

Rebecca Green

OCLC Online Computer Library Center, Inc., Library of Congress, Washington DC USA

See-also Relationships in the Dewey Decimal Classification^{1*}

Abstract: This paper investigates the semantics of topical, associative see-also relationships in schedule and table entries of the Dewey Decimal Classification (DDC) system. Based on the see-also relationships in a random sample of 100 classes containing one or more of these relationships, a semi-structured inventory of sources of see-also relationships is generated, of which the most important are lexical similarity, complementarity, facet difference, and relational configuration difference. The premise that see-also relationships based on lexical similarity may be language-specific is briefly examined. The paper concludes with recommendations on the continued use of see-also relationships in the DDC.

*Also published as: Green, Rebecca. 2011. See-also Relationships in the Dewey Decimal Classification. *Knowledge Organization* 38: 335-41.

1.0 Introduction

This paper investigates the semantics of topical, associative see-also relationships in schedule and table entries of the Dewey Decimal Classification (DDC) system. This study is part of a larger, ongoing assessment of relationships in the DDC, whose purpose is to establish a more logical and powerful representation of the scheme.

Introduced to the tables and schedules in DDC 20, see-also relationships have increased in number from edition to edition, being now found in over 40% more records than in DDC 20 (Dewey, 1989). As they have not been strictly defined and are not a key element of the structural hierarchy, the use of see-also relationships has undergone less scrutiny than other relationships.

Topical, associative see-also relationships are but one type of see-also relationship in the DDC and also but one kind of note in schedule and table entries that lead from one class to another. In addition to their use in schedule and table entries, see-also relationships also occur in the DDC's Manual entries and Relative Index displays. In the Manual, a note instructing the user to "*See also discussion at*" indicates that another Manual entry gives further information on the use of notation described in the source Manual entry. In the Relative Index, see-also references refer to headings where additional relevant numbers may be found. The headings led to are typically either broader terms or preferred (synonymous) terms. In thesauri, see relationships are usually employed to refer from non-preferred (lead-in) terms to preferred terms. In the DDC Relative Index, however, all entries list at least one associated number, making the see-also relationship the appropriate relationship for dealing with equivalence relationships there.

Entries in the tables (Tables 1–6) and schedules (000–999) include several types of notes for topics found elsewhere. One type—the do-not-use note—explains irregularities in the use of

¹ DDC, Dewey, and Dewey Decimal Classification are registered trademarks of OCLC Online Computer Library Center, Inc.

regular standard subdivision notation or add table provisions. Another set—relocation notes and discontinuation notes—gives the current location of topics previously classed at that number. Another type—the see reference—is used for subordinate parts of comprehensive or interdisciplinary topics found outside the notational hierarchy under a number. Class-elsewhere notes indicate where interrelated topics are found; in particular, they lead to numbers for comprehensive or interdisciplinary treatment of a topic and also clarify the meaning of numbers within the same notational hierarchy. Thus, the see-also relationship is only one of several types of relationships leading to numbers for related topics.

As the introduction of the DDC (Dewey, 2011) explains, see-also relationships “are reminders that minor differences in wording and context can imply differences in classification” (p. lvii); in thesaural parlance, they are associative relationships. (For further information on notes in the DDC describing what is found in other classes, see Chan and Mitchell [2003], pp. 25–30; for further discussion on the relationship between see-also references and associative relationships in the DDC, see Mitchell [2001], pp. 217–218.)

The remainder of the paper is organized as follows: Section 2 explores the semantic nature of see-also relationships, both in knowledge organization systems generally (especially in controlled vocabularies) and then in the DDC specifically, from a theoretical perspective. Section 3 reports on an empirical study of topical, associative see-also relationships in the DDC; a by-product of this study is a semi-structured inventory of the sources of the DDC’s see-also relationships. This section also briefly examines if some see-also relationships are language-specific. The final section of the paper concludes by proposing a path forward for the future use of see-also relationships in the DDC.

