

Categorization is the process of dividing the world into groups of entities whose members are in some way similar to each other. Recognition of resemblance across entities and the subsequent aggregation of like entities into categories lead the individual to discover order in a complex environment. Without the ability to group entities based on perceived similarities, the individual's experience of any one entity would be totally unique and could not be extended to subsequent encounters with similar entities in the environment. Consider a situation in which each separate entity—each tree, each flower, or each drop of rain—was distinct from all other entities and carried its own unique set of defining characteristics. As Markman (1989) observes, the individual would not be able to handle the variety and complexity of her day-to-day interactions with the environment. By reducing the load on memory and facilitating the efficient storage and retrieval of information, categorization serves as the fundamental cognitive mechanism that simplifies the individual's experience of the environment.

Categorization divides the world of experience into groups or categories whose members share some perceptible similarity within a given context. That this context may vary and with it the composition of the category is the very basis for both the flexibility and the power of cognitive categorization. Zerubavel (1993) contends that the individual finds order and meaning in the environment by imposing boundaries—by splitting and lumping objects of experience so as to create distinct "islands of meaning" (p. 5). How an entity is categorized creates a context or conceptual frame that not only provides information about the entity but also shapes the in-

dividual's interaction with it. For example, the historic period known as the English Renaissance (1500–1650) is perceived as fundamentally different from the English Middle Ages even though England in the sixteenth century was, in many respects, quite similar to England in the fifteenth century. Splitting the sixteenth century from the fifteenth century by labeling them as belonging to two distinct historical periods focuses attention on the differences between them rather than on their similarities and provides the information that, in England, these differences were of greater import than differences between the fourteenth and fifteenth centuries.

Barsalou (1987) points out that this ability to manipulate the environment through the creation of categories allows the individual to forge new relationships and thus to create new information whose value exceeds the simple grouping of objects in the environment. He proposes that, because different features or properties are used to represent the same category at different times and in different contexts, the information associated with a particular category varies across individuals and across contexts. Thus the set of features associated with a category on any given occasion is composed of both context-dependent and context-independent information. Context-dependent information is relevant only within a particular context. For example, a high temperature of 50 degrees Fahrenheit might be described as cold on a summer day in southern Indiana, but warm or even hot on a winter day in the same locale. To say that it is cold outside conveys context-dependent information that is meaningful only in relation to the seasonal context. In contrast, context-independent information provides information about a category that is relevant across contexts. Even when used metaphorically, for example, the word "fire" connotes heat, light, and energy. The apparent instability of categories is therefore a reflection of the flexibility and the plasticity that are the power of the cognitive process of categorization and of the individual's ability to create and modify the informational content of a category as a function of immediate context, personal goals, or past experience.

The acquisition and transmission of information are dependent not only on the cognitive ability to create new categories—and thus new information—through the discovery of new patterns of similarity across entities, but also on the ability to capture information about these patterns through the medium of language. With the accumulation of more specialized knowledge and the creation of disciplinary domains, however, these categories and the relations between them have a tendency to become formalized (Jacob, 1994). The need to ensure that disciplinary knowledge is consistent across individuals and across time privileges the stability of reference provided by well-defined classes. As experientially-based categories evolve into well-defined, domain-specific classes that facilitate sharing of knowledge without loss of information, they lose their original flexibility and plasticity as well as the ability to respond to new patterns of similarity.

The Classical Theory of Categories

Until Rosch's publication in the 1970s of her seminal work on categories and categorization (Rosch, 1973, 1975), research in the area of categorization had focused on concept formation not as a process of creation but as a process of recognition. The world of experience was assumed to consist of a set of predetermined categories, each defined by a set of essential features represented by a category label; and all members of a given category were assumed to share a set of essential features that was identified by the category label and could be apprehended by all members of the linguistic community. Thus Hull (1920) wrote of the child's discovery of meaning in the word "dog" as the gradual recognition of a preexisting and invariant concept: "The 'dog' experiences appear at irregular intervals. . . . At length the time arrives when the child has a 'meaning' for the word dog. Upon examination this meaning is found to be actually a characteristic more or less common to all dogs and not common to cats, dolls and 'teddy-bears'" (Hull, 1920, pp. 5–6; cited in Brown, 1979, p. 188).

The presumption that a category is determined by a set of defining criteria is known as the "classical theory of categories." This is a simple but powerful theory that rests on three basic propositions (Smith & Medin, 1981; see also Taylor, 1989):

1. The intension of a category is a summary representation of an entire category of entities.
2. The essential features that comprise the intension of a category are individually necessary and jointly sufficient to determine membership within the category.
3. If a category (A) is nested within the superordinate category (B), the features that define category (B) are contained within the set of features that define category (A).

Proposition I states that the definition (*intension*) of a category is the union of the essential features that identify the membership (*extension*) of that category. Furthermore, because all members of a single category must share this set of essential features, each member is equally representative of the category as a whole. For this reason, the internal structure of a category is said to be ungraded, or without rank, because no member can be more typical or more representative of a category than any other member.

Proposition II states that, because each member of the category must exhibit all of the essential features that comprise the intension of the category, possession of the set of features that defines the category is sufficient to determine membership in the category. And, because there is a binary, either/or relationship that exists between an entity and a category such that an entity either is or is not a member of a particular category, the boundaries of categories are said to be fixed and rigid.

Proposition III identifies the inheritance relationship that exists be-

tween categories in a hierarchical structure: any member of a category that is a subset of a superordinate category must exhibit not only the set of essential features that determine membership in the subset but also the set of essential features that determine membership in any superordinate category within which the subset is nested.

In its most rudimentary form, categorization can be defined as the placement of entities in groups whose members bear some similarity to each other. Within the framework of the classical theory of categories, however, categorization is the process of systematically dividing up the world of experience into a formalized and potentially hierarchical structure of categories, each of which is defined by a unique set of essential feature(s). Because the intension of a category defines the set of essential features that each member of the category must exhibit, the classical theory maintains that intension equals extension—that membership within a particular category (extension) entails possession of the essential and defining character (intension) of the category. For example, if the intension of the category "bird" consists of the features "lays eggs," "has wings," "flies," and "builds nests in high places," every member of the category must exemplify the complete set of defining features. If an entity does not fly, it cannot be accorded membership in the category "bird" even if it does lay eggs, have wings, and build nests in high places. And, because all members of the category are defined by the same set of features, no one bird can be more typical or more representative of the category than any other bird. Thus, according to the classical theory, a parrot, a pigeon, and a puffin would be equally representative of the category "bird."

