

Sameness and Difference

A Cultural Foundation of Classification

Hope A. Olson

The idea of sameness is used to gather material in classifications. However, it is also used to separate what is different. Sameness and difference as guiding principles of classification seem obvious but are actually fundamental characteristics specifically related to Western culture. Sameness is not a singular factor, but has the potential to represent multiple characteristics or facets. This article explores the ramifications of which characteristics are used to define classifications and in what order. It explains the primacy of division by discipline, its origins in Western philosophy, and the cultural specificity that results. The Dewey Decimal Classification is used as an example throughout.

The duality of sameness and difference is an underlying principle of classification as we construct and practice it in Western culture. We try to group similar things together and separate them from things that are different. This principle is taught at an early age. In children's books and television shows, we learn to identify "which of these things is not like the other." In newspaper comic pages and activity books used in school or to keep children amused on long trips, we are given two nearly identical pictures and asked to find the details that are different. Once we learn to view the world in this manner, classification that groups similar things together seems to be an almost natural or innate way of organizing things. Indeed, for those of us who have been acculturated to identify sameness and difference, we find classification an extremely useful arrangement for browsing. It is so ingrained that we do not even think of it as a "real" way of finding information. It is not uncommon to hear people deprecating their searching skills by admitting that in a library they just find a call number and then browse the shelves. They take the classification for granted as though it were a natural landscape rather than a well-manicured lawn that is the product of intellectual labor.

Classification gathers things according to their commonalities. In doing so it demonstrates the effectiveness of this sameness/difference-principle duality. However, a large body of library literature suggests that classifications embody the biases most common in our culture. This literature, summarized by Olson and Schlegl (1999), documents bias in the placement of topics outside of mainstream North American and European culture and the omission of topics associated with marginalized groups. A. C. Foskett posits that this bias exists because classifications reflect the views and values of the classificationists who create them (1971; 1984). I suggest that a concentration on sameness can explain at least part of this bias and has been an unquestioned presumption of most classificationists. Two questions will help in exploring this possibility:

1. What is the same?
2. Which sameness takes precedence?

Hope A. Olson (hope.olson@ualberta.ca) is Associate Professor, School of Library and Information Studies, University of Alberta, Canada.

The author thanks Michael Taft for assistance in understanding the distinction made between literature and folk literature and Connie Winther for her assistance with the manuscript.

This paper builds on work presented at the session: One-size-fits-all Subject Access Systems: Tailoring General Schemes to Meet the Needs of Specific Communities of Searchers, sponsored by the Subject Access Committee of the Association for Library Collections and Technical Services at the ALA Annual Conference, Washington, D.C., June 27, 1998.

Manuscript received October 31, 2000; accepted for publication January 4, 2001.

In this article I will first examine the sameness/difference duality of classification in more depth, especially its cultural role in creating disciplines, our first level of classificatory gathering. I will then probe each of the two questions I have posed. Finally I will consider how we might look towards solutions, rejecting the idea of universal solutions and pointing in directions where some potential solutions might lie.

The Dewey Decimal Classification (DDC) is a useful example that I will use throughout this exploration of classification because of its familiarity to most of us; because of the way its notation reflects its structure making that structure more visible than in other classifications; and because Melvil Dewey, its creator, was articulate in his reasons for constructing the classification as he did. However, the same duality arises in other classification schemes as well, both within libraries and in a broader social and cultural context.

The Duality of Sameness and Difference

The sameness/difference duality has been with us at least since the ancient Greeks. Before Plato and Aristotle, Parmenides proposed the notion that any given thing either is or is not. It exists or does not exist. While Parmenides would not have envisioned anything resembling our present-day classifications as a result of this either/or choice, Aristotle did (for a fuller discussion see Olson 1999). Aristotle applied the notion of being or not being as something being or not being part of a category. So something either belongs to category X because it is in some way the same as other things in category X, or it does not belong because it is different. Sameness is the privileged factor in this pair because X defines sameness, while difference is defined negatively as not-X. Certainly there are other complexities in the construction of classification, but this duality is a requisite first principle for classification in our Western culture. For us, classifications, as their most fundamental function, gather things that are the same and separate things that are different.

As mentioned above, classifying by sameness as opposed to difference seems to fit the way we use information. However, it is at odds with other principles that we employ, and problems begin to appear when we look at what we are classifying. The body of information-bearing documents that we organize in libraries is not based on sameness and difference. Published documents tend to offer something new. We typically value novelty over redundancy. Publishers deciding what to publish and librarians selecting what to collect will look for new topics and fresh approaches to old topics. However, once we collect this innovative material we try to organize it by gathering what is the same.

