The Semantic Distance Model of Relevance Assessment

Terrence A. Brooks
University of Washington, Seattle, Washington

Abstract

This paper presents the Semantic Distance Model of Relevance Assessment. It is a cognitive model of the relationship between semantic distance and relevance assessment. Premises of the model such as the subjective nature of information and the metaphor of semantic distance are discussed. Empirical results illustrate the effects of semantic distance and semantic direction. Relevance auras, a combination of vertical and horizontal relevance assessments are also presented.

An ongoing series of experiments (Brooks, 1995a, 1995b, 1997) has investigated the relationship between semantic distance and relevance assessments. This essay cumulates these results and adds some new experimental results in presenting the Semantic Distance Model (SDM) of Relevance Assessment. The SDM is a cognitive model describing the relevance relationships between bibliographic database records and topical subject descriptors.

The SDM proposes that relevance assessment systematically declines with semantic distance. Such a decline of stimuli effect as a factor of distance in psychological space is typical of many cognitive models (Shepard, 1987). The SDM provides empirical support for the general indexing/cataloging notion that subject-term hierarchies express a conceptual arrangement from broad to narrow. The SDM also provides empirical support for the notion that the "appropriate" subject descriptor will be perceived as more relevant than those at greater semantic distance. A semantically distant descriptor might be one that resides at either a broader or narrower position in the same subject-term hierarchy.

A number of unanticipated findings have also been observed: (1) It appears that the SDM can be switched off and on depending on the context of relevance assessment, (2) It appears that topical subject expertise steepens the rate of decline of relevance assessment, and (3) It appears that relevance assessment is contingent on semantic direction; that is, the rate of decline differs if one is assessing a series of increasingly narrow descriptors, or increasingly broad descriptors.

The archetype trial of the experiments supporting the SDM has been the comparison of two texts. One text has been bibliographic records trimmed to title and abstract, while the other text has been topical subject descriptors. Relevant assessments were harvested from library science, economics, engineering and mathematics graduate students as they assessed the relevance relationship between subject descriptors and trimmed bibliographic records from the literatures of education, economics, engineering and mathematics.

INFORMATION AS A SUBJECTIVE PHENOMENON

A fundamental assumption of the SDM is that information is a subjective phenomenon, a product of the interpretation of texts (Cole, 1994; Dervin & Nilan, 1986). This assumption is widely supported. The transactional model of reading asserts that the active negotiation between readers and texts produces meaning (Straw, 1990). Jacob and Albrechtsen (1997, p. 42) describe reading "not a passive reception of meaning unpacked from a neutral and fixed, or stable, reference but an active process of contextualizing the word within the particular set of relativized and competing definitions that inform the moment of expression". Bertrand-Gastaldy, et al. (1995, p. 56) consider text as an interweave of multiple semiotic systems: "It is not the character strings that are significant; rather, it is their properties specific to each of these systems and interpreted by a cognitive agent."

Since information is a subjective phenomenon, it is hardly surprising that readers will disagree over the meanings of text. The disagreement exhibited by indexers and catalogers has already been widely documented (Broadbent & Broadbent, 1978; Thorngate & Hotta, 1990). There appear to be a number of

cognitive factors at work. The frequency distribution of names for objects prompted Furnas, et al. to report that "no single name, no matter how well chosen, can do very well" (1983, p. 1,796). Some of the variability of naming is due to fuzzy natural categories that permit verbal concepts to be modified to suit changing situations (McCloskey & Clucksberg, 1978); other variability is due to the graded structure of categories that permits category members to be ranked according to the category ideal (Rosch, 1975; Shoben, 1976).

Recognizing the subjective nature of information does not deny the possibility of agreement about the meaning(s) of a text. If one were to control factors like vocabulary, culture, subject expertise, and so on, a group of readers may agree about the meaning of a text. The work of indexers/catalogers is premised on this potential agreement. It also gives force to the concept of topical relevance. Analogous agreement underlies the basic category of abstraction of ordinary objects:

> We will argue that categories within taxonomies of concrete objects are structured such that there is generally one level of abstraction at which the most basic category cuts can be made. In general, the basic level of abstraction in a taxonomy is the level at which categories carry the most information, possess the highest cue validity, and are, thus, the most differentiated from one another. (Rosch, et al., 1976, p.382)

Online searching is one battleground where the subjective nature of information and textual core meanings clash. The archetypal problem of online searching is disagreeing with the indexer's labeling of a record; or put another way, being surprised by the records fished out of a database by a given descriptor.