Brown (1979) observes that within the formalized and rigidly constrained ordering of reality established by the classical theory, category membership is absolute: ". . . any given thing is either in or out of the set" (p. 189). It is this stipulation that is the source of the classical theory's explanatory power: because it requires that intension equals extension—that membership in a category demonstrates possession of the set of essential features that define the category—the classical theory of categories would provide a simple yet elegant explanation for both the internal structure of cognitive representations and the semantic meanings of words.

Until recently, the classical theory of categories exemplified "the 'right way' to think about categories, concepts, and classifications" (Gardner, 1987, p. 340). But empirical research conducted over the past thirty years has challenged the validity of the assumptions on which this theory is founded. Critics of the classical theory have argued that the inability of subjects to identify the defining characteristics of an entity (Hampton, 1979; Rosch & Mervis, 1975) not only undermines the assumption that the set of essential features determining category membership is absolute but also calls into question the notion that these features are available to and can be specified by all members of a linguistic community. Demonstration of graded typi-

cality effects—the observation that subjects do judge certain members to be more representative of a category than others (McCloskey & Glucksberg, 1978; Rips, Shoben, & Smith, 1973; Rosch, 1973, 1975)—controverts the assumption that category structure is ungraded because all members are equally representative of the category. There is evidence, too, that subjects are able to rank both members and nonmembers of a category on a single continuum of representativeness. For example, Barsalou (1987) demonstrated that subjects could rank a robin, a pigeon, an ostrich, a butterfly, and a chair on a single continuum of representativeness for the category "bird"—a continuum extending from the most typical member of the category (robin) to the most atypical member (chair). The evidence for graded structure of categories points to the lack of fixed and determinate boundaries separating members of a category from nonmembers; and, buttressed as it is by demonstrations of category membership based on family resemblance (Rosch & Mervis, 1975), graded structure casts doubt on the classical assumption that there is an explicit inclusion/exclusion relationship between an entity and a category.

Classification

In LIS, the term "classification" is used to refer to three distinct but related concepts: a system of classes, ordered according to a predetermined set of principles and used to organize a set of entities; a group or class in a classification system; and the process of assigning entities to classes in a classification system. The focus here is on the first of these—on the classification system as a representational tool used to organize a collection of information resources—but a full appreciation of the implications of classification for information environments requires a basic understanding of the classification process itself.

Classification as process involves the orderly and systematic assignment of each entity to one and only one class within a system of mutually exclusive and nonoverlapping classes. This process is lawful and systematic: lawful because it is carried out in accordance with an established set of principles that governs the structure of classes and class relationships; and systematic because it mandates consistent application of these principles within the framework of a prescribed ordering of reality. The scheme itself is artificial and arbitrary: artificial because it is a tool created for the express purpose of establishing a meaningful organization; and arbitrary because the criteria used to define classes in the scheme reflect a single perspective of the domain to the exclusion of all other perspectives.

Taxonomic Classification.

Classification is perhaps best exemplified by the discipline of taxonomy. Broadly defined, taxonomy is the science of classification or, as Mayr (1982) defines it, "the theory and practice of delimiting kinds of organisms"

(p. 146). The objectives of taxonomic investigation are to provide an orderly and systematic organization of knowledge about the biological world; to identify the defining characteristics that distinguish a biological entity; and, based on those characteristics, to place the entity within a hierarchical ordering of mutually exclusive superordinate and subordinate classes in accordance with a set of established and widely accepted principles.

Taxonomic classification establishes stability of nomenclature through the aegis of a formalized and universally accepted language that facilitates transmission of knowledge across time and the barriers of natural language. Each class in the taxonomic scheme is given a unique name that is used to refer to all entities that display the complete set of features defining the class. And, because it is universally employed to identify all members of a given class, this label provides access to the accumulated knowledge about those entities, not as individuals but as members of a particular class. The taxonomic name establishes a relationship of equivalence between the set of features that define the class (its intension) and the set of entities that are members of the class (its extension). Using the taxonomic name, a member of a biological class is recognizable wherever it occurs, regardless of natural language or the local name(s) by which it may be known.

Through the inheritance of definitional criteria made possible by enforcing a principled structure of superordinate and subordinate classes, taxonomic classification also serves as an external cognitive scaffolding (Clark, 1997; Jacob 2001, 2002) that provides for the economical storage and retrieval of information about a class of entities. For example, the observation that Bleu is a poodle provides information about Bleu that is associated with the class "poodle." More importantly, however, it also provides information about Bleu that is available from the hierarchical structure within which the class "poodle" is nested—information associated with the superordinate classes dog, mammal, vertebrate, etc.

The essential observation, however, is that the practice of taxonomy is carried out within the arbitrary framework established by a set of universal principles. For example, while the naturalist Adanson, a contemporary of Linneaus, proposed a method for organizing botanical phenomena based on the identification of differences between individual specimens (Foucault, 1970), Linneaus advocated a systematic approach based on similarity of reproductive structure. For the naturalist following Linneaus's lead, any physical differences between two specimens not directly related to the process of reproduction would be irrelevant: for example, differences of leaf, stem, or root structure that might be used to distinguish between two plants would be ignored if the plants exhibited similar reproductive structures.

Taxonomic classification supports the efficient storage and retrieval of information about a class of entities, but reliance on a systematic approach such as that advocated by Linneaus constrains the information context by limiting the identification of knowledge-bearing associations to hierarchi-

cal relationships between classes. Furthermore, class definitions based on a single feature such as reproductive structure effectively reduce the amount of meaningful information that can be represented about each class in the taxonomy.

Classification Schemes.

A classification scheme is a set of mutually exclusive and nonoverlapping classes arranged within a hierarchical structure and reflecting a predetermined ordering of reality. Because a classification scheme mandates that an entity can be a member of one and only one class, it provides for communication of meaningful information through the systematic and principled ordering of classes. Furthermore, it establishes and enforces stability of reference by providing each class with a unique label that links individual members of the class to the class definition.

Shera (1951/1965) observes that, throughout history, attempts to classify knowledge have relied on four basic assumptions: universal order, unity of knowledge, similarity of class members, and intrinsic essence. The assumption of universal order posits an immutable conception of reality that serves as a unifying framework for all knowledge. The assumption of unity of knowledge presupposes that past, present, and future knowledge can be represented within a single, inclusive hierarchy of superordinate and subordinate classes. The assumption of similarity of class members holds that a class can be defined by a set of essential features and that these features are shared by all members of the class and distinguish that class from all other classes in the structure. And the assumption of intrinsic essence maintains that there is a set of individually necessary and jointly sufficient features that is intrinsic to all members of a class and that these features constitute the essence of the class.

