Kwasnik, B. (1992). The Role of Classification Structures in Reflecting and Building Theory. 3rd ASIS SIG/CR Classification Research Workshop, 63-82. doi:10.7152/acro.v3i1.12597

PROCEEDINGS OF THE 3rd ASIS SIG/CR CLASSIFICATION RESEARCH WORKSHOP

The Role of Classification Structures in Reflecting and Building Theory

Barbara H. Kwasnik
School of Information Studies
Syracuse University
Syracuse, NY 13244
bkwasnik@suvm.acs.syr.edu

INTRODUCTION

In general, we can say that the role of a classification scheme is to provide a descriptive and explanatory framework for ideas and a structure for the relationships among the ideas. We create classificatory schemes to organize our knowledge of the world in such a way as to be useful in communicating and using this knowledge. It follows then, that a felicitous classification scheme has embedded within its content and structure a great deal of information — not only about the entities themselves (representing individual phenomena and concepts) but also about the relationships among these phenomena and concepts, that is, how these things *go* with respect to each other. As such, classifications are really very much like theories. Like theories, classification schemes can provide an explanatory shell for looking at the world from a contextually determined perspective. Classification schemes not only *reflect* knowledge by being based on theory and displaying it in a useful way (as, for example, in the phylogenetic tree based on Darwinian theory), but also classifications *in themselves* function as theories do and serve a similar role in inquiry: that is, the role of explanation, parsimonious and elegant description, and the generation of new knowledge.

In this paper I examine the strong relationship of theories and classification schemes. Two classification schemes: The DSM Classification (for mental disorders) and the Periodic Table of Elements are offered as two examples of this relationship. Next, I examine three classification structures and their properties: hierarchies, trees, and faceted classifications as examples of how classificatory structure and theory interact.

THE RELATIONSHIP OF THEORY AND CLASSIFICATION

Abraham Kaplan describes theories as the "symbolic dimension of experience as opposed to the apprehension of brute fact" [Kaplan, 1963:294]. When we build theory, he says, we are making sense of a disturbing situation. A classification scheme functions in a similar manner by first, identifying the phenomena of interest naming them; second, partitioning these phenomena into meaningful clusters following systematic rules for discrimination and inclusion; and third, by constructing a structure that reflects a framework of relationships among the phenomena. So, just as a theory "explains" a particular fact by relating it to other facts [Kaplan, 1963:297], a classification strives to represent knowledge about entities by relating them to other entities, and just like a theory is not just "economy of thought" or a "mental shorthand," but rather the introduction of order into a congeries of fact [ibid., 302], a classification also not only a way of representing entities, but rather, a way of imposing order on them.

ISSN: 2324-9773

Kwasnik, B. (1992). The Role of Classification Structures in Reflecting and Building Theory. 3rd ASIS SIG/CR Classification Research Workshop, 63-82. doi:10.7152/acro.v3i1.12597

The Diagnostic and Statistical Manual of Mental Disorders (DSM)

The field of psychology has attempted to classify mental disorders, with what seems to me to be singular lack of success. The early attempts were "descriptive," classifying disorders according to behavioral symptoms, such as disorientation, alcohol abuse, and sleeplessness. The mandate for this classification was politically and economically motivated: government agencies, insurance companies, benefits programs, and others wanted to be able to differentiate and "tag" patients with mental disorders unambiguously for the purpose of reimbursement, legal action, confinements and so on. As a practitioner put it: "The classification assumes each person responds similarly for similar conditions — it considers the individual without considering individuality."

Early versions (DSMI and DSMII) were extremely impoverished in terms of reflecting the complexities of human behavior, such as environment, past history and societal norms. It also ignored much of psychological theory in terms of the underlying motivations and causes of mental disorders. Practitioners hated it. They were unable to "fit" clinical cases into the arbitrary categories. The scheme did not aid diagnosis or treatment — in fact, it constrained it. Furthermore, the scheme did not bring together knowledge and insight in such a way as to make further knowledge and insight easier.

Structural changes were introduced in the third edition and the revised third edition (DSMIII and DSMIII-R) included the addition of a multiaxial scheme. There are five axes, each representing an important aspect of diagnosis. Each axis (or we might think of it as a facet) follows its own internal logic of organization. Figure 1 shows an example:

EXAMPLE 1		
Axis I:	296.23	Major Depression, Single Episode, with Melancholia
	303.93	Alcohol Dependence, In Remission
Axis II:	301.60	Dependent Personality Disorder (Provisional, rule out Borderline Personality Disorder)
Axis III:	Alcoholic cirrhosis of liver	
Axis IV:	Psychosocial stressors: anticipated retirement and change in residence with loss of contact with friends Severity: 4—Moderate	
Axis V:	Highest level of adaptive functioning past year: 3—Good	
EXAMPLE 2		
Axis I:	304.03	Heroin Dependence, In Remission
Axis II:	301.70	Antisocial Personality Disorder (Principal diagnosis); prominent paranoid traits
Axis III:	None	
Axis IV:	Psychosocial stressors: No information Severity: 0—Unspecified	
Axis V:	Highest level of adaptive functioning past year: 5—Poor	

Fig. 1. An Example of an Multiaxial Evaluation [from DSMIII, p.30]

Kwasnik, B. (1992). The Role of Classification Structures in Reflecting and Building Theory. 3rd ASIS SIG/CR Classification Research Workshop, 63-82. doi:10.7152/acro.v3i1.12597

The three main axes are:

Axis I - Clinical Syndromes (such as "Depression")
Axis II- Personality Disorders and Specific Developmental Disorders
Axis III - Physical Disorders (that bear on the mental disorder)

In addition the following two supplemental axes provide additional information "useful for planning treatment nd predicting outcome" [p.23]:

Axis IV - Severity of Psychosocial Stressors (such as loss of job)
Axis V - Highest Level of Adaptive Functioning in Past Year

Axes I and II follow an enumerative, partially hierarchical structure, similar to that used in medicine. Axis III follows accepted international disease classification terminology. Axis IV measures severity of psychosocial stressors on a seven-point scale ranging from "None" to "Catastrophic," and an eighth position for "Unspecified." Axis V measures functioning on a five-point scale ranging from "Superior" to "Poor." Figure 2 shows an example from the classification of Axis 1: Clinical Syndromes [DSMIII, p. 18].