Relevance assessment research evolves by investigating textual meanings under controlled conditions. The SDM can be viewed as a device for calibrating agreement about the relationship of bibliographic records and subject descriptors. It describes how the average reader will assess relevance relationships. It does not predict how any particular reader will assess the relationship between a specific subject descriptor and bibliographic record.

THE METAPHOR OF SEMANTIC DISTANCE

Semantic distance is a psychological construct that has been used to locate concepts along various dimensions of meaning (Schvaneveldt, Durso & Mukherji, 1982). Placing database records in n-dimensional space, or docuverse (Das Neves, 1997, Wallmannsberger, 1991) has been an ancient and seductive information conceit. As early as 1973, Charles Bachman urged database programmers to forsake their linear record model in favor of n-dimensional data spaces. Twenty five years later, the spatial metaphor is still a mainstay of theorizing. Chen et al, (1998), for example, proposed a concept space approach to internet searching.

The SDM used a variety well-known trees of descriptors and established hierarchies of subject terms, as well as non-library hierarchies, to provide semantic distance. The witness of the SDM with a variety of hierarchies, including non-library ones, suggests its robustness. So far, an SDM effect has been produced by the following subject hierarchies and descriptor trees:

* The Thesaurus of ERIC descriptors in Brooks (1995a, 1995b, 1997).
* The Index of Economic Articles in Journals and Collective Volumes in Brooks (1995b).
* Mathematical Reviews in Brooks (1995b).
* INSPEC Thesaurus, 1993. The Institution of Electrical Engineers. (As yet unpublished, but a portion of the results appears below.)
* Thesaurus of Psychological Index Terms. American Psychological Association, 1994. (As yet unpublished, but a portion of the results appears below.)
* Help-menu terms presented by Kreigh, Pesot and Halcomb (1990) in Brooks, (1995b). Help-menus are a feature of many computer interfaces that permit computer users to select appropriate help information. A sample hierarchy of help-menu terms would be "Science", "Technology", "Agriculture", "Field", "Grain" and "Wheat" (Kreight, Pesot & Halcomb, 1990, p. 652).

RELEVANCE AND ITS ASSESSMENT

Relevance has remained a primitive information science construct despite decades of struggle (Mizzaro, 1997). The current unresolved state of theorizing about relevance inhibits the conduct of empirical experimentation, which requires an operationalization of the dependent variable, relevance assessment. The SDM operationalized relevance assessment as a similarity between the prima facie meaning of a bibliographic record and the prima facie meaning of a subject descriptor. Other studies have used similarity in semantic matching tasks. Pierce, et al. (1992, p. 654) found "relatedness" and "semantic similarity" help experimental subjects match a target word to a set of alternatives in a computer menu display. Atkinson (1984, 1989) used intertextual similarity to describe the decision making of collection developers.

Substituting “similarity” for “relevance” merely replaces one poorly understood word for another. It does have a practical advantage, however. While “relevance assessment” may be information science jargon that dazzles the outsider, no graduate student of mathematics, economics, engineering or library science felt stymied by the request to judge the similarity of a bibliographic record and a subject descriptor. No experimental subject objected to this request, nor wished to engage me in a philosophical discussion about the meaning of text, or the ways in which texts could be similar. My observation is that ordinary people feel perfectly competent to read two texts and judge their similarity of meaning.

A typical experiment was constructed as follows. First a descriptor tree or hierarchy was selected. An online search then located a bibliographic record captured by the top term in a hierarchy. A check was made to ensure that this record was not indexed by any other term in the same hierarchy. A similar online search located a bibliographic record captured by the bottom term in the hierarchy that was not indexed for any other term in the same hierarchy. The result of this procedure was two bibliographic records for each hierarchy: a top record and a bottom record. The records were trimmed to title and abstract fields. A computer program stored the trimmed bibliographic records and index terms from the hierarchy. An experimental trial consisted of the joint presentation of a trimmed bibliographic record and an index term. Randomization of the trials nullified any order effects in the presentation of records and descriptors.