With the possible exception of universal order, Shera's exposition of the assumptions that support efforts to organize knowledge can be interpreted in terms of the three propositions that constitute the classical theory of categories: the assertion that a category is defined by a summary representation (Proposition I) is a statement of the essential similarity of class members; the assertion that a category is defined by a set of essential features (Proposition II) is a statement of the intrinsic essence of a class; and the assertion that defining features are inherited in a hierarchical structure of categories (Proposition III) is a statement of the unity of all knowledge. It is instructive that, although the classical theory of categories is unable to account for the variability and flexibility of cognitive categorization, it does provide an elegant accounting of the fundamental assumptions on which classification schemes have historically been constructed.

Bibliographic Classification Schemes.

Traditionally, bibliographic classifications have been deductive, top-down schemes that enumerate a set of mutually exclusive classes. An enu-

merative classification scheme begins with a universe of knowledge and a theory of organization or set of principles that establishes the conceptual structure of the scheme. Whether the universe encompasses all knowledge or is limited to a specific domain, construction of the scheme involves the logical process of division and subdivision of the original universe such that each class, or each level of classes in the structure, is differentiated by a particular characteristic or property (e.g., the property "color" or "shape"). The result is a hierarchical structure of generic (genus/species) relationships wherein each subordinate class is, theoretically, a true species of the superordinate within which it is nested.

Faceted (analytico-synthetic) classification systems are inductive, bottom-up schemes generated through a process of analysis and synthesis. Construction of the faceted structure begins with analysis of a universe of knowledge to identify the individual elements—properties and features—of the universe. These elements are then organized into mutually exclusive groups on the basis of conceptual similarity, and these groups are, in turn, arranged in successively larger groupings to form facets (aspects) that can be used to represent entities in the universe. In this way, meaningful relationships are established not only between the elements in a group but between the groups themselves. The result is not a classification scheme but a controlled vocabulary of concepts and their associated labels that can be used, in association with a notation and a prescribed citation order, to synthesize the classes that will populate the classification scheme. A faceted vocabulary for classifying cars might include mutually exclusive facets for "color" (red, blue, black), "body style" (sedan, convertible, minivan), and "transmission" (manual, automatic). Following the citation order *body style—transmission—color*, classes would be constructed by selecting a single value, or isolate, from each facet. Examples of the classes that could be constructed in this faceted scheme would be convertible—manual—red and minivan—automatic—blue.

Because a faceted classification scheme adheres to a fixed citation order during the construction of individual classes, the resulting structure, like an enumerative scheme, is necessarily hierarchical. In fact, it is the hierarchical nature of bibliographic classification schemes that allows for the arrangement of physical resources on the shelves of the library. "Reading" a classification scheme involves moving down the hierarchy, from superordinate to subordinate and from left to right, to generate a series of relationships between classes that can be translated into the linear order of the library shelf. It is just this linear structure that Ranganathan captured in the notion of APUPA (or Alien-Penumbral-Umbral-Penumbral-Alien). The umbral class (U) represents the focal topic; penumbral classes (P) are those most closely related to the focal topic; and alien classes (A) are those removed from and therefore unrelated to the focal topic. When the individual reviews a collection of resources arranged in classified order, she

generally begins with the most relevant class or focal topic (U); moving either to the right or the left, she progresses from resources on the focal topic through closely related materials (P) to those resources which are unrelated (A). In this fashion, the linearity inherent in the hierarchical structure of a classification scheme is used to create a meaningful context by bringing into proximity those classes within the hierarchical structure which are theoretically most closely related.

Linearity is, in fact, the first of seven properties that Shera (1953/1965) identifies as characteristic of a bibliographic classification scheme: linearity; inclusivity of all knowledge within the classification's universe; well-defined, specific, and meaningful class labels; an arrangement of classes that establishes meaningful relationships between them; distinctions between classes that are meaningful; a mutually exclusive and nonoverlapping class structure; and an infinite hospitality that can accommodate every entity in the bibliographic universe. Each of these properties contributes to Shera's definition of a bibliographic classification scheme as

> a list of terms which are specifically and significantly different each from the other, capable of describing the subject content of [resources], inclusive of all knowledge, infinitely hospitable, in an arrangement that is linear, unique, and meaningful, and which when applied to [resources], usually, though not necessarily, through the medium of a notation, results in their arrangement on the shelves according to the logical principles that inhere in the schematism. (Shera, 1953/1965, p. 99)

In other words, a bibliographic classification establishes a controlled vocabulary in the form of a set of uniquely labeled classes that serve both to define and to organize the intellectual content of a collection of resources. Furthermore, this vocabulary determines the conceptual boundaries of the scheme's universe by including only the knowledge that is relevant within the immediate universe. The resulting arrangement is meaningful precisely because it constitutes a principled context for information—a context shaped by class definitions, by information-bearing, hierarchical relationships and by meaningful distinctions between classes and, by extension, between the concepts that those classes represent.

Classification as a Disciplinary Language.

A classificatory structure frequently inheres in a disciplinary language when it is used to establish a specific conceptual context that both defines and organizes the domain of investigation (Foucault, 1970; Jacob, 1994). The language serves to prescribe the boundaries of the domain; to determine both the subject matter of the domain and the relationships that obtain between phenomena of investigation; to legitimize specific concepts and methodologies; to ensure effective transmission of knowledge by stabilizing the vocabulary; and to foster a domain-specific perspective or dis-

ciplinary episteme. Because a disciplinary language reflects the underlying classificatory structure of the domain, the meaning of any class term can only be apprehended within the conceptual context established by the classificatory structure.

The Difference between Classification and Categorization

Although there are obvious similarities between classification and categorization, the differences between them have significant implications for the constitution of an information environment. Failure to distinguish between these two systems of organization appears to stem from the misconception that they are, in fact, synonymous—a misconception that may be reinforced by the fact that both are mechanisms for organizing information.