2.0 Theory of associative relationships

2.1 Associative relationships in knowledge organization systems

The ANSI/NISO *Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies* (NISO, 2005) are partially relevant to the investigation of see-also relationships in the DDC. On the one hand, these guidelines communicate “recommendations based on preferred techniques and procedures” regarding, *inter alia*, the treatment of associative relationships in knowledge organization systems. On the other hand, the guidelines specifically target controlled vocabularies (e.g., thesauri); many of the recommendations are not directly applicable to the DDC because of differences between controlled vocabularies and classification schemes. Specifically, controlled vocabularies are structured around terms, while classification schemes are structured around classes. Thus, in controlled vocabularies, associative relationships involve “terms [that] are semantically or conceptually associated to such an extent that the link between them *should* be made explicit in the controlled vocabulary, on the grounds that *may* suggest additional terms for use in indexing or retrieval” (p. 51). But associative relationships in a classification scheme are not meant to lead to additional classes that may be relevant, but to distinguish between the topics in two or more classes, so the relevant class may be identified.

What associative relationships in controlled vocabularies and classification schemes do have in common is that they are difficult to characterize. Hence Dextre Clarke's (2001) comment on the associative relationship (typically represented as a related term [RT] in thesauri): "The presence of an RT link . . . depends more on . . . what will serve the users than on a precise semantic analysis" (p. 46). And, according to Svenonius [2000, 162], "the possibility remains that really helpful [associative] relationships . . . do not lend themselves to formalization." However, the ANSI/NISO guidelines suggest that it is important to make the nature of associative relationships explicit so they may be used consistently (p. 51).

The ANSI/NISO guidelines further note that for some controlled vocabularies it may be "desirable" to "refine" the expression of associative relationships "to make the nature of [specific types of associative] relationships explicit" (p. 57). It could be useful to do so in the DDC, since see-also relationships there are currently not well-defined. In order to give clear direction on when they should be used (if indeed, they should be used at all), it is first necessary to identify specific types of relationships that are being expressed as see-also relationships. Only then can thoughtful consideration be given to which specific types of relationships ought to be expressed as see-also relationships.

2.2 Associative relationships in the DDC

A see-also relationship in the DDC directs the user to another class in relation to a specific topic; the semantic nature of the see-also relationship is dependent on the relationship that exists between a topic in the class containing the see-also relationship (the source topic) and the topic named in the see-also relationship (the target topic), which should be classed in the other number. For example, 364.164 Violent offenses against property contains the following see-also relationship: *See also 364.166 for copyright piracy*; it is the relationship between piracy, a topic named in the including note ("Including piracy, sabotage") and copyright piracy, the topic named in the see-also relationship that is of concern to us. (Source topics are found in captions and such note types as class-here notes, including notes, and variant-name notes.)

3.0 Empirical study of associative relationships in the DDC

3.1 Methodology

Overall, 1678 see-also references are found in 1383 DDC table and schedule records, of which only 360 are reciprocal (that is, 180 pairs of numbers refer to each other through see-also relationships). Of these, a random sample of 100 records with see-also relationships was drawn. These 100 records include 118 see-also relationships, all of which were investigated as part of the study.

An initial goal of the paper is to identify subtypes of the relationships presented through see-also references. In assessing the semantic nature of a see-also relationship, the first order of business is to identify the topic in the source class that the topic named in the see-also relationship is related to. In most cases this topic is readily apparent: the two topics typically will share at least one word in common, or have similar-sounding words in common, or contain words that are synonyms or enter into some other strong lexical relationship. (In one case, the topic of the

source class governing the see-also relationship was identified on the basis of a reciprocal see-also relationship in the target class.) In the sample, source topics are twice as likely to be found in the caption as in all other locations combined.