The "classes" are based on more-or-less consensual labels, the organization more or less reflecting the Western paradigm of mental disorders being like other illnesses in their presentation of observable behaviors (symptoms). The Introduction to DSMIII, for instance, explains how this manual is meant to serve as a:

1. common language for clinicians and research investigators
2. a tool for accurate diagnostic assessment for treatment, and
3. a standard vocabulary for patients groups being compared in various treatment modalities

However, in trying to serve a broad spectrum of practitioners who represent a wide variety of theoretical perspectives, the DSM is, by its own admission, atheoretical in terms of the etiology of mental disorders. That is, it does not support or propose any one particular theoretical framework for understanding mental disease. As a result, practitioners do not use this tool to build further knowledge. In the structure, the lack of specified relations among the classes in Axes I and II and the lack of suggested interrelations among the five Axes allows a practitioner little opportunity for making any inferences or analogies and consequently restricts any diagnosis made on the basis of this classification to a rather shallow description.

Structurally, the DSM fails in its lack of internal consistency and coherence. Most categories have a "NOS" ("Not further specified") category that acts as a catchall and is frequently used when a disorder does not fit easily into another existing category. From a theoretical standpoint, at the very least, such a lack of comprehensiveness compromises a great deal of descriptive (and certainly explanatory) strength that a classification may provide. That is, this classification does not adequately "handle" all entities of interest to the practitioner.

Practitioners report that the DSM is cumbersome to use, awkward to learn, and does not reflect the way they really think or talk about their patients and the patients' illnesses. Colby [1982:12] suggests that this is the case because "in treatment, a clinician uses his fact space, not the official diagnosis space." That is, a clinician takes account of a global picture that changes over time and presents itself in a different way in different situations, even for the same patient — all aspects of diagnosis and treatment that are not present in the DSM. Furthermore, the DSM is criticized for

ISSN: 2324-9773

Kwasnik, B. (1992). The Role of Classification Structures in Reflecting and Building Theory. 3rd ASIS SIG/CR Classification Research Workshop, 63-82. doi:10.7152/acro.v3i1.12597

Other specific affective disorders
301.13 Cyclothymic disorder
300.40 Dysthymic disorder (or Depressive neurosis)

Atypical affective disorders
296.70 Atypical bipolar disorder
296.82 Atypical depression

ANXIETY DISORDERS

Phobic disorders (or Phobic neuroses)
300.21 Agoraphobia with panic attacks
300.22 Agoraphobia without panic attacks
300.23 Social phobia
300.29 Simple phobia

Anxiety states (or Anxiety neuroses)
300.01 Panic disorder
300.02 Generalized anxiety disorder
300.30 Obsessive compulsive disorder (or Obsessive compulsive neurosis)

Post-traumatic stress disorder
308.30 acute
309.81 chronic or delayed

300.00 Atypical anxiety disorder

SOMATOFORM DISORDERS
300.81 Somatization disorder
300.11 Conversion disorder (or Hysterical neurosis, conversion type)
307.80 Psychogenic pain disorder
300.70 Hypochondriasis (or Hypochondriacal neurosis)
300.70 Atypical somatoform disorder (300.71)

DISSOCIATIVE DISORDERS (OR HYSTERICAL NEUROSES, DISSOCIATIVE TYPE)
300.12 Psychogenic amnesia
300.13 Psychogenic fugue
300.14 Multiple personality
300.60 Depersonalization disorder (or Depersonalization neurosis)
300.15 Atypical dissociative disorder

PSYCHOSEXUAL DISORDERS
Gender identity disorders
Indicate sexual history in the fifth digit of Transsexualism code: 1 = asexual, 2 = homosexual, 3 = heterosexual, 0 = unspecified.

302.5x Transsexualism, _____
302.60 Gender identity disorder of childhood
302.85 Atypical gender identity disorder

Paraphilias

302.81 Fetishism
302.30 Transvestism
302.10 Zoophilia
302.20 Pedophilia
302.40 Exhibitionism
302.82 Voyeurism
302.83 Sexual masochism
302.84 Sexual sadism
302.90 Atypical paraphilia

Psychosexual dysfunctions

302.71 Inhibited sexual desire
302.72 Inhibited sexual excitement
302.73 Inhibited female orgasm
302.74 Inhibited male orgasm
302.75 Premature ejaculation
302.76 Functional dyspareunia
306.51 Functional vaginismus
302.70 Atypical psychosexual dysfunction

Other psychosexual disorders

302.00 Ego-dystonic homosexuality
302.89 Psychosexual disorder not elsewhere classified

FACTITIOUS DISORDERS

300.16 Factitious disorder with psychological symptoms
301.51 Chronic factitious disorder with physical symptoms
300.19 Atypical factitious disorder with physical symptoms

DISORDERS OF IMPULSE CONTROL NOT ELSEWHERE CLASSIFIED

312.31 Pathological gambling
312.32 Kleptomania
312.33 Pyromania
312.34 Intermittent explosive disorder
312.35 Isolated explosive disorder
312.39 Atypical impulse control disorder

Figure 2. A section of Axis 1

Kwasnik, B. (1992). The Role of Classification Structures in Reflecting and Building Theory. 3rd ASIS SIG/CR Classification Research Workshop, 63-82. doi:10.7152/acro.v3i1.12597

being biased unfairly with respect to women, homosexuals, minorities, children and the elderly. These biases are reflected not only in Axes I and II but especially in Axis III (Psychosocial Stressors). It can be argued, then, that DSM does not even do a good job of *reflecting* existing knowledge, let alone being a tool for shaping and creating new knowledge.