The literature on categorization is riddled with passages where the terms "classification" and "categorization" are used indiscriminately to refer to the same process. Rosch et al. (1976) provides an illustrative example of how these two terms are used indiscriminately:

> . . . one purpose of *categorization* is to reduce the infinite differences among stimuli to behaviorally and cognitively usable proportions. It is to the organism's advantage not to differentiate one stimulus from others when that differentiation is irrelevant for the purposes at hand. The basic level of *classification,* the primary level at which cuts are made in the environment, appears to result from the combination of these two principles; the basic *categorization* is the most general and inclusive level at which categories can delineate real-world correlational structures. (Rosch et al., 1976, p. 384. Emphasis added)

This lack of distinction between *category/categorization* and *class/classification* is frequently compounded by the use of *concept* as yet another synonym for *category* (e.g., Gardner, 1987, p. 340). Unfortunately, this terminological imprecision obscures the fact that researchers are actually dealing with two similar but nonetheless distinct approaches to organization.

Although systems of classification and categorization are both mechanisms for establishing order through the grouping of related phenomena, fundamental differences between them influence how that order is effected—differences that do make a difference in the information contexts established by each of these systems. While traditional classification is rigorous in that it mandates that an entity either is or is not a member of a particular class, the process of categorization is flexible and creative and draws nonbinding associations between entities—associations that are based not on a set of predetermined principles but on the simple recognition of similarities that exist across a set of entities. Classification divides a universe of entities into an arbitrary system of mutually exclusive and nonoverlapping classes that are arranged within the conceptual context established by

a set of established principles. The fact that neither the context nor the composition of these classes varies is the basis for the stability of reference provided by a system of classification. In contrast, categorization divides the world of experience into groups or categories whose members bear some immediate similarity within a given context. That this context may vary—and with it the composition of the category—is the basis for both the flexibility and the power of cognitive categorization (Jacob, 1992).

Figure 1 identifies six systemic properties that serve as a starting point for comparing systems of classification and categorization: (i) process, (ii) boundaries, (iii) membership, (iv) criteria for assignment, (v) typicality, and (vi) structure.

(i) The process of classification involves systematic arrangement of classes of entities based on analysis of the set of individually necessary and jointly sufficient characteristics that defines each class. In contrast, the process of categorization is generally unsystematic but inherently creative in that it need not rely on predetermined definitions but is able to respond to similarity assessments based on immediate context, personal goals, or individual experience.

Figure 1. Comparison of Categorization and Classification.

Categorization	Classification
Process	
Creative synthesis of entities based on context or perceived similarity	Systematic arrangement of entities based on analysis of necessary and sufficient characteristics
Boundaries	
Because membership in any group is non-binding, boundaries are "fuzzy"	Because classes are mutually-exclusive and non-overlapping, boundaries are fixed
Membership	
Flexible: category membership is based on generalized knowledge and/or immediate context	Rigorous: an entity either *is* or *is not* a member of a particular class based on the intension of a class
Criteria for Assignment	
Criteria both context-dependent and context-independent	Criteria are predetermined guidelines or principles
Typicality	
Individual members can be rank-ordered by typicality (graded structure)	All members are equally representative (ungraded structure)
Structure	
Clusters of entities; may form hierarchical structure	Hierarchical structure of fixed classes

(ii) Systems of classification and categorization are also distinguished by the boundaries imposed on groupings. Because the classes in a classification system are rigidly circumscribed by the intension of the class and further constrained by the requirement that they be mutually exclusive and nonoverlapping, boundaries between classes are fixed, determinate, and persistent. In a categorization system, however, membership of an entity in any one category is nonbinding and does not prohibit membership in any other category. Thus the membership of any two or more categories in a system of categorization may overlap or vary across time in response to changing contexts. This is possible because category boundaries are not simply fuzzy but are, in fact, mutable and potentially fluid.

(iii) and (iv) Membership and criteria for assignment are two closely related characteristics that distinguish systems of classification from systems of categorization. In a classification system, criteria for class assignment—the set of necessary and sufficient features that constitutes the intension of a class—are governed by principles that establish the conceptual framework of the system. Membership in a class is rigorous in that it is determined by the intension of the class: an entity either is or is not a member of any class in the system. More importantly, however, membership in a class is absolute simply because an entity can belong to one and only one class. In contrast, the criteria for category assignment employed by a system of categorization are potentially variable, allowing the membership of a category to respond to the demands of the context in which it is used. In this way, the membership of a category may vary across time based on the combination of context-dependent and context-independent information that is used to define category membership.

Differences in the criteria for assignment emphasize an important distinction between classification and categorization. In systems of classification, class assignment relies on definitions that are "idealizations" or "theoretical abstractions" (Barsalou, 1987) to determine class membership. In systems of categorization, however, category assignment is flexible and dynamic, reflecting the ability of the individual to modify category definitions in response to variations in the immediate environment. Thus Barsalou argues that

> . . . the concepts that theorists "discover" for categories may never be identical to an actual concept that someone uses. Instead, they may be analytic fictions that are central tendencies or idealizations of actual concepts. Although such theoretical abstractions may be useful or sufficient for certain scientific purposes, it may be more fruitful and accurate to describe the variety of concepts that can be constructed for a category and to understand the process that generates them. (Barsalou, 1987, p. 120)

(v) Typicality is closely related to the characteristics of membership and criteria for assignment. However, typicality is potentially ambiguous: on the

one hand, typicality is used as an indication of the individual's assessment as to how representative a member is of its particular class or category; and, on the other hand, it is used as a reflection of the assumptions regarding membership and membership criteria that govern a system of classification or categorization. Because empirical research indicates that subjects are capable of ranking members according to typicality even when working with well-defined, either/or classes such as *odd number* or *even number* (Armstrong, Gleitman, & Gleitman, 1983), attempting to distinguish between classification and categorization on the basis of an individual's typicality judgments would be an exercise in futility. In contrast, systemic assumptions governing membership do provide an important point of distinction between classification and categorization.

In a system of classification, all members of a class must display the full set of essential features prescribed by the class definition (see Proposition I of the classical theory). It follows, then, that all members are assumed to be equal and therefore equally representative of the class. For this reason, the internal structure of a class is said to be ungraded because no entity can be a "better" member of the class than any other member. However, in a system of categorization, there is no assumption of equality of membership. The fact that individuals can identify particular members as more typical of a category reflects the dynamic nature of category definitions and the corresponding variability of category membership as a reflection of immediate context. The internal structure of a category is said to be ungraded because it is possible to rank category members as to how typical or representative they are of the category as a whole.