The evidence used to identify the source topic is often useful as well in determining the semantic nature of the relationship between source and target topics. Another important source of evidence for this task is the Relative Index (RI). Because the Relative Index is constructed in adherence with a principle of consistency, RI terms sometimes normalize a relationship that is not so clearly expressed in natural language. A prime example of this normalization is the case where one RI term consists of a main heading, without subdivision, and another RI term consists of the same main heading with a subdivision. For example, at 621.55 Vacuum technology is a see-also relationship to 533.5 for vacuum physics. The semantic nature of this relationship becomes clearer when we see that Vacuums—engineering is a Relative Index term for 621.55, while Vacuums—unsubdivided—is a Relative Index term for 533.5. The absence of a subdivision in the latter RI term indicates that 533.5 is the interdisciplinary number for vacuums, while 621.55 is the number for vacuums considered from an engineering perspective. (This RI scenario is also an example of a see-also relationship in which DDC editorial rules have been misapplied. Unless they are in the same hierarchy, the relationship between the interdisciplinary number for a topic and the same topic treated in another discipline is to be expressed as a see reference, while the relationship between a topic treated within a single discipline and its interdisciplinary treatment is to be expressed using a class-elsewhere note (that is, if it *is* elsewhere).

Several limitations involved in the process of assessing the semantic nature of the see-also relationship should be acknowledged. First, only one person made the assessment. It is uncertain if other persons would evaluate the relationships similarly. Second, the set of relationship types evolved during the assessment process. Third, the relationship types identified in the study are not strictly defined and indeed are not identified as subtypes of the see-also relationship in the editorial rules of the DDC.

But the significance of these limitations is itself constrained. First, no attempt is made here to make statistically significant statements. Second, there is no serious alternative to identifying relationship types inductively. The evolution of the relationship inventory represents efforts to normalize relationship assessments over time. Moreover, some, but not all, of the relationship types used here had been identified in a predecessor study. Lastly, at this point it is not nearly so important to identify all types of see-also relationships as to identify the most common ones.

Source classes in the random sample are distributed as follows: Table 1 (2), Table 2 (12), Table 3 (0), Table 4 (0), Table 5 (2), Table 6 (0); 000 (3), 100 (5), 200 (3), 300 (23), 400 (2), 500 (9), 600 (19), 700 (14), 800 (2), and 900 (4). This distribution generally mirrors the proportion of see-also references throughout the tables and schedules; in terms of absolute numbers (as opposed to proportions) only the underrepresentation of the 500s and the overrepresentation of the 700s in the sample are worth mention. In particular, the larger numbers of see-also references from Table 2, the 300s, and 600s occurring in the sample accurately represent the overall distribution of see-also references.

3.2 Inventory of see-also relationship types

A semi-structured inventory of sources of see-also relationships has been generated on the basis of the random sample and assessment procedures described above. The following major sources of see-also relationships were found: lexical similarity, complementarity, facet difference, and relational configuration difference. Each of these sources accounts for roughly 10-25% of the sample cases (each see-also relationship in the sample was assigned a primary source).

3.2.1. Lexical Similarity

The lexical similarity category covers two major see-also relationship subtypes, even of which is manifest in different ways. One is the use of the same name or a similar name for entities belonging to different classes; typically the entities are unrelated except for the lexical similarity, although that is not always the case. Consider, for example, the following four examples:

2—719 2 *Northwest Territories (1870–1999)

See also —7193 for Northwest Territories (1999–)

2—764 252 Austin County

See also —76431 for Austin (city)

5—975 Peoples who speak, or whose ancestors spoke, Siouan, Iroquoian, Hokan, Chumash, Yuki languages

...