In summary, DSMI through DSMIII-R, and the soon to be published DSM-IV, in bypassing theory for the practical expedient of pigeonholing disorders, have effectively constrained any heuristic or explanatory function a more theory-based classification may have afforded. The realities of the profession, however, along with requirements of the law and administrators, and the paradigm that governs the general view of mental disorders in general, make a stronger tool difficult to achieve.

The Periodic Table of Elements

By contrast, the Periodic Table of Elements, commonly attributed to Mendeleyev, is an example of a classification that has served not only to describe the "reality" of the elements in a systematic, coherent and corroborated way, but has actually been of "inestimable value in the development of chemistry" [New Encyclopaedia Britannica, 1988:949].

During the 19th Century there was a rapid growth in knowledge about the elements, their properties and their behavior. The result was a vast amount of individual "facts" in need of classification in order to make some sense of them. These facts included such things as correlations of properties, understanding of atomic weights and valency. In other words, there was already a great deal of knowledge out there. In 1869, Mendeleyev proposed a table in which "the elements arranged according to the magnitude of atomic weights show a periodic change of properties" [New Encyclopaedia Britannica, 1988, p.950]. Speaking in classification terms, a partitioning of the entities of interest (elements) based on a salient factor (atomic weight) yielded a regular pattern of distinctions. Speaking in terms of theory, the table "...so divides its subject matter that it can enter into many and important true propositions about the subject matter..."[Kaplan, 1963: 50]. Other scientists were proposing similar classifications and the Table underwent several revisions and amplifications, but has remained relatively stable since then (see Figure 3). Clearly this was a felicitous structure. Not only did it reflect and meet with the consensus of contemporary scientists, but it has also endured

According to Kaplan, theory is a "symbolic construction which can provide vicarious experience never actually undergone." What this means is that theory has a heuristic function, not just in explaining facts already known, but also those as yet undiscovered — "...laws propagate when they are united in a theory: theory serves as a matchmaker, midwife, and godfather all in one." In generating new questions for investigation we say that theories are "fruitful." The Periodic Table functioned in this way. The particular juxtaposition of elements suggested by the Table, yielded new and corroborating evidence not only for the Table itself but also for new knowledge about the elements and about chemistry. For instance, it was later discovered that elements occurring in close proximity in rows and columns (such as the inert gases), shared certain properties. Furthermore, new elements were predicted based on the regularity of the knowledge about existing elements. Thus the classification enabled a clearer understanding not only of each element with respect to other elements, but also of the entire universe of elements.

ISSN: 2324-9773

Kwasnik, B. (1992). The Role of Classification Structures in Reflecting and Building Theory. 3rd ASIS SIG/CR Classification Research Workshop, 63-82. doi:10.7152/acro.v3i1.12597

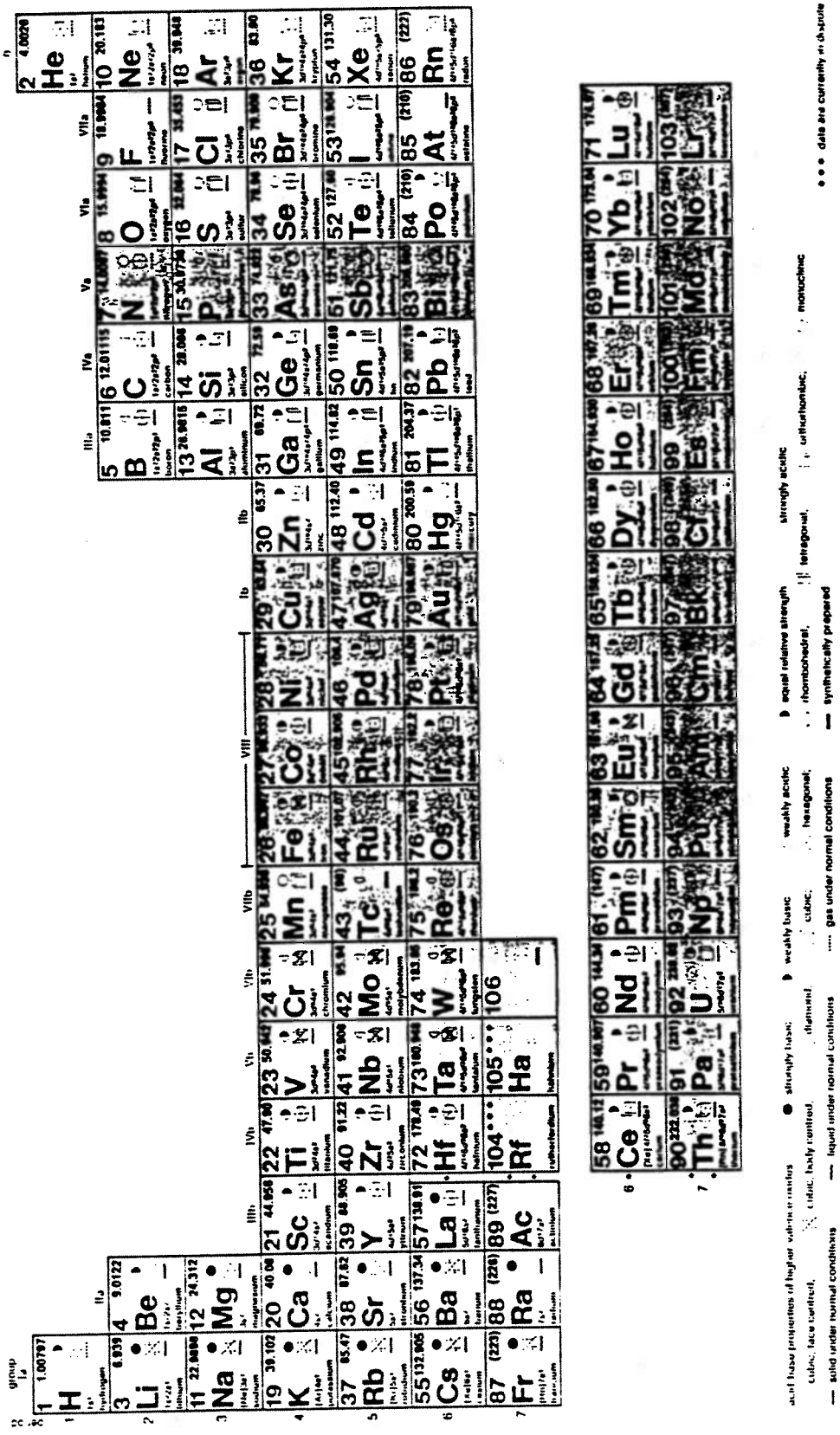

Fig. 3. The Periodic Table of Elements

ISSN: 2324-9773

Kwasnik, B. (1992). The Role of Classification Structures in Reflecting and Building Theory. 3rd ASIS SIG/CR Classification Research Workshop, 63-82. doi:10.7152/acro.v3i1.12597