(vi) Structure is perhaps the single most important characteristic that can be used to discriminate between systems of classification and categorization because it is influenced by distinctions based on process, boundaries, membership, and criteria for assignment. A classification system is generally a hierarchical structure of well-defined, mutually exclusive, and nonoverlapping classes nested in a series of superordinate-subordinate or genus-species relationships. The structure of a classification system provides a powerful cognitive tool—an external scaffolding (Clark, 1997; Jacob 2001, 2002)—that minimizes the cognitive load on the individual by embedding information about reality through the organization of classes within the system. For example, because an entity either is or is not a member of a particular class in a system of classification, it provides for determination of class membership as a relatively simple pattern-matching or pattern-completing activity. At a more complex level, the structure of the classification system establishes information-bearing relationships between classes: vertical relationships between superordinate and subordinate classes that are subject to the mechanism of inheritance illustrated above in the example of the poodle Bleu; and lateral relationships between coordinate classes that occur at the same level in the hierarchy and, when taken together, consti-

tute the immediately superordinate class within which they are nested. In this fashion, the structure of a classification system serves as a medium for the accumulation, storage, and communication of information associated with each class in the structure; and, by capitalizing on the hierarchical and lateral relationships between classes, it minimizes the information that must be stored with each class and reduces the load on memory.

In contrast, the structure of a categorization system consists of variable clusters of entities that may or may not be organized in a hierarchical structure. Because categories are not constrained by a requirement for mutual-exclusivity, membership in one category does not prohibit membership in any other category. More importantly, however, the very plasticity that is the creative power of categories may actually prohibit the use of categorization as a persistent information structure. The potentially transitory and overlapping nature of categories provides that any relationships established between categories are themselves mutable. Thus a system of categorization creates a conceptual framework whose meaning may be short-lived and ephemeral—a conceptual framework that cannot function as cognitive scaffolding and whose ability to serve as a medium for the accumulation, storage, and communication of information is limited.

Ordering, Grouping, and Organization

A system for ordering (Jacob & Loehrlein, 2003) provides access to resources by arranging them in some recognizable order. Typically, these systems will employ alphanumeric or chronological sequences because these arrangements generate syntactic patterns that are familiar to a majority of individuals. Although such a system is intended to support access to known items, it may appear to create groupings of similar resources (e.g., all individuals with the last name *Smith* or alumni who graduated in the year 2000), but the imposition of sequential order is nonetheless a purely syntactic device that cannot create meaningful relationships either between individual entities or between groups of entities.

In contrast, a system of organization (Jacob & Loehrlein, 2003) is a unified structure that establishes a network of relationships among the classes or categories that comprise the system. These relationships are meaningful and information-bearing because they specify principled connections between two or more groups within the same system. Thus, with a single possible exception, classification systems are systems of organization because they provide for the conceptual arrangement of a set of mutually exclusive and nonoverlapping classes within a systematic structure of hierarchical, genus-species relationships.

The exception is a constitutive classification (Jacob, Mostafa, & Quiroga, 1997) consisting of a set of mutually exclusive classes that comprise the totality of a given universe but lack nested, superordinate-subordinate relationships. For example, the classes freshman, sophomore, junior, and

senior comprise the universe of college undergraduates. These classes appear to evince a hierarchical ordering (e.g., from freshman to senior), but they fail to demonstrate meaningful, information-bearing relationships: although a senior may be assumed to have been a junior at some point in time, the class junior is not a true species of its purported superordinate senior. Thus a constitutive classification does not qualify as a system of organization because, even though it is comprised of a set of mutually exclusive and nonoverlapping classes that constitute the totality of a particular universe, it fails to establish meaningful relationships between its constituent classes. It is interesting, too, that neither a hierarchical nor a constitutive classification can serve as a system for ordering: because the distinctions between classes are conceptual, the classes cannot conform to a recognizable, syntactic pattern of arrangement. Furthermore, both hierarchical and constitutive systems of classification require an index or other auxiliary mechanism to support access, whether to unique resources or to individual classes in the structure.

A system of categorization may or may not be a system of organization. Although a categorization system groups entities on the basis of similarity, the example of a constitutive classification demonstrates that the simple identification of a set of categories without the establishment of meaningful, information-bearing relationships does not constitute a system of organization. But, even though a categorization system does not indicate meaningful relationships, it is not a system for ordering: the simple fact of grouping entities into categories does not support access. Because categorization reflects conceptual distinctions between groups of entities, it, too, requires an auxiliary mechanism to provide access, whether to individual categories or to unique category members.

If a system of categorization does not impose a systematic, syntactic order on its member categories and if it does not establish meaningful relationships between categories, then it is simply a mechanism for grouping. For example, dividing the items on a shopping list into categories defined by place of purchase (e.g., grocery store, gas station, and five-and-dime store) is a mechanism for grouping that simplifies the individual's interaction with her environment but neither creates meaningful relationships between categories nor imposes any recognizable order on them. A constitutive classification is also an example of a simple mechanism for grouping: in this case, for dividing a universe of entities into a set of well-defined and mutually exclusive groups without the identification of any meaningful relationships among them.

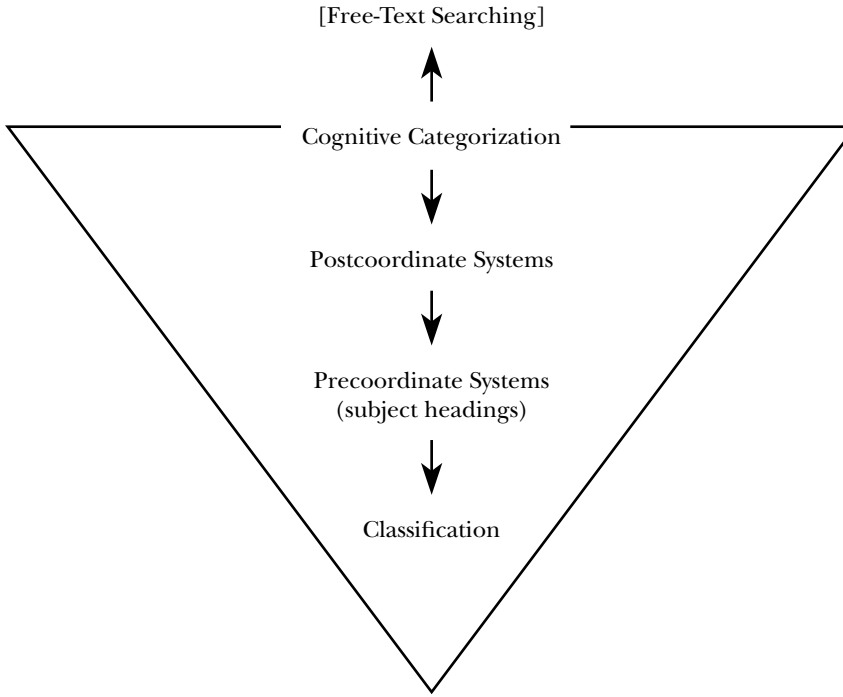
Implications of Structure

The functional role of structure in the creation and enhancement of information contexts can be addressed through analysis of four general approaches to the organization and retrieval of resources: free-text search-