Two different types of experiments were used. One experiment collected the time needed to find the "relevant" descriptor from a list of descriptors. Rips, et al. (1973) also used time to calibrate semantic distance, and found that the time spent evaluating the semantic relation between two nouns increases with the semantic distance between them. Another type of experiment used sliding bars that permitted subjects to express degrees of relevance, or “partial” relevance (Spink and Griesdorf, 1997). A transformation turned each subject’s raw scores into $\underline{z}$ scores, standard normal variates centered on zero with a standard deviation of one. As a consequence of this standardization, the negative half of the distribution of $\underline{z}$ scores was arbitrarily defined as “non-relevant”, while the positive half falling above the mean was arbitrarily defined as “relevant”. In the results presented below, a negative mean relevance assessment indicates that experimental subjects gave it lower scores on average, or assessed it as nonrelevant. A positive mean relevance assessment indicates that experimental subjects gave it higher scores on average, or assessed it as relevant.

THE SEMANTIC DISTANCE MODEL

Semantic Distance Effect

A semantic distance effect is an inverse correlation between semantic distance and relevance assessment: Near descriptors are assessed as relevant, far descriptors are assessed as non-relevant. Further, relevance assessments are expected to decline systematically with semantic distance. One implication of a semantic distance effect is that the indexer-generated, topically relevant term would be assessed as most relevant.

Table 1: Relevance Assessments for ERIC Descriptors and Records

Semantic Steps	Trials	Means	Standard Deviation
0	680	0.70	0.92
1	680	0.37	0.89
2	680	-0.08	0.92
3	680	-0.27	0.89
4,5,6	1020	-0.48	0.87

The relevance assessments of descriptors and bibliographic records from the ERIC database (Brooks, 1995a) in Table 1 present a typical semantic distance effect. The topically relevant term at zero semantic steps has a mean z-score of 0.70 based on 680 experimental trials. As semantic distance increases, relevance assessments decline systematically: 0.70, 0.37, -0.08, -0.27, -0.48. Between one and two semantic steps, relevance assessment flips from relevant to non-relevant.

Table 2: Relevance Assessments for Mathematics Records

	Mathematicians			Non-mathematicians				
Semantic distance	Mean	Standard deviation	Trials	Mean	Standard deviation	Trials	t	p
0	0.30	0.64	50	0.35	0.53	150	-0.53	0.59
1	0.11	0.78	50	0.39	0.52	150	-2.80	0.00
2	-0.04	0.67	50	0.20	0.48	150	-2.80	0.00

Table 2 presents the relevance assessments based on a classification scheme used by Mathematical Reviews. In this experiment (Brooks, 1995b) mathematics graduate students were compared to non-mathematicians (i.e., library school and economics graduate students). The mathematics classification scheme permitted only a semantic distance of only three steps. The mathematicians display a semantic distance effect across three semantic steps (0.30, 0.11, -0.04), while the non-mathematicians did not display a semantic distance effect (0.35, 0.39, 0.20).

Table 2 illustrates that the SDM is sharpened by topical subject expertise. While the relevance assessments of mathematicians and non-mathematicians do not differ at 0 semantic steps ($p = 0.59$), they differ significantly at steps 1 and 2 ($p = 0.00$). Subject experts can spy partially relevant (mean relevance assessment at 1 semantic step is 0.11) and non-relevant subject terms (mean relevance assessment at 2 semantic steps is –0.04). Non-mathematicians can do neither of these.

Another experiment illustrated that the SDM is contingent on the context of relevance assessment. Brooks (1995b) presented single-word help-menu terms in a random mix with multiword classifications. An example of a multiword classification is "Theory of the Household (consumer demand)" while an example of a help-menu term is "Science". Table 3 illustrates that the single-word help-menu terms suffered in the comparison. In the presence of multiword phrases, experimental subjects assessed them all negatively. No semantic distance effect was exhibited (-0.78, -0.44, -1.55, -0.72). The identical help-menu terms and records were then used in a follow-up experiment (Brooks, 1995b). This time they were randomly mixed with single-word ERIC descriptors. In that company, they showed a strong semantic distance effect (0.638, -0.138, -0.380, -0.697). These results illustrate that relevance assessment can be switched off and on depending on the context of assessment. Harter described relevance assessment as "time-, order-, and situation-dependent." (1996, p. 39)