Of course, our perspective is not really this naive. We

gather things that have one or more elements or facets in common. We build our classifications using these facets that bring things together according to some kinds of sameness. So a facet of time brings together things relating to the same chronological period. A facet of space brings together things relating to the same location. The things we gather are the same in limited ways as Dewey pointed out:

The skeme givs us for each topic, as it wer, a case of 9 pijeonholes, with a larj space at the top; and we uze them as every practical business man uzes such pijeonholes about his desk. . . . If [a businessman] insisted on having a different case made to order for each use, it wud cost over twice as much; he cud not group them together or interchanje them, and they wud not fit offis shelvs (Dewey, DDC13 1932).

Classifications such as the DDC set up pigeonholes for certain samenesses. Dewey himself recognized that dividing all of knowledge into tens is absurd in any theoretical sense but asserted that because it is efficient in practice—like the pigeonholes on a nineteenth-century businessman's desk—it is justifiable. Unfortunately in their efficiency, preformed pigeonholes do not conform to the shape of the information in our documents. Pigeonholes will accommodate some facets and not others—that is, they will give preference to some samenesses over others. S. R. Ranganathan in explaining his conception of faceted classification, echoed Dewey:

A characteristic used as the basis for the classification of a universe should differentiate some of its entities. . . . The characteristics relevant to the purpose of classification are usually many. Practical considerations, however, will restrict us to the inclusion of only a few of them . . . (1967, 146)

Topics may have large numbers of characteristics, facets, or samenesses, but it is not feasible to express all of them in classification. While a few are included, the remainder are excluded.

Disciplines—The Primary Facet in Classification

In the DDC, the nine main classes (those other than 0 Generalities) represent the primary facet of discipline. The introduction to the DDC states clearly that “the parts of the Classification are arranged by *discipline*, not by subject” (Dewey for Windows 1998). There has been much discussion about how these classes came to be considered the first level (or facet) and how they came to be in the

sequence in which we find them. Figure 1 suggests that the roots of the classes in the DDC, as an example, go deep into Western philosophy.

Looking at Dewey's main classes, we see an arrangement that came from William T. Harris's classification for the St. Louis public school library system that is allegedly a reverse order of Francis Bacon's classification of knowledge (Comaromi 1976). Bacon's Renaissance views came from Aristotle directly and via the medieval classification of knowledge for the sake of pedagogy into Trivium, Quadrivium, and Theology. Our current disciplines bear considerable resemblance to this medieval scheme. However, Bacon added a more philosophically sound ordering to them derived through a dialectical method (Jardine 1974). He viewed Memory as the basic store of experience that, when compared and contrasted via Imagination, can be processed through Reason into Knowledge. So Bacon's classification of Knowledge is based on History as an emanation of Memory, Poetry as an emanation of Imagination, and Philosophy as an emanation of Reason. Bacon, then, saw knowledge as the ultimate end of this development—it is basically an epistemological foundation for classification.

Since Harris was also a scholar of Georg Friedrich Wilhelm Hegel and Dewey himself was heavily influenced by Hegelians at Amherst College (Wiegand 1996), one might also compare the DDC main classes to Hegel's classification of knowledge. Hegel's Logic enumerated three categories of existence: Being (*Sein*), Essence (*Wesen*), and Idea (*Begriff*). Harris (1895) wrote an explication of Hegel's Logic in which he explained this three-part ontology. Harris describes Hegel's concept of Being as an illusory condition in which a person perceives reality as though it is only what it appears to be—it is the immediate perceived and requires no mediation. Hegel sees Being giving way to Essence—a disillusioned state in which a person recognizes that what appears is not independent reality but is the dependent results of various forces hidden behind the illusions of independent Being. So Essence is a condition in which the obvi-

ous is mediated by a skeptical understanding. Finally, Essence leads to the self-activity of Idea—a higher category that is the revelation of a higher being. Hegel describes plants, animals, and "man" as "imperfect Idea" (Harris 1895, 144–45). Idea transcends these forms to become a revelation of Hegel's theistic view of God. With this in mind, it is possible to see how there is a rough correlation (see figure 1) between History and Being as simple observation, the Fine and Mechanical Arts and Essence as an understanding of illusions, and the Sciences and Religion and Philosophy and Idea as states that are beyond and independent of both material being and illusion. Interestingly, Harris and other classificationists have put the highest state first rather than last. While he describes Hegel's ontological progression from Being to Essence to Idea, Harris reverses this sequence in his classification, placing the pinnacle of human understanding in a position of primacy at the beginning (much as the Soviet classification placed Marxism-Leninism first).