See also —979 for Yuchi

583.98 *Campanulales

Including Campanulaceae (bellflower family), Goodeniaceae, Lobeliaceae, Stylidiaceae; bluebells, Campanula, Indian tobacco

See also 583.94 for bluebells of forget-me-not family; also 584.32 for bluebells of lily family

The see-also relationship in the first of these examples ties together the two classes for the same-named geopolitical entity (Northwest Territories) before and after another geopolitical entity (Nunavut) separated from it (see-also relationships similarly relate political parties of different time periods). While the hierarchy usually provides adequate context, see-also references are provided in Table 2 in a limited area (state or province) to distinguish between the same or similar geographic names that refer to different geographic entities. Thus, the see-also relationship is supplied in the second example to distinguish between Austin County and the city of Austin (which is not in Austin County, but in Travis County). Similarly in the third example,

it is typical that a name of another language (or a name in another language) may be represented in multiple ways in a second language: it would not be surprising if Yuki and Yuchi referred to the same language. But in fact they do not. Yuchi is a language isolate spoken by a Native American people living in Oklahoma, while Yuki (also known as Ukiah) is spoken by a Native American people living in California. In the fourth example, popular biological/botanical names may suggest a closer relationship than are recognized in current biological classifications. The 580s and 590s are replete with such see-also relationships (for which reciprocal see-also relationships are generally provided).

The lexical similarity category also covers homonymy (the use of the same lexical form for unrelated meanings) or, more frequently, polysemy (the use of multiple senses of a single term). The shared lexical form in each of the two examples below is polysemous: “biscuit” can refer to two different food items, one a cookie, the other a quick bread; “mystery” can refer to two different kinds of plays, one a religious play of historical vintage, and the other a drama of modern origin.

641 .865 4 Cookies

Variant name: biscuits

See also . . . 641.8157 for biscuits (quick breads)

792.16 †Religious and morality plays

Including miracle, mystery, passion plays

See also . . . 792.27 for modern mystery plays

3.2.2 Complementarity

Complementarity—a term borrowed from linguistics, where it refers to a form of antonymy—has two distinct manifestations in the DDC. (In the absence of specific instructions to the contrary, the general rule for an antonym in the DDC is to class it in the same number as its opposite.) We adopt the term here to emphasize that the two topics related by the see-also relationship together form a whole; they complement each other. The first manifestation of complementarity occurs frequently in Table 2 (as seen below) with its many see-also relationships between land masses and adjoining bodies of water. On the one hand, land and water contrast with one another; on the other hand, a land mass and adjoining body of water form a larger geographic unit.

2—598 Indonesia and East Timor

Class here Malay Archipelago, Sunda Islands

See also —16473 for inner sea of Malay Archipelago; also —16474 for seas adjoining southern Sunda Islands

The second manifestation binds together two entities, one of which is more-or-less loosely defined in terms of the negation of the other. For instance, at the heart of the relationship between 181 and 190 in the first example below is the contrast between eastern and noneastern; together eastern and noneastern cover the entire geographic scope of philosophy. In the second example, “other plastic arts” are defined by context as all plastic arts that are not sculpture. As with the previous example, one of the topics related by the see-also relationship is defined in terms of the other topic, and taken together, the two topics cover the entire scope.

190 Modern western and other noneastern philosophy

See also 181 for eastern philosophy

> 736–739 Other plastic arts

...

See also 731–735 for sculpture

...

3.2.3 Facet difference

The facet difference category covers cases where the two topics related by the see-also relationship represent different facets of a subject. Alternatively, we may say that each of the related topics has a different focus within some general framework. In the first example below, the see-also relationship is dependent on the relationship between medicine and health: the goal of medicine is to promote good health; thus, 616.9803 emphasizes the means, while 613.92 emphasizes the end or goal. The workings of the see-also relationships in the second example is most apparent by looking at the Relative Index terms assigned to the topics at the two ends of the see-also relationship: Milling tools vs. Milling metals. At 621.91, the emphasis is on the tools used for milling; at 671.35 the emphasis is on the object that is milled.

616.980 3 Industrial and occupational medicine

...

See also 613.62 for industrial and occupational health . . .

621. 91 Planing and milling tools

See also 671.35 for machining metal

3.2.4 Relational configuration difference

In the fourth category we find see-also relationships linking complex topics that include the same topical components, but that relate those topical components in different ways. The end result is that the relationships among the topical components form different configurations. In the first example below, both 379.158 and 371.26 concern schools and evaluation, but 379.158 concerns the evaluation of schools, while 371.26 concerns evaluations of students, as administered in schools. In the second example block (which shows context for not-quite reciprocal relationships), see-also relationships distinguish between diagnosis based on analysis of blood at 616.07561 and diagnosis used to identify diseases of blood at 616.15075.