The Periodic Table is a marvel of structure. It conforms to what is known and predicts that which is not known. It accommodates new knowledge and yet is able to comprehensively account for all existing knowledge. It is parsimonious and elegant — clearly and unambiguously expressing important and far-reaching general principles, yet simple enough for the interested layperson to understand. In this way it is a tool for science and for communication.

While it is perhaps not fair to compare the difficulties of classifying the complexities of human behavior with the challenge of classifying chemical elements, these two classification schemes demonstrate the strong relationship between theory and classification. In Chemistry, there seemed to be consensus about the "facts" at least and also, to a great extent, about the theory. In Psychology, there is general agreement about almost nothing at all. In Chemistry, there are agreed-upon instruments and techniques for precise measurement which are lacking in the social sciences. Consequently, in Chemistry there is a large body of knowledge with which to fill out a theoretically based classification. In Psychology, we are still debating about the measurement, the instruments, and even the entities themselves.

CLASSIFICATION STRUCTURES

"Theory is not the aggregate of the new laws but their connectedness, as a bridge consists of girders only in that the girders are joined together in a particular way" [Kaplan, 1963:297]. Similarly the entities incorporated into a classification reflect knowledge to the extent that they are arranged in a particular structure. In our culture, we have become accustomed to two ubiquitous classificatory structures: hierarchies and trees. In fact, a hierarchical structure or a tree is often considered synonymous with "classification" itself. While there are many other structures available to us: paradigms (matrices), networks, fuzzy sets, and "tangled" versions of both hierarchies and trees, in this paper I will deal with the two better-known structures and their properties.

Hierarchies.

In a hierarchy we arrange the names of the phenomena being classified into a structure in which the relationship among the items is one of class inclusion and attribute differences *(species/ differentia)*. This sort of structure has many advantages and strength as a tool in representing knowledge and creating new knowledge. For instance, we know that a Manx cat is like all cats in certain important ways, and that all cats are like all felines in certain important ways. We also know that cats, leopards and lions have some features in common and also that they differ in some *predictable way*, one from the other. If, then, we are faced with classing a new entity, say, an unfamiliar fossil, we look for the place in the classification scheme it fits best, using the rules for inclusion and differentiation. Having found this place, we now have a great deal more information about that fossil which we learn from the information embedded in the structure of the classification scheme and which may not have been apparent from the artifact itself. That is, we can make some inferences, and we can fill in the details of description that may not be immediately evident.

For instance, if we establish that an newly discovered creature is a kind of "cat," we then also know that this creature must bear live young, and that it breast-feeds its young. We know this from information about all cats, which is in turn, information about all felines, which is in turn,

ISSN: 2324-9773

Kwasnik, B. (1992). The Role of Classification Structures in Reflecting and Building Theory. 3rd ASIS SIG/CR Classification Research Workshop, 63-82. doi:10.7152/acro.v3i1.12597

information about all mammals, and so on. These resemblances discovered from the entity's position in the classification may be of more theoretical significance than the originally observed resemblances. In this way, a hierarchy that reflects theory well helps inquiry, corroborates theory and generates new knowledge.

If a new entity refuses to be easily classified within such a classification (as has been the case with viruses, for instance), a clear signal is given that either the observation is inaccurate, or that the classification scheme itself (and therefore also its underlying theory) needs reassessment. In this way a strong classification helps to identify gaps in knowledge and to test theoretical explanations of existing facts.

The problem with this structure is that it may offer an incomplete representation, that is, it may not always accommodate a representation of *other aspects* of the entities that are being classed. For instance, the example of the "mammals—felines—cats" type of classification, mentioned above, does not differentiate among creatures with different temperaments or in different habitats (such as zoos) or from different time periods (such as prehistory). Furthermore, a scientific classification of cats requires that we adhere to the principles of mutual exclusivity and, therefore, does not permit allocation of entities to more than one category — for instance the fact that a cat is a feline, but it may also be a pet, or an icon in a work of art. And finally, such a structure may be inhospitable to new knowledge. If we learned that, for example, leopards were more like dogs in some important, newly discovered way (such as a molecular component of the DNA) we would not be able to amend this classification without disrupting its basic structure and without redefining the rules of inclusion and differentiation.

Trees

The other traditional classificatory structure is that of the tree. Entities classified in a tree are divided into categories based on one dimension of differentiation at a time. The structure displays the phenomena of interest in such a way as to *make explicit the patterns of some salient characteristics*. For instance, in the classical phylogenetic tree, animals are divided into two groups: Vertebrates and Invertebrates, based on the presence or absence of a spine. A tree based on this particular rule of division is very good at displaying the distribution of the entities according to this criterion.

In an excerpt (adapted from Merritt, 1991) of a classification of "vegetarians," the categories are formed for various kinds of eaters by determining whether or not they eat or refrain from eating certain kinds of foods, such as meats, dairy products, eggs, foods produced in soils fertilized by animal wastes, and so on. It is one way of describing and distinguishing eaters of all kinds. The choice of which criterion to apply first, second, and so on, is determined by the importance of that criterion in discriminating between one set of entities and another. In the case of classifying vegetarians, the consumption of foods from animal sources is considered the most important, or "first cut" criterion.