ing, postcoordinate indexing, precoordinate indexing, and classification (see Figure 2). Although cognitive categorization serves as the baseline for this analysis, it is removed from consideration as a system of organization, not because it lacks semantic foundation or relational structure, but because, contrary to the arguments proffered by Shera (1956/1965), the organization imposed on cognitive categories is so dynamic and responsive to changes in context that it cannot establish persistent, knowledge-bearing relationships between categories.

Of the four general approaches to organization, free-text searching is the least constrained. Although it shares with systems of classification the creation of mutually exclusive, nonoverlapping, and rigidly bounded classes whose membership is constrained by an explicit criterion of assignment (i.e., the alphanumeric search string used to query the system), free-text searching lacks an established set of principles that governs the structure of classes and class relationships. It can be described as a system of categorization in the very broadest sense, but it is, at best, a very elementary mechanism for grouping. Even as a grouping mechanism, however, it has two significant flaws. In the first place, the basis for grouping is purely syntactic: because the criterion for group assignment involves the simple match-

Figure 2. Systems of Organization.



ing of alphanumeric strings, groups produced by this process share a superficial similarity without deeper semantic implications. In the second place, the process of free-text grouping is binary in that it generates only two groups of entities—those that match the query string and those that do not. However, because free-text searching lacks a semantic base, it cannot support meaningful distinctions between these two classes, and, because it exemplifies the very simplest of structures (i.e., two antonymous classes), a free-text retrieval system cannot contribute to an information environment that will support or enhance the value of system output through the establishment of meaningful context.

Unlike free-text searching, postcoordinate systems, precoordinate systems, and classification systems are all indexing systems in that each involves the assignment to a resource of one or more descriptors intended to represent the intellectual content of that resource. These descriptors are usually drawn from a controlled vocabulary or indexing language that normalizes the vocabulary used in representation and retrieval by creating an indexical, one-for-one correspondence between a descriptor and the concept to which it points. The indexing language also provides for communication between the system and the individual by specifying the set of authorized terms or subject strings that can be used to pose search queries to the system. Although a descriptor may be a class label, a subject heading or a single term or phrase, depending on the nature of the system, each descriptor serves to identify or describe the intellectual content of a group of resources. Unlike an access point in a system for ordering that supports the retrieval of a unique entity, a descriptor is a surrogate for (or a pointer to) the intellectual content shared by a group of resources. Indeed, indexing, like categorization, would be impossible if every resource were to be treated as a unique entity.

In the progression from postcoordinate indexing systems through precoordinate indexing systems to systems of classification, organizational structure becomes increasingly more constrained (see Figure 2). It is appropriate, then, to begin this analysis with classification, the most highly constrained of these three systems, and to work back through the less constrained systems toward the baseline of cognitive categorization.

Theoretically, a classificatory structure epitomizes a system of organization because it creates a principled structure of well-defined classes that are linked by a system of hierarchical, genus-species relationships. Although practice does not always adhere to theory in the development of classification schemes, classification is nonetheless the most rigid of organizational systems because its structure of mutually exclusive and nonoverlapping classes mandates an absolute relationship between a resource and its class: each resource may be assigned to one and only one class in the structure. Thus the process of classification is inherently systematic because it is gov-

erned by a set of principles that serves as a persistent conceptual framework for the creation of meaningful, structural relationships between classes.

Although the well-defined structure of a classification system provides for the creation of meaningful, information-bearing relationships between classes—relationships that facilitate the use of classification as an external cognitive scaffolding—it places powerful limitations on communication between the individual and the information system. In an information system whose class structure is predetermined, the retrieval set returned for any query posed to the system is necessarily limited to the membership of a single class. Thus the structure of the classification system constrains the questions that can be presented to the system by prescribing the set of possible answers before a query has actually been posed. Within a classificatory structure, then, communication is one-way—from the system to the individual—and the individual must rely on her understanding of or intuitions about the structural relationships between classes in order to interact with the system in an effective and meaningful way.

Information systems are identified as precoordinate when the categories or classes that comprise the system are either assigned or built by the indexer at the time of indexing. A classification system is obviously a precoordinate system because its classes are either established by the classificationist during scheme generation or built by the classifier at the time of class assignment using a faceted vocabulary and a fixed citation order. A subject heading system is also a precoordinate system but it is generally less constrained—and less constraining—than a classification system. Whereas classification mandates assignment of a resource to one and only one class, a precoordinate system of subject headings does not require individual groups to be mutually exclusive. Rather, subject heading systems allow for the assignment of multiple descriptors to a single resource, thereby providing multiple access points for each entity rather than the single access point (the unique class label) prescribed by a classification system.

Because it does not demand a well-defined and absolute relationship between a resource and a subject heading—because it does not require that the groups of entities associated with individual subject headings will necessarily be mutually exclusive—a precoordinate subject heading system is, in fact, a system of categorization. Categories formed by the subject heading system are not rigidly bounded but frequently overlap, with individual members spilling over into penumbral and even alien categories. Although allowing multiple descriptors for a single resource provides for greater variability in the range of resources that can be retrieved with a single query, the questions that can be posed to the information system are nonetheless limited, as they are in a system of classification, by the authorized set of subject heading strings that comprise the system. And, as with a classification system, the retrieval set generated in response to a query is determined

by the indexer: the assignment of subject headings as descriptors not only constrains the questions that can be posed to the system but serves to establish the specific set of resources that can be retrieved in response to each query posed to the system.