379.158 School standards and accreditation

Class here educational evaluation, school accountability

See also . . . 371.26 for examinations and tests

616.075 Diagnosis and prognosis

616.075 61 Blood analysis

See also 616.15075 for diagnosis of diseases of blood

616.15 Diseases of blood

See also 616.07561 for use of blood analysis in diagnosis of diseases in general

616.150 75 Diseases of blood—diagnosis

3.2.5 Other see-also relationship sources

The four sources of see-also relationships discussed above yield scenarios in which classifiers might legitimately need assistance in understanding the scope of one class by contrasting it with a topic that is classed elsewhere. There are additional circumstances in which see-also relationships have been supplied, where it is not as clear that the relationship between topics is sufficiently regular to warrant an explicit relationship. One group of these depends on world knowledge or knowledge of the DDC. In the first example below, the see-also relationship depends on world knowledge that dilatation and curettage (D&C) has been used (more frequently in the past than now) as a method of surgical abortion. As D&Cs have many other motivations, giving this see-also relationship is somewhat suspect. Giving a see-also relationship in the second example implies an assumption that some users might think that intellectual property is deemed to be a “specific [item]” in the DDC. Just over 10% of the see-also relationships in the sample depend on world knowledge or DDC knowledge.

618.145 8 Dilatation and curettage

See also 618.88 for surgical abortion

364 .162 8 Theft of specific items

See also 364.1662 for theft of intellectual property

3.2.6 Misapplication of DDC editorial rules

As indicated previously, see references and class-elsewhere notes are to be given under some very specific circumstances. For example, a see reference should be given from the interdisciplinary number for a topic to numbers for the topic in other disciplines, while a class-elsewhere relationship should be given from the number for the treatment of a topic in a discipline to the interdisciplinary number (unless these are the same number). Relative Index terms can be used to identify interdisciplinary numbers and numbers for topics in specific disciplines. The rules about which type of note to give have been misapplied in 10-15% of the sample cases; that is, see-also relationships have been given when a see reference or class-elsewhere note was called for. Approximately 5% more of the cases involve the use of see-also relationships between the treatment of a topic in two disciplines where no interdisciplinary number is given. Here no general decision has been made as to which kind of note to give. As it may not always be possible to identify the most appropriate interdisciplinary number for a given topic, guidance needs to be given on referring between different disciplinary treatments of a topic in the absence of an interdisciplinary number: what kind of note should be used, and where should they be placed?

3.3 Language-specific nature of some see-also relationships

See-also relationships based on lexical similarity may not always be relevant in translations of the DDC. This premise was investigated in a preliminary manner by examining the German (Dewey, 2005) and Italian (Dewey, 2009) translations of the DDC for some of the see-also relationships in the sample that are based on lexical similarity. Given that the vocabulary of English is largely drawn from Germanic and Romance sources, the degree to which the German and Italian translations avoid the lexical similarity issues found in the English is likely to be magnified in the languages of some other translations, for example, Arabic, Greek, Hebrew, Icelandic, Russian, and Vietnamese. If the sources of see-also relationships were annotated, translators would know when it might not be necessary to translate a see-also relationship.

The German and Italian translations were each examined for a dozen cases of lexical similarity in English. For the most part, the lexical similarity was duplicated in the translations. It is not immediately clear to what extent this duplication resulted from explicit attempts on the parts of translators to mirror the English translation as closely as possible or resulted from the same lexical similarity existing in the other languages. But there are situations where the relevance of the see-also relationship might be called into question. For example, in English, the including note for popular names at 583.94 mentions “Virginia cowslip (bluebell), with see-also relationships to 583.98 for bluebells of bellflower family and to 584.32 for bluebells of lily family. In the Italian translation, “campanule” are mentioned in both see-also relationships, but there is no “campanule” in 583.94’s including notes.