ISSN: 2324-9773

Kwasnik, B. (1992). The Role of Classification Structures in Reflecting and Building Theory. 3rd ASIS SIG/CR Classification Research Workshop, 63-82. doi:10.7152/acro.v3i1.12597

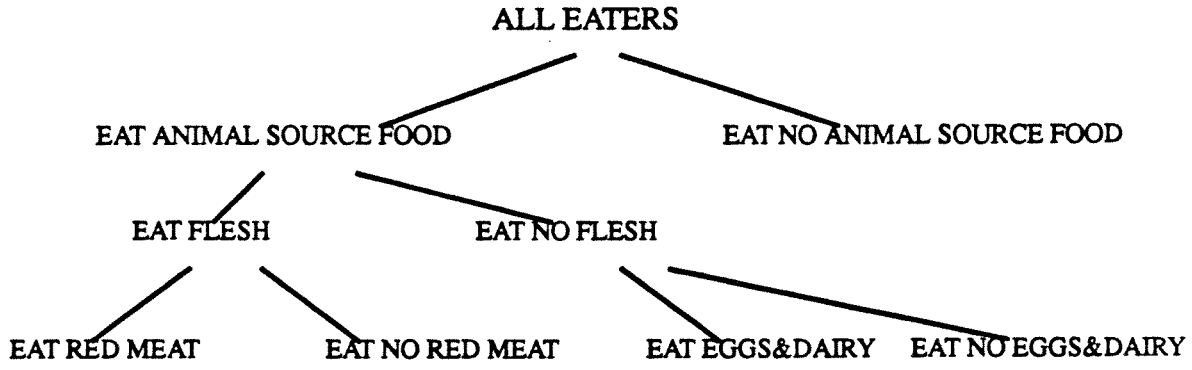

Fig. 4. A Tree Display of Eaters

We can see that such a tree is successful at presenting very clearly the pattern of *abstinence or consumption*, which is an important criterion for categorizing eaters. At the same time, we may also know that there may be another dimension that is important for distinguishing among eaters, but that is not represented in this scheme. For instance, a person may have different *motivations* for not eating flesh. Thus we may have a category of people who refrain from eating meat because they abhor the thought of killing another animal, and a category of people who refrain from eating meat because they think it is not good for their health. That is, the same action can be caused by two different motivations. In order to accommodate the various other dimensions into this scheme, the display would have to become fairly complex and extensive. At each node, for instance, we might have to further divide by motivation, by cultural influence, and so on. In addition, besides the added complexity that could occur at each node, the author of this scheme pointed out (only half-seriously) that there is another dimension that is superimposed from left to right over the entire scheme — that of "increasing sanctimoniousness of the eater." What that means is that there is an important dimension that is accommodated neither in the nodes or their links, but rather in the overall scheme. Such a dimension is difficult to represent in a tree.

Both hierarchies and trees are effective when we have sufficient knowledge about a given domain to be able to class the entities gracefully and effectively. In hierarchies, we must have sufficient knowledge about the entities, the rules for differentiation and class inclusion, and the relationship among the classes; in trees, we must know the important criteria by which to divide the entities into classes and the right order in which to invoke the criteria. In domains that are guided by theory and relative stability and consensus, such classificatory schemes not only reflect the salient aspects of the existing state of knowledge, but also help to build further theoretical understanding. In new and emerging domains, however, or when we have incomplete knowledge, or when the field is likely to undergo rapid change in the state of knowledge, the building of hierarchies and trees that are useful and that resonate with the state of affairs is extremely problematic.

ISSN: 2324-9773

Kwasnik, B. (1992). The Role of Classification Structures in Reflecting and Building Theory. 3rd ASIS SIG/CR Classification Research Workshop, 63-82. doi:10.7152/acro.v3i1.12597

FACET ANALYSIS

While hierarchies and trees are desirable for their strong descriptive and predictive properties, it is frequently the case that we have, like Mendeleyev, a large and ever-increasing body of facts with a need for creating some order out of chaos, but, unlike Mendeleyev, we are lacking that one salient characteristic, that "atomic weight" as it were, that would allow us to build so beautiful and useful a classification as the Periodic Table. Or, we are lacking the theoretical glue that would allow us to build a coherent and stable framework which we might then fill in with the facts. Moreover, we now realize that *all* classifications miss something or are unable to capture some aspect of the phenomena of interest, or in attempting to do so, become so complex as to defy use.

Ranganathan's contribution to classification theory is not only his innovative principles, canons, and techniques for notation, but also his acknowledgment that all classification is tentative in nature, that is, he developed his scheme with the understanding that there is no *one* way to view the world. He recognized, of course, that there were canonical, or traditional, ways of viewing the universe of knowledge, and that these ways were useful until they no longer reflected consensus, but, according to his view, all basic categories and their subdivisions were provisional and constantly being changed in a dynamic way by new ways of combining and modifying them. Thus, while he talks of basic subjects such as Mathematics, and distilled subjects, such as Research or Management, he also posits that such divisions are provisional. Over time, nontraditional subjects (such as Library Science) can become traditional subjects; compound subjects (such as Biophysics) can become basic subjects and so on. Categories can change their status and their relationship to each other as knowledge and needs change and evolve.

In order to manage such complexity and change, however, it is necessary to find some systematic way of creating order. Ranganathan's "fundamental categories" can be viewed as such a framework. These fundamental categories: Personality, Matter, Energy, Space, and Time (PMEST for short) are, according to his view, inherent in all complex phenomena. That is, any complex entity can be viewed from the point of view of its personality, its relation in space and time, and so on. These fundamental categories are the basic building blocks into which any phenomenon can be analyzed. They are commonly referred to as facets. To be precise, Ranganathan's definition of a facet was quite a bit broader:

> A generic term used to denote any component — be it a basic subject or an isolate — of a Compound Subject, and also its respective ranked forms, terms, and numbers (Ranganathan, 1967, p.88).

It is the notion of analyzing a phenomenon in terms of fundamental aspects and then resynthesizing it into a useful expression, however, that has come to be known as **facet analysis**.