Unlike the systematic and principled structure of a classification system, the structure of a subject heading system is frequently unprincipled, unsystematic and polyhierarchical. And, unlike the relationships established between well-defined and mutually exclusive classes in a classification, any relationships created between the categories of a subject heading system cannot be assumed to be either meaningful or information-bearing. An example from *Subject Headings for Schools and Public Libraries* (Fountain, 2001) illustrates the lack of knowledge-bearing relationships that characterizes many subject heading systems. The heading "Rats as carriers of disease" combines two broader concepts: "rats" and "disease." Although it is obvious that "Rats as carriers of disease" is somehow related to both rats and disease, this heading is neither a kind of "Rat" nor a kind of "Disease." Because the specific value of any relationship that might link this heading to its broader concepts is unidentified, the relationship must be supplied by the individual if the heading is to be linked in a meaningful way to other concepts in the subject heading system.

Although subject heading systems appear to create relationships between headings, these relationships are often descriptive, idiosyncratic, and, sometimes, potentially meaningless. For example, the *Library of Congress Subject Headings* (Library of Congress. Cataloging Policy and Support Office, Library Services, 2002) identifies the subject heading "Humanities" as the broader term for the heading "Philosophy." It then proceeds to list "Humanism" as the broader term for "Humanities" and "Philosophy" as the broader term for "Humanism." Thus the supposed nesting structure is circular: "Philosophy" > "Humanities" > "Humanism" > "Philosophy." Obviously, the absence of either a well-defined indexing language or principled and meaningful relationships between subject headings undermines the ability of the system to establish a context that can contribute to the apprehension of information.

As with classification, communication between the individual and a subject heading system tends to be one-way—from the system to the individual—but the unprincipled structure of many subject heading systems and the general lack of a prescriptive conceptual framework that can support information-bearing relations undermines the potential for meaningful communication between the user and the system. This is an important distinction between subject heading systems and the more structured classification system that can be explained, in part, as a difference between the processes of identification and predication. Classification involves a process of identification (or definition) in that it asserts a meaningful, one-for-one relationship between an entity and its class, but a precoordinate system of

subject headings involves a process of predication (or description) that allows for multiple assertions to be ascribed to a single resource. While a system based on predication demonstrates greater creativity, flexibility, and hospitality than the well-defined structure of a system based on identification, the very rigidity of the latter actually supports the creation and persistence of information-bearing relationships that are simply not possible in the looser structure of the former.

Precoordinate systems constrain communication between the individual and the system through the establishment of a finite collection of class labels or subject headings that serve as the complete set of possible search queries and predetermine the composition of retrieval sets. In contrast, postcoordinate systems predetermine neither the queries nor the retrieval sets but allow the individual to build her own category definitions that can be presented to the system as search queries at the time of retrieval. Descriptors representing the intellectual content of a resource are assigned by the indexer at the time of indexing. During retrieval, the individual builds her own search categories by combining descriptors with Boolean logic.

By allowing the individual to generate her own queries, the postcoordinate system supports a more interactive form of communication between searcher and system. In most postcoordinate systems, descriptors are assigned from a controlled vocabulary. In others, however, communication between the individual and the information system is complicated by the fact that the indexing language does not exist as a controlled vocabulary but is extracted by the indexer from terms occurring in the resource being indexed. Generally, however, the generation of category definitions as postcoordinate search queries is limited only by the set of individual terms that comprise the indexing language. Although the resources that participate in a retrieval set are determined by the indexer's assignment of descriptors, communication between the system and the individual is greatly enhanced by her ability to create her own queries that will capture her immediate information need.

Unfortunately, however, the flexibility of category generation, like the process of cognitive categorization, goes hand-in-hand with the absence of meaningful relationships. As with any free-text information system, posing a query to a postcoordinate system simply divides the collection into two groups: the set of resources whose assigned descriptors match the search query and the remaining resources whose descriptors do not match the query. Obviously, postcoordinate systems, like free-text systems, are simply mechanisms for grouping, not systems of organization. Unlike free-text systems, however, the basis for grouping in a postcoordinate system is semantic, not syntactic. Although the postcoordinate system is simply matching strings, the indexer imposes a certain level of conceptual control by assigning simple descriptors from an indexing language that establishes an indexical, one-for-one relationship between a descriptor and its referent.

The individual is empowered to create unique and potentially idiosyncratic search categories precisely because the system itself does not establish any but the simplest categories—those defined by the individual descriptors assigned by an indexer. Because the system fails to establish a principled framework that provides for the establishment of information-bearing relationships between categories, the postcoordinate system can neither create nor contribute to an information context precisely because there is no persistent structure that could support meaningful relationships between categories.

Conclusion

This very preliminary review of the properties and features of the different approaches to organizing, ordering, or simply grouping information resources has barely scratched the surface in addressing structural distinctions between systems of classification and systems of categorization and how these distinctions affect interaction with the system as an information environment.

For example, at a very superficial level, the strength of classification is its ability to establish relationships between classes that are stable and meaningful. But the rigidity of structure that supports these relationships has its corresponding disadvantages. In particular, traditional classification systems are context-independent: because the relationships established by classification are invariant and persist across time and space, these systems are resilient to the context of use and severely constrain the individual's ability to communicate with the system in a meaningful and productive manner. In contrast, systems of categorization, and especially postcoordinate systems, are highly responsive to—even dependent on—the immediate context. The utility of these systems as information environments depends ultimately on provisions for effective communication with the individual. But the responsiveness and flexibility of the postcoordinate system effectively prohibit the establishment of meaningful relationships because categories are created by the individual, not the system, and are thus fleeting and ephemeral.

It is important for philosophers, theoreticians, and developers to work toward a more in-depth and comprehensive understanding of how the structure of an information system contributes to the establishment of semantic context; how different forms of organization support communication between the searcher and the system; and how concrete organizational structures and specific types of relationships contribute to the production of meaningful information environments. The search for adequate explanations of these issues will ultimately contribute to a deeper understanding of the "dynamics of information" (Floridi, 2002) and the implications that the structure of information systems holds for the composition of and interaction with information environments.

ACKNOWLEDGMENTS

I would like to thank Aaron Loehrlein for his thoughtful reading of and commentary on preliminary drafts of this paper and for the many conversations that contributed so richly to its theoretical content. I would also like to thank Ken Herold for his very careful and insightful reading of the final draft. This exploration of the role of structure in the generation of a semantically meaningful information environment is in its nascent stages, and I would like to thank Ken for the opportunity to develop these ideas for presentation in this venue.

REFERENCES

Armstrong, S. L., Gleitman, L. R., & Gleitman, H. (1983). What some concepts might not be. *Cognition, 13,* 263–308.