It is no surprise then that Dewey's approach is reflective of both the epistemological and ontological presumptions of Western philosophy. Taking Dewey's ten big pigeonholes as a beginning one can see in figure 1 how he interpreted Harris's interpretation of Bacon's epistemological dialectic and Hegel's ontological model of logic. This arrangement of disciplines sets the first level of sameness (and difference) in the DDC and is very similar in other classifications, notably the Library of Congress (LC) classification. Although we no longer equate philosophy and science or technology and poetry, we do still think of the world of knowledge as carved up into much the same categories as Dewey defined with his disciplinary pigeonholes, and these disciplines reflect the tradition from which they have grown.

Discipline—as the primary facet in our classifications—is the fundamental sameness. Within each discipline in a classification the subdivision reflects the discourse of specialists. Dewey depended upon specialists at Amherst College to help him devise the contents of each class. The current DDC follows disciplinary practice as in its recent major changes in the life sciences so that they are divided

| 0 Generalities | | | | | | | | | | |
|----------------|---------------------------|----------|-----------------|-------------|---------------------------|------------|----------|-------------------------|-----------------------|---------|
| DDC | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | |
| | Philosophy and Psychology | Religion | Social Sciences | Language | Natural Sciences and Math | Technology | The Arts | Literature and Rhetoric | Geography and History | |
| Bacon | Reason | | | Imagination | | | Memory | | | |
| | Philosophy | | | Poetry | | | History | | | |
| Hegel | Idea | | Imperfect Idea | | | | Essence | | Being | |
| Harris | Science (Philosophy) | | | | Art (Poetry) | | | | | History |

Figure 1. DDC and Western Philosophy

first by process rather than by organism. LC classification was originally structured by subject specialists and has a long tradition of literary warrant, reflecting the patterns of publication by specialists. The discourses of individual disciplines, then, determine the citation order—the internal structuring of classes according to priority amongst samenesses.

Clare Beghtol suggests that “increasing multidisciplinary knowledge creation makes it critical to reconsider the traditional reliance on discipline-based classification and to try to solve the problems that orientation has created” (1998, 2). The following discussion will demonstrate some of these problems as evidence that bias is linked to the DDC’s disciplinary structure and its discourses surrounding sameness and difference. The final section below will propose some directions for seeking solutions.

Two Problems in the Sameness/Difference Duality

What Is the Same?

How does the primacy of sameness play out in the current DDC? As an example, we can look at a seemingly universal discipline such as literature. In the DDC, the 800s are designated for literature. It is interesting, however, to look at the difference between literature and folk literature in the DDC. The definition, expressed in the following entry from the DDC manual, undoubtedly reflects dominant North American and European cultural values:

Manual Entry: 800 vs. 398.2 Literature (Belles-lettres) and rhetoric vs. Folk literature

Notes:

Folk literature consists of brief works in the oral tradition and is classed in 398.2. **Whatever literary individuality the folk literature may once have had has been lost** to the anonymity that the passage of time brings. Anonymous **classics**, however, are not considered to be folk literature. Despite the fact that their authorship is unknown, such works have a **recognized literary merit**, are almost always **lengthy**, and form a **part of the literary canon**. Therefore, they are classed in 800, e.g., *Chanson de Roland* 841.1, *Cantar de mio Cid* 861.1, *Kalevala* 894.54111. (*Dewey for Windows* 1998—emphasis added)

This definition implies that literature is created by individuals, not built up through the contributions of many people over generations. Here, DDC makes the same presumptions as other institutions, such as our intellectual property regimes, that individuals should be credited with

the production of texts even when they are drawn from long cultural traditions. Other cultures do not necessarily have these same presumptions (which are actually relatively recent even in dominant Western culture). Hence, DDC makes an exception for classics of “recognized literary merit” if they are part of the “literary canon.” The result is that a work like the *Kalevala*, a collection of anonymous fragments from an oral tradition drawn together by an individual in the nineteenth century, is treated as literature because it is long and a recognized classic while comparable uncollected fragments would be folk literature. The *Kalevala* is identified with Elias Lönnrot, the nineteenth-century scholar who collected and recorded the fragments and structured them into a lengthy narrative that ceased to resemble the oral originals just as it established their canonicity. It came into a published international canon as it gained recognition through the work of Henry Wadsworth Longfellow, a recognized contributor to that canon. Recent recognition of the validity of oral literary traditions and the questioning of existing literary canons suggest that this definition of literature is exclusive rather than inclusive. It is defined by difference as much as by sameness. Literature is defined as a limited set, the literary canon, and what does not meet the subjective criteria of canonicity is excluded.