Table 3: Help-menu terms and ABI/INFORM records

	Multiword Descriptors		Single-word Descriptors	
Semantic Distance	Mean	Trials	Mean	Trials
0	-0.78	60	0.638	224
1	-0.44	60	-0.138	224
2	-1.55	60	-0.380	224
3	-0.72	60	-0.697	224

The robustness of the SDM was investigated by collecting relevance assessments from disturbed hierarchies. A disturbed hierarchy has one or more terms missing. The research question was: If the semantic steps of a term hierarchy of the ERIC thesaurus were disturbed by one or missing terms, would there still be a semantic distance effect? The treatments were a bibliographic record in the context of a five-term hierarchy, the same record with the same hierarchy missing one term, and the same record with the same hierarchy missing two terms. The three treatments can be visualized as follows:

A five-term hierarchy:	The same hierarchy missing one term:	The same hierarchy missing two terms:
Humanities Fine arts Visual arts Drafting Technical illustration	Humanities Fine arts **[Missing Term]** Drafting Technical illustration	Humanities **[Missing Term]** Visual arts **[Missing Term]** Technical illustration

Table 4 shows a typical semantic distance effect with the full five-term hierarchy (.524, .170, .043, -.262, -.502). Both of the disturbed hierarchies produce semantic distance effects as well, suggesting the robustness of the SDM. Disturbed hierarchies have not heretofore been investigated. It is intriguing research question why the most disturbed hierarchy, the one missing two terms, would flip to nonrelevance sooner than the full five-term hierarchy.

Table 4: Semantic Distance and Disturbed Hierarchies

	Full Five Terms		Missing One Term		Missing Two Terms	
Semantic Steps	Trials	Mean	Trials	Mean	Trials	Mean
0	149	.524	139	.562	154	.591
1	149	.170	139	.033	Missing	Missing
2	148	.043	Missing	Missing	157	-.119
3	146	-.262	140	-.421	Missing	Missing
4	150	-.502	137	-.566	158	-.335

Semantic Direction Effect

A semantic direction effect occurs when the distance to non-relevance systematically varies up or down a hierarchy of index terms. Table 5 presents relevance assessments for top terms in trees of ERIC descriptors (Brooks, 1995a). The semantic steps represent narrower descriptors residing lower in the same tree of descriptors. There is a clear semantic distance effect with relevance assessments flipping to non-relevance at one semantic step.

Table 5: Relevance Assessments for Top Records

Semantic Steps	Trials	Means	Standard Deviations
0	340	0.24	0.86
1	340	-0.01	0.82
2	340	-0.41	0.87
3	340	-0.64	0.85
4,5,6	510	-0.82	0.75

Table 6 presents relevance assessments for records gathered for ERIC descriptors at the bottom of trees of descriptors (Brooks, 1995a). The semantic steps represent broader descriptors residing higher in the same tree of descriptors. There is a clear semantic distance effect with relevance assessments flipping to non-relevance after three semantic steps. (Note the orientation of the table from bottom to top.)

Table 6: Relevance Assessments for Bottom Records

Semantic Steps	Trials	Mean	Standard Deviation
4,5,6	510	-0.14	0.85
3	340	0.10	0.78
2	340	0.25	0.84
1	340	0.75	0.78
0	340	1.16	0.73

The difference in distance to non-relevance in Tables 5 and 6 represent a semantic direction effect. The semantic direction effect occurs when descriptors below a bibliographic record degrade into non-relevance more quickly than the descriptors above a record. To coin a phrase, the distance downward to non-relevance is shorter than the distance upwards to non-relevance. For example, the semantically narrow descriptor "Pine Trees" would be discounted for a semantically broad record about Northwest Forests, while the heading "Northwest Forests" might be considered appropriate for a bibliographic record narrowly focused on pine trees. The timed experiment (Brooks, 1995a) found supporting evidence. Semantically narrower descriptors were harder for subjects to match to bibliographic records than indexer-chosen descriptors or semantically broader descriptors. The proportion of successful broader term searches (81.4%) was nearly equal to the proportion of successful indexer-term searches (83.3%), but the proportion of narrower term searches was significantly smaller (64.4%, $p < 0.001$).