There may not be that many cases where DDC translational practice diverges from patterns present in the English. But where the motivation for a see-also relationship in English does not exist in a translation, the see-also relationship should not be included in the translation. By the same token, where there is motivation for a see-also relationship in a translation but not in English, the see-also relationship should be added to the translation. For example, the German translation has added a see-also relationship from 583.625, with “Violaceae (Familie der Veilchengewächse)” in the including note, to 583.675 for Alpenveilchen and to 583.95 for Usambaraveilchen. Translation teams have been given such direction on a case-by-case basis, but the general principles have not yet been formally established as part of DDC translation practice, so implementation of the language-specific principles tends to be inconsistent.

4.0 Conclusion

The foregoing analysis of see-also relationships in the DDC has enumerated several sources of this type of relationship, including lexical similarity, complementarity, facet difference, relational configuration difference, and world knowledge. A first step that is needed going forward is a series of editorial decisions about which of these potential sources of see-also relationships to recognize in the DDC. The circumstances of their use need to be well-defined so these relationships can be created, maintained, and interpreted consistently. A second step is to review current (and new) see-also relationships against those editorial decisions and to code the motivating source of those retained. In this process, some see-also relationships may be deleted or converted into other kinds of relationships. A third step is to codify the general principle that see-also relationships should be included in a translation if and only if the circumstances outlined by the editorial decisions from step one are met in the translation. With these actions, the see-also relationship can take its place within the well-understood semantics of the DDC.

References

- Chan, Lois Mai, and Mitchell, Joan S. 2003. *Dewey decimal classification: principles and application*. Dublin, Ohio: OCLC.
- Dextre Clarke, Stella G. 2001. Thesaural relationships. In Bean, Carol A. and Green, Rebecca, eds., *Relationships in the organization of knowledge*. Dordrecht: Kluwer Academic Publishers, pp. 37–52.
- Dewey, Melvil. 1989. *Dewey decimal classification and relative index* (20th ed.; John P. Comaromi, Julianne Beall, Winton E. Matthews, Jr., & Gregory R. New, eds.). Albany, N.Y.: Forest Press.
- Dewey, Melvil. 2005. *Dewey-Dezimalklassifikation und Register* (22nd ed.; Joan S. Mitchell, Julianne Beall, Giles Martin, Winton E. Matthews, Jr., & Gregory R. New, eds.). Deutsche Ausgabe. München: K. G. Saur.
- Dewey, Melvil. 2009. *Classificazione decimale Dewey e indice relative* (22nd ed.; Joan S. Mitchell, Julianne Beall, Giles Martin, Winton E. Matthews, Jr., & Gregory R. New, eds.). Edizione italiana. Roma: Associazione Italiana Biblioteche.
- Dewey, Melvil. 2011. *Dewey decimal classification and relative index* (23rd ed.; Joan S. Mitchell, Julianne Beall, Rebecca Green, Giles Martin, & Michael Panzer, eds.). Dublin, Ohio: OCLC.

Rebecca Green. 2011. See-also relationships in the Dewey Decimal Classification. In Smiraglia, Richard P., ed. *Proceedings from North American Symposium on Knowledge Organization*, Vol. 3. Toronto, Canada, pp. 41-51.

Mitchell, Joan S. 2001. Relationships in the Dewey Decimal Classification System. In Bean, Carol A. and Green, Rebecca, eds., *Relationships in the organization of knowledge*. Dordrecht: Kluwer Academic Publishers, pp. 221–226.

National Information Standards Organization. 2005. *Guidelines for the construction, format, and management of monolingual controlled vocabularies*. Bethesda, Md.: NISO Press. (ANSI/NISO Z39.19-2005).

Svenonius, Elaine. 2000. *The intellectual foundation of information organization*. Cambridge, Mass.: MIT Press.