Barsalou, L. W. (1987). The instability of graded structure: Implications for the nature of concepts. In U. Neisser (Ed.), *Concepts and conceptual development: The ecological and intellectual factors in categorization* (pp. 101–140). Cambridge: Cambridge University Press.

Bateson, G. (1979). *Mind and nature: A necessary unity.* New York: Dutton.

Brown, R. (1979). Cognitive categories. In R. A. Kasschau & C. N. Cofer (Eds.), *Psychology's second century: Enduring issues* (pp. 188–217). New York: Praeger.

Clark, A. (1997). *Being there: Putting brain, body, and world together again.* Cambridge, MA: MIT Press.

Floridi, L. (2002). What is the philosophy of information? *Metaphilosophy, 33*(1/2). Retrieved August 20, 2003, from http://www.wolfson.ox.ac.uk/~floridi/pdf/wipi.pdf.

Floridi, L. (in press). Open problems in the philosophy of information. *Metaphilosophy.* Retrieved August 20, 2003, from http://www.wolfson.ox.ac.uk/~floridi/pdf/oppi.pdf.

Foucault, M. (1970). *The order of things: An archaeology of the human sciences.* New York: Vintage Books.

Fountain, J. F. (2001). *Subject headings for school and public libraries: An LCSH/Sears companion* (3rd ed.). Englewood, CO: Libraries Unlimited.

Freyd, J. J. (1983). Shareability: The social psychology of epistemology. *Cognitive Science, 7,* 191–210.

Gardner, H. (1987). *The mind's new science: A history of the cognitive revolution.* New York: Basic Books.

Hampton, J. A. (1979). Polymorphous concepts in semantic memory. *Journal of Verbal Learning and Verbal Behavior, 18,* 441–461.

Hull, C. L. (1920). Quantitative aspects of the evolution of concepts. *Psychological Monographs, 28,* Whole No. 123. [Cited in Brown, 1979]

Jacob, E. K. (1992). Classification and categorization: Drawing the line. In B. H. Kwasnik & R. Fidel (Eds.), *Advances in classification research, Vol. 2. Proceedings of the 2nd ASIS SIG/CR Classification Workshop: Held at the 54th ASIS Annual Meeting, Washington, DC, October 27–31, 1991* (pp. 67–83). Medford, NJ: Learned Information.

Jacob, E. K. (1994). Classification and crossdisciplinary communication: Breaching the boundaries imposed by classificatory structure. In H. Albrechtson & S. Oernager (Eds.), *Knowledge organization and quality management: Advances in knowledge organization, vol. 4* (pp. 101–108). Frankfurt/Main: Indeks Verlag.

Jacob, E. K. (2001). The everyday world of work: Two approaches to the investigation of classification in context. *Journal of Documentation, 57*(1), 76–99.

Jacob, E. K. (2002). Augmenting human capabilities: Classification as cognitive scaffolding. In M. J. López-Huertas & F. J. Muñoz-Fernández (Eds.), *Challenges in knowledge representation and organization for the 21st century. Integration of knowledge across boundaries: Proceedings of the Seventh International ISKO Conference, 10–13 July 2002, Granada, Spain* (pp. 38–44). Würzburg, Germany: Ergon Verlag.

Jacob, E. K., & Loehrlein, A. (2003). What ontologies are not: A [draft] theoretical framework for the analysis of representational systems. Paper presented in the SLIS Colloquium

Series, School of Library and Information Science, Indiana University-Bloomington. Retrieved August 20, 2003, from http://ella.slis.indiana.edu/~aloehrle/repsys.ppt.

Jacob, E. K., Mostafa, J., & Quiroga, L. M. (1997). An approach to the evaluation of automatically generated classification schemes. In P. Solomon (Ed.), *Advances in classification research, Vol. 7. Proceedings of the 7th ASIS SIG/CR Classification Workshop: Held at the 59th ASIS Annual Meeting, Baltimore, MD, October 20, 1996* (pp. 78–98). Medford, NJ: Information Today.

Library of Congress. Cataloging Policy and Support Office. (2002). *Library of Congress Subject Headings* (25th ed.). Washington, DC: Library of Congress, Cataloging Distribution Service.

Markman, E. M. (1989). *Categorization and naming in children: Problems of induction.* Cambridge, MA: MIT Press.

Mayr, E. (1982). *The growth of biological thought: Diversity, evolution, and inheritance.* Cambridge, MA: Harvard University Press.

McCloskey, M. E., & Glucksberg, S. (1978). Natural categories: Well defined or fuzzy sets? *Memory and Cognition, 6*(4), 462–472.

Rips, L. J., Shoben, E. J., & Smith, E. E. (1973). Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior, 12,* 1–20.

Rosch, E. (1973). Natural categories. *Cognitive Psychology, 4*(3), 328–350.

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General, 104,* 192–233.

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

Rosch, E., Mervis, C. B., Gray, W., Johnson, D., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8*(4), 382–439.

Shera, J. H. (1951/1965). Classification as the basis of bibliographic organization. In *Libraries and the organization of knowledge* (pp. 77–96). Hamden, CT: Archon. (Reprinted from *Bibliographic organization,* pp. 72–93, 1951, Chicago, IL: University of Chicago Press)

Shera, J. H. (1953/1965). Classification: Current functions and applications to the subject analysis of library materials. In *Libraries and the organization of knowledge* (pp. 97–111). Hamden, CT: Archon. (Reprinted from *The subject analysis of library materials,* pp. 29–42, by M. F. Tauber, Ed., 1953, New York: Columbia School of Library Science)

Shera, J. H. (1956/1965). Putting knowledge to work. In *Libraries and the organization of knowledge* (pp. 51–62). Hamden, CT: Archon. (Reprinted from *Special Libraries, 47,* pp. 322–326, 1956)

Shera, J. H. (1960/1965). What lies ahead in classification. In *Libraries and the organization of knowledge* (pp. 129–142). Hamden, CT: Archon. (Reprinted from *Proceedings of the Allerton Park Institute,* pp. 116–128, 1960, Champaign: University of Illinois Bookstore)

Smith, E., & Medin, D. (1981). *Categories and concepts.* Cambridge, MA: Harvard University Press.

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

Taylor, J. R. (1989). *Linguistic categorization: Prototypes in linguistic theory.* Oxford: Clarendon Press.

Zerubavel, E. (1993). *The fine line: Making distinctions in everyday life.* Chicago, IL: University of Chicago Press.