These general effects of semantic distance and direction have already been observed in another computer-supported searching function. Snowberry, Parkinson and Sisson (1985) reported that experimental subjects experienced difficulty associating terms when they were separated by deep help-menu hierarchies (i.e., the effect of semantic distance) and when moving down help-menu hierarchies seeking a narrow term (i.e., the effect of semantic direction). Kreigh, Pesot and Halcomb (1990) replicated these effects.

Related Term Effects

The SDM can be extended by inquiring about the relationship between relevance assessment and semantic distance expressed “horizontally”; that is, using related terms to move horizontally across various terms hierarchies. Many hierarchies of subject descriptors provide related terms. Related terms are links to other, closely associated terms that are not hierarchical relationships (Green, 1995). Related terms may give part-whole and near-synonym links designed to be helpful to the online searcher (Houston, 1990).

Constructing an experiment with related-term linkages is more difficult than constructing an experiment with hierarchies of terms. Hierarchies can be found already constructed. The experimenter must construct related-term linkages, however, which usually involves selecting one related term from several. This procedure has the risk of introducing unwanted researcher bias. Brooks (1997) arbitrarily used the first-listed related term under each ERIC descriptor. The only caveat was avoiding mutually referencing terms that lead back and forth to each other. In this case, the second listed term was chosen.

Table 7 shows relevance assessments for both vertical and horizontal semantic distances. There is a clear semantic distance effect for the five vertical steps; there is a semantic distance effect for two horizontal steps. The positive relevance assessments of the top records is shaded. This relevance aura for top records extends one semantic step vertically and horizontally.

Table 7 Relevance Aura for Top Records (Shaded Area)

		Horizontal Semantic Distance			
	Top Record	0	1	2	3
	0	0.895	0.037	-0.294	-0.139
Vertical	1	0.000			
Semantic	2	-0.420			
Distance	3	-0.427			
	4	-0.603			
	5	-0.683			

Table 8 presents the corresponding data for bottom records. There is a semantic distance effect extending for five vertical steps; there is a semantic distance effect extending for three horizontal steps. The positive relevance assessments are shaded. This relevance aura for bottom records extends upwards five steps and two steps horizontally.

Table 8: Relevance Aura for Bottom Records (Shaded Area)

	5	0.265			
	4	0.462			
Vertical	3	0.529			
Semantic	2	0.558			
Distance	1	1.125			
	0	1.164	0.537	0.47	-0.114
	Bottom Record				
		0	1	2	3
		Semantic Distance in Horizontal Dimension			

Semantic distance effects have been exhibited in another study that included descriptors from ERIC, INSPEC and the Thesaurus of Psychological Index. In this study, the related term linkages for top records (1.04, 0.15, 0.13, -0.08) showed a semantic distance effect up to three steps away.

Tables 7 and 8 present relevance auras in two dimensions, but also beg an interesting research question: What about the related terms that would fill up the two-dimensional matrix of terms? How many of the interior terms of these two-dimensional term matrixes would be considered relevant?

In an as yet unpublished experiment, Brooks build two-dimensional term matrixes by gathering related terms for each term of a hierarchy. An example is presented in Figure 1. The primary hierarchy runs from "Logic" down to "Differential Equations". For each term in the hierarchy, related terms, and then related terms of the related terms, etc., were found. An example horizontal term linkage runs from "Logic" outwards to "Comprehension". As before, just the first listed related term was used, unless it involved a return to a term already used.

Table 9: Relevance Aura for Top Records

Vertical Semantic Steps	Horizontal Semantic Steps 0	1	2	3
0	 **1.04**	**.15**	**.13**	**-.08**
1	**.21**	-.18	-.37	-.54
2	**-.20**	-.57	-.41	-.39
3	**-.40**	-.31	-.34	-.63
4	**-.41**	-.23	-.31	-.33

Relevance Aura: Shaded Relevance Assessments
Semantic Distance Model: Bold Relevance Assessments

Table 9 presents the relevance assessments of 27 graduate students of engineering and 27 graduate students of library and information science for top records. There is a semantic distance effect extending downwards four steps; there is a semantic distance effect extending horizontal three steps. The positive relevance assessments are shaded. This relevance aura extends one step downwards and two steps horizontally. None of the interior matrix terms were considered relevant.