While literature is classed in the 800s, folk literature is classified in 398.2 in the DDC, placed hierarchically under customs, etiquette, folklore. The manual tells us:

Manual Entry: 390 Customs, etiquette, folklore

Notes:

Customs, etiquette, and folklore are among the **raw material of the social sciences**, particularly of anthropology and sociology—the descriptive and analytical aspects of the study of the behavior of mankind in general social groups. **Melvil Dewey considered customs to be the culmination of social activity** and classed them in 390 (1998—emphasis added)

So folk literature is part of this culmination of social activity, the raw material of the social sciences, while literature consists of individuals’ refined products. The valuation placed on individuality and individual production in our culture suggests that folk literature is devalued by being separated out. Its anonymity and social nature place it outside of the category of *Literature*. It is different, not the same. Interestingly, the tradition of Bacon, Hegel, and Harris puts folk literature in a position nearing Hegel’s summit of enlightenment, the Idea, rather than in the illusory realm of the Essence. This valuation of folk literature over canonical literature is no longer a part of our discourse, and its remnant fails to reflect the original intent. However, the

arrangement serves to differentiate rather than gather, even though the status of the concepts is reversed.

This particular example illustrates how much the DDC reflects the culture (past and present) of which it is a product in the way that it allocates its pigeonholes. It is a problem of what we consider the same and what we consider different. We need to be aware of how we define our samenesses and how those definitions are culturally grounded.

What Is the Hierarchy of Samenesses?

The hierarchy or order of precedence of samenesses is the other problem rooted in the sameness/difference duality that I will explore. It both comes from and creates the hierarchical structure of classification. As discussed above, few things that we classify are made up of one aspect or facet. More typically they have many facets and in arranging them we privilege one facet over another. That is, we divide first by one facet, then by another and another and so on in a prescribed citation order. The result is a hierarchical arrangement that gathers effectively by the first facet following the idea that we gather what is the same and separate what is different. However, this arrangement then separates what is the same in subsequent facets. Continuing with the previous example, in the DDC main class 800 for literature the first facet is language, followed by genre and then by period with arrangements for geography sometimes interposed between language and genre. So all literature in German regardless of genre or period is together, but poetry of different languages is not all together. Genre is scattered according to the language in which it was written. Poetry in English is in 821, poetry in German is in 831, poetry in French is in 841, poetry in Italian is in 851, and so on for other languages. Further, if you want literature of the twentieth century it is even more widely scattered with twentieth-century German literature in 830.8009 (collections), 830.9009 (criticism), 831.9 (poetry), 832.9 (drama), 833.9 (fiction), 834.9 (essays), twentieth-century Spanish literature in 860.8006, 860.9006, 861.6, 862.6, and so on for all of the possible permutations of language, genre, and topic that can be further subdivided by period. So even something fairly specific such as an interest in twentieth-century European poetry will be diasporized across the 800s.

One problem with this hierarchy of facets in a one-size-fits-all standard such as our major library-classification schemes is that we do not always want to choose the same sameness first. That is, we do not always want what the citation order gives us. For example, if you are studying German literature you will probably want literature written in German together. However, if you are studying Canadian, South African, or U.S. literature you will want literature

grouped by geographic origin.

Looking up Canadian literature in the DDC 21 index you will find:

Canadian literature—English 810

Canadian literature—French 840

Canadian literature—Inuit 897.12

And of course it is possible to build numbers for Canadian literature in other languages—Cree, Chinese (Mandarin or Cantonese), Hindi, Ukrainian, and so on. However, taking only the three languages listed in the index illustrates the problem of studying Canadian literature. It becomes scattered because it is written in a variety of languages (see figure 2).

The same thing happens with South African literature:

South African literature—Afrikaans 839.36

South African literature—Bantu 896.39

South African literature—English 820

The problem demonstrated here is that the cultural background influences not only how we define our samenesses but which samenesses are primary for the organization of the classification. The fact that the major colonial languages define the majority of space in the 800s whereas the literatures of hundreds of different languages are crammed into the 890s is indicative of both bias and literary warrant in North American collections. However, the focus on language is also indicative of the largely colonial perspective of the classification of literature. It reflects the dominant discourse of the discipline. Colonizing countries have been more likely—in the past—to use a single language and impose that language on colonized countries. As Britain colonized many colonies, its language, English, is used in countries all around the world. Countries like Canada, South Africa, and the United States that were colonized by more than one European power have multiple European languages in addition to indigenous languages.

| | | 800s | | | | | | | | |
|---------------|---------|------|---------|-----------|--------|------|------|------|------|-------|
| | | 810s | 820s | 830s | 840s | 850s | 860s | 870s | 880s | 890s |
| Canadian | 810 | | | | 840 | | | | | 897.1 |
| | English | | | | French | | | | | Inuit |
| South African | | | 820 | 839.36 | | | | | | 896.4 |
| | | | English | Afrikaans | | | | | | Bantu |

Figure 2. DDC Literary Classification