Ranganathan's Fundamental Categories

Briefly summarized, the fundamental categories, as expressed by Ranganathan, are traditionally arranged from most concrete (Personality) to the most abstract (Time). They are presented here in reverse order (from Satija, 1989):

ISSN: 2324-9773

Kwasnik, B. (1992). The Role of Classification Structures in Reflecting and Building Theory. 3rd ASIS SIG/CR Classification Research Workshop, 63-82. doi:10.7152/acro.v3i1.12597

- **Time.** The chronological factor in a subject, e.g., 20th Century poetry, Medieval science;
- **Space.** The manifestation of geographical areas, e.g. Central Europe, London, hilly areas;
- **Energy.** Action activities of all kinds, e.g. editing, dissemination, extraction, welding;
- **Matter.** Covers three subcategories; Matter-Property (fundamental or inherent properties such as inertia, public health, occurrence), Matter-Method (various methods of actions such as processes, increase, decrease) and Matter-Material (matter *qua* matter, such as, hydrogen, carbon).
- **Personality.** It is what is left over when the former four categories have been analyzed.

Out of context, it is rather difficult to ascertain unambiguously what each fundamental category describes. From the start, Ranganathan emphasized that the analysis proceeds in a given domain and under the aegis of that domain's assumptions and norms. Still, these fundamental categories have undergone considerable debate, and the definition of what belongs in each category has shifted somewhat over time in order to clarify procedures for applying the scheme in practice. However, the basic notion of these five as being sufficient to describe and effectively analyze complex concepts has withstood the test of time. Ranganathan's *approach* to classification — that is, the underlying philosophy and the techniques of facet analysis have been extended to many diverse areas of application.

Facet Analysis in Practice

It is important to note that the process introduced by Ranganathan is referred to as "analysis" and not as "decomposition" or "division." To divide something into its components is to break it down into entities that are each in and of themselves different than the whole. Thus, the *components* of a pie are flour, fruit, shortening, water, and so forth. To *subdivide* something is to create ever more specifically differentiated categories: *subdivisions* of fruit are pears, apricots and kiwis; pears are further subdivided into bosc and anjou pears. To *analyze* something, on the other hand, in the sense it is being used here, is to view it from all angles — the same entity, but with emphasis on a different dimension or facet. Thus, the *facets* of a pie may be its ingredients, taste, process of cooking, nutritional value, aesthetic appeal, and so on. Each of these facets may be expressed following the rules specific to it. Thus, "taste" may be a facet that has as its possibilities: "sweet," "sour," "...." "salty," etc., whereas "process of cooking" may have as its possibilities "baking," "frying," "boiling," "...," "roasting," etc. The former facet deals with attributes, the latter with processes. The advantage is that in trying to classify some phenomenon such as a pie, it is not necessary in facet analysis to choose only one way, only one scheme, only one logic of division and aggregation. What is necessary is to choose the appropriate facets and to develop the facets with schemes appropriate to them.

Following a previous example, it is possible, therefore, by facet analysis to create a richer, more complex classification of eaters. Each eater can be described by a number of facets: e.g., consumption/abstinence; motivation; sanctimoniousness; cultural determinants, and so on. The "consumption/abstinence" facet can be displayed effectively as a tree, while the

ISSN: 2324-9773

Kwasnik, B. (1992). The Role of Classification Structures in Reflecting and Building Theory. 3rd ASIS SIG/CR Classification Research Workshop, 63-82. doi:10.7152/acro.v3i1.12597

"sanctimoniousness" facet might better be described by a scale of high to low. Such a classification loses some heuristic properties but is very flexible because it admits new, yet unknown kinds of eaters.

One might say that the five axes of the DSM are facets, and considering the state of theoretical understanding and flux in the field, this is probably appropriate. But then, one might question the appropriateness of the way in which each axis itself is organized, or even of the appropriateness of the choice of those particular five axes. That is, the choice of a faceted analysis for each mental disorder may be the right choice of classificatory structure, but perhaps different facets should be chosen, and the ones that are retained, say, Axis 4, Psychosocial Stressors, be structured in such a way as to better reflect understanding of what a stressor is and how it interacts with mental dysfunction.

THREE EXAMPLES OF FACETED ANALYSIS

As mentioned previously, when we are confronted with creating a classification scheme for entities and domains that are new, or about which we have insufficient theoretical knowledge, or which are complex in terms of the number of salient dimensions, or which appear in evolving areas, facet analysis is a good candidate for a classificatory structure that is flexible, but which, if it is carefully constructed, can have coherence and integrity in terms of validly representing knowledge.

The Art and Architecture Thesaurus

The *Art and Architecture Thesaurus* (AAT) covers terminology for art and architecture of the Western world from antiquity to the present. It covers not only the terminology for the objects of our material culture but also the words used to describe these objects [AAT, 1989, pp. 25-26; Bearman & Petersen, 1991]. It is immediately obvious that the realm of material culture is vast and ever-changing, and that the way of looking at this culture differs from one person to another and one application to another. There are thousands of as-yet uncreated objects that might fit into the scope of this thesaurus. Thus, the faceted scheme that was devised seems especially appropriate.

The conceptual framework has been worked out to include seven facets and forty hierarchies. Thus far, twenty-three have been worked out, and work continues on the rest. These are shown in Figure 5.

Kwasnik, B. (1992). The Role of Classification Structures in Reflecting and Building Theory. 3rd ASIS SIG/CR Classification Research Workshop, 63-82. doi:10.7152/acro.v3i1.12597

ASSOCIATED CONCEPTS FACET
- Associated Concepts

PHYSICAL ATTRIBUTES FACET
- Design Attributes
- Design Elements
- Colors

STYLES AND PERIODS FACET
- Styles and Periods

AGENTS FACET
- People and Organizations

ACTIVITIES FACET
- Disciplines
- Functions
- Events
- Processes and Techniques

MATERIALS FACET
- Materials

OBJECTS FACET
- Built Environment
  - Settlements, Systems and Landscapes
  - Built Complexes and Districts
  - Single Built Works and Open Spaces
  - Building Divisions and Site Elements
  - Built Work Components
- Furnishings and Equipment
  - Tools and Equipment
  - Hardware and Joints
  - Furniture
- Visual and Verbal Communication
  - Image and Object Genres
  - Drawings
  - Photographs
  - Document Types

Fig. 5 The Seven Facets of the Conceptual Framework of the Art and Architecture Thesaurus

Within each facet, the further subdivisions are arranged each following its own logic. For instance, subdivisions with STYLES AND PERIODS are arranged chronologically, Tools and Equipment in the OBJECTS facet are arranged by function, and so on. An object can be described by one descriptor chosen from the schedules, by two or more in combination (a modified descriptor), or by a syntactic construction of several descriptors.

In building the syntactic constructions (called strings), a focus term is chosen. Only one focus term is allowed per string. Modifiers are chosen from other facets and are arranged preceding the focus

ISSN: 2324-9773

Kwasnik, B. (1992). The Role of Classification Structures in Reflecting and Building Theory. 3rd ASIS SIG/CR Classification Research Workshop, 63-82. doi:10.7152/acro.v3i1.12597

term (as in natural language) in the order presented above. That is, a modifier describing styles and periods comes before a modifier describing materials. Examples of analyzed expressions (A&AT, Guide for Use, p. 57)are:

> asymmetrical blue pressed glass serving bowls
> large Baroque hunting lodges

In each case the focus term is underlined. The facet order assures unambiguous strings and provides an almost infinite array of possible expressions.

The difficulty of establishing the basic seven facets should not be underestimated, however. Clearly, Ranganathan's principles, although not slavishly followed, have been used as the underlying structure for this faceted scheme.

Software Reuse

The domain described in this section is software reuse [Prieto-Diaz and Jones, 1990]. In software engineering, the need for reusing software components is a growing problem. In order for these components to be retrievable from a software library, they must be described and classified in such a way that a programmer can find what he or she needs.

In the system described by Prieto-Diaz and Jones, the terms and concepts for the classification are gathered over time from users and the documents they generate. This process is continued even after the system is in place as a way of updating and keeping current with new developments. A classical classification scheme was rejected because, in this environment, a given software module might belong to many categories. Searching detailed trees or hierarchies would also require extremely specific knowledge on the part of the librarian in charge of storing and retrieving the modules.

Instead a faceted classification scheme was devised, consisting of six facets. The first three describe program function, the second three describe program environment.

> The faceted approach offers a very attractive method for classifying reusable software. In addition to tailoring classifications to the specific subject, facets may be ordered by their relevance to the users of the collection. For example, the classification scheme for Structured Systems Programming could list 'activity' as the first facet and 'entities" as the second....This feature enhances search and retrieval performance when used to organize a database [Prieto-Diaz & Jones, 1990, p. 157].

Figure 6 shows a partial listing of facets designed for a collection of 200 small, general-purpose software components

ISSN: 2324-9773

Kwasnik, B. (1992). The Role of Classification Structures in Reflecting and Building Theory. 3rd ASIS SIG/CR Classification Research Workshop, 63-82. doi:10.7152/acro.v3i1.12597

FACETS

FUNCTION - e.g., add, append, close, compare, complement, compress,...join...etc.
OBJECTS - e.g., arguments, arrays, backspaces, blanks, buffers,... lists...etc.
MEDIUM - e.g., array, buffer, cards, disk, file, keyboard, line, list, mouse,...tape...etc.
SYSTEM-TYPE - e.g., assembler, code-generation, compiler,...line-editor...etc.
FUNCTIONAL-AREA - e.g., accounts-payable, auditing, customer-information...etc.
SETTING - e.g., advertising, aircraft-manufacture, appliance-store,...cleaning...etc.

Fig. 6. Facets for Classifying Multi-Purpose Software Components

In some respects, this is a much simpler application than that used in the A&AT and it does not conform as closely to Ranganathan's PMEST categories. For one thing, the classification is not used as a thesaurus and does not attempt to build in a specific citation order to the facets. Secondly, it does not seem that within each facet there is a development of internal structure for that facet. The terms are arranged alphabetically. Nevertheless, the scheme offers a "building-blocks" approach to analyzing software components and thereby offers a more flexible way of later retrieving these components. Although not mentioned explicitly in this article, it also offers the indexer a framework within which to consider each new entity being added to the system.

Use of Facet Analysis in Describing Research Results

The final example comes from the classification problems inherent in summarizing and describing the results of research. Qualitative research in particular, where large amounts of interview protocols are gathered are especially problematic in this respect. The researcher is faced with a mass of data and must coherently and validly summarize that data to make some general statements about what she or he found. The situation is one in which there may be no existing framework upon which to build a hierarchy or a tree. Thus, in exploratory, descriptive studies, such as the one described here, the researcher using the facet analysis approach attempts to analyze the data into the most fundamental components and then resynthesize these components into observed patterns and statements that will abstractly, but specifically, explain the phenomenon under investigation.

The study described here, had as its goal the description of how people organize documents in their own offices [Kwasnik, 1989]. Eight university faculty members were asked to describe their own offices in terms of the organization of what each of them defined as documents. Each respondent was also asked to sort a day's mail, simulating as closely as possible the usual way in which this task was done, but "thinking out loud" and giving as much detail as possible. The outcome of data collection was eight transcripts of interviews and the thinking-out-loud protocols gathered from the respondents.

The analysis of the data showed that documents are identified and classification choices are made in situations that can be described by a variety of dimensions. In an attempt to extract these dimensions from the data, first, each instance in which a participant identified a "document" was marked. The modifying phrases they used were interpreted by the researcher and summarized by

ISSN: 2324-9773

Kwasnik, B. (1992). The Role of Classification Structures in Reflecting and Building Theory. 3rd ASIS SIG/CR Classification Research Workshop, 63-82. doi:10.7152/acro.v3i1.12597

brief terms and labels. The labels were defined into coding categories after multiple passes through the data, and a codebook was built up iteratively. Finally, the entire corpus was coded using these coding categories. Each coding category represented a dimension along which classificatory decisions were made.

For example, the following two instances of classificatory decisions:

> on the top shelf are books that are very seldom used
> correspondence I must deal with immediately goes into my briefcase

can be described by the same set of coding categories: LOCATION, FORM, TIME and USE.

on the top shelf	—LOCATION—	into my briefcase
books	—FORM—	correspondence
very seldom	—TIME—	immediately
used	—USE—	deal with

This analysis yielded an inventory of document labels and an inventory of dimensions. Once the entire corpus had been coded, it was possible to merge and rearrange the categories so that extremely fine levels of distinction that accounted for a very small proportion of data were collapsed into more inclusive categories.

The following is a list of the seven fundamental categories and their subdivisions. These categories "emerged" from the data through the inductive, constant-comparison method of analysis. Each dimension represents a possible criterion for the classification of documents in an office.

ISSN: 2324-9773

Kwasnik, B. (1992). The Role of Classification Structures in Reflecting and Building Theory. 3rd ASIS SIG/CR Classification Research Workshop, 63-82. doi:10.7152/acro.v3i1.12597

SITUATION ATTRIBUTES
- Access
- Circumstance
- Need/Requirement
- Ownership of the document
- Related to me
- Room/Space
- Source
- Use/Purpose

DOCUMENT ATTRIBUTES
- Author
- Form (book, file, notebook, etc.)
- Topic
- Physical Attribute (e.g., color, size, etc.)

DISPOSITION
- Change
- Discard
- Keep
- Locate
- Postpone decision

ORDER/SCHEME
- Accumulation
- Arrangement
- Group
- Separate
- Unfinished arrangement

TIME

VALUE
- Important
- Interesting
- Needs improvement
- Not valuable
- Secret/Confidential
- Unspecified value
- Works for me

COGNITIVE STATE (of person making the classificatory decision)
- Don't know
- Want to remember
- "Just know"

Fig. 7. Fundamental Categories and Their Subdivisions of Dimensions along Which People Classified Personal Documents

ISSN: 2324-9773

Each of the subfacets (e.g., USE/PURPOSE) can be further analyzed into more specific kinds of uses and purposes, such as, Use under Certain Circumstances, Use Defined by Time, Two Uses Together, and so on.

Once this classification was in place it was possible to study the data more comprehensively. For instance, it was possible to establish the frequency of the occurrence of each facet in the data. This revealed that SITUATION ATTRIBUTES were at least as important, if not more important in terms of frequency of citation, as DOCUMENT ATTRIBUTES. Furthermore, it was possible to describe patterns of criteria. Thus, for example, urgency (TIME) and closeness (LOCATION) are frequently paired.

The inductive "discovery" of the facets, through the filter of the researcher's analytical techniques, was an attempt to express the description of the classification of documents in the respondent's terms or categories rather than in terms dictated by a deductive model. For this purpose, the technique of facet analysis was invaluable. In fact, even though these categories were derived strictly from the data, without reference to Ranganathan's PMEST, it can be seen that the fundamental categories are not so different, at least in underlying principle, from those outlined by Ranganathan.

SUMMARY

The aim of classification is to provide a meaningful clustering of experience. We need to classify in order to impose some order on an overwhelming array of facts and knowledge and to be able to communicate about them. These aspects of classification are almost intuitively understood by most people. As a tool in inquiry, the role of classification is to provide coherent, useful, systematic, explanatory, heuristic, and theoretically sound representation of the entities and relationships among entities in a given domain. As such, classification is intimately bound to theory, which has similar functions. The extent to which a classification reflects and "nourishes" theory in a given domain determines its usefulness. Classifications have structural properties that lend themselves to representing knowledge in a given situation. The traditional trees and hierarchies are powerful, but are difficult structures to construct in domains that are not fully understood, changing, complex, and multivariate. Facet analysis proves to be a useful tool for building classifications in such circumstances and, if carefully and rigorously applied, allows for movement towards theoretical understanding of the domain it classifies.

REFERENCES

American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. Third Edition. Washington, D.C.: APA, 1980.

Art and Architecture Thesaurus. Guide for Use, 1990.

Bearman, David & Petersen, Toni. "Retrieval requirements of faceted thesauri in interactive information systems." In: Humphrey, Susanne M. & Kwasnik, Barbara H. (eds.) *Advances in Classification Research: Proceedings of the 1st ASIS SIG/CR Classification Research Workshop*. Medford, NJ: Learned Information, 1991:9-23.

ISSN: 2324-9773

Kwasnik, B. (1992). The Role of Classification Structures in Reflecting and Building Theory. 3rd ASIS SIG/CR Classification Research Workshop, 63-82. doi:10.7152/acro.v3i1.12597

Colby, Kenneth Mark. The Fundamental Crisis in Psychiatry. Springfield, IL: Charles C. Thomas, 1983.

Davies, Roy. The Creation of new knowledge by information retrieval and classification. *Journal of Documentation* 45(4) (Dec. 1989):273-301.

Kaplan, Abraham. *The Conduct of Inquiry: Methodology for Behavioral Science.* New York: Harper & Row, 1963.

Kwasnik, Barbara H. *The Influence of Context on Classificatory Behavior.* Ph.D. Thesis. School of Communication, Information and Library Studies. Rutgers. The State University of New Jersey, 1989.

Merritt, Nina. A classification of vegetarians. (Unpublished paper), 1991.

The New Encyclopaedia Britannica. Vol 15. "Periodic Law and Table." Chicago: Encyclopaedia Britannica, Inc., 1988.

Prieto-Diaz, Ruben and Gerald A. Jones. Breathing life into old software. In: Tracz, Will (Ed.), *Software Reuse: Emerging Technology* (Tutorial), Los Alamitos, CA: IEEE Computer Society Press, 1990.

Ranganathan, S.R. *Prologomena to Library Classification.* 3rd ed. Bombay, Asia Publishing House, 1967.

Satija, M.P. *Colon Classification: 7th Edition. A Practical Introduction.* New Delhi, India: Ess Publications, 1989.

ISSN: 2324-9773