Table 10: Relevance Aura for Bottom Records

Vertical Semantic Steps	Horizontal Semantic Steps 0	1	2	3
4	.71	.05	-.12	-.33
3	**.45**	-.13	-.02	-.48
2	**.65**	.05	-.25	.00
1	**.83**	.65	.22	-.20
0	**1.21** 	**.98**	**-.02**	**-.04**

Relevance Aura: Shaded Relevance Assessments
Semantic Distance Model: Bold Relevance Assessments

Table 10 presents the relevance assessments for the engineers and librarians for the bottom records. There is a semantic distance effect extending upwards three steps; there is a semantic distance effect extending horizontal three steps. The positive relevance assessments are shaded. This relevance aura extends upwards three steps and horizontally one step. Several of the interior matrix terms have been considered relevant. A comparison of tables 9 and 10 reveals that a these data are also showing a semantic direction effect.

Further analysis on a record by record basis will explain the scatter visible in Table 10. An example of this micro analysis is illustrated by Figure 1. Presented is the record itself, which is from the ERIC database (Record number: EJ515482). Both the results for the fourteen engineers and the fourteen librarians that assessed the record are given for each term in the two-dimensional matrix. The terms that were positively assessed are shaded.

“Logic” is the topically relevant term, the one given to the record by an indexer. The engineers find satisfaction in this term, as well as “Logical Thinking”. The librarians disdain all of the terms, even the topically relevant one. One can imagine these librarians fishing this record out of the ERIC database and exclaiming with disgust “Who indexed this?” The engineers, who may be less familiar with this record, seem to agree with the indexer. This example illustrates a systematic clash between core meanings of text and personal interpretations of groups of users/searchers. It is a prime example for future research in relevance assessment.

Figure 1: Relevance Assessments for ERIC Record EJ515482 TM519186

TI: The Logic of Educational Policy.
AB: Some arguments of E. Callan for the need for common education and the importance of early training are well-founded, but his conclusion that limited separate education may be advisable is not. This discussion on the logic of educational policy defends a common schooling to obtain a common education.

Engineers (N = 14)

 Logic .14	Logical Thinking .19	Abstract Reasoning -.20	Comprehension -.07
Mathematical Logic -.88	Computational Linguistics -1.05	Automatic Indexing -.87	Automation -1.06
Mathematical Formulas -.87	Computation -.56	Algorithms -.68	Mathematics -.27
Equations (Mathematics) -.75	Functions (Mathematics) -.69	Mathematical Concepts -.68	Conservation (Concept) -.55
Differential Equations -.98	Calculus -.79	Analytic Geometry -.88	Algebra -.74

Librarians (N = 14)

 Logic -.05	Logical Thinking -.13	Abstract Reasoning -.23	Comprehension -.05
Mathematical Logic -.98	Computational Linguistics -.77	Automatic Indexing -1.04	Automation -.96
Mathematical Formulas -.93	Computation -.81	Algorithms -.79	Mathematics -.76
Equations (Mathematics) -.86	Functions (Mathematics) -.96	Mathematical Concepts -.72	Conservation (Concept) -.56
Differential Equations -.89	Calculus -.95	Analytic Geometry -.94	Algebra -.81

FURTHER RESEARCH

The foregoing presentation has sketched the major features of the SDM. Obviously there is a lot of experimentation and analysis yet to do. But this sketch raises a number of interesting research avenues that are outlined below:

Quantity of text The SDM has been developed with subject descriptors and trimmed bibliographic records. How would increasing or decreasing the amount of text affect the SDM? I would venture that as text is decreased, the results should approximate the reports in the cognitive psychology literature that have compared similarity among words.

Contingency of SDM The effects of semantic distance and direction are contingent psychological effects based on reading in context. What are the factors that will provoke the SDM? The degree of subject expertise appears to be one factor that will sharpen the semantic distance effect. Another is the

context of comparison of the terms, the example being help-menu terms that do not show a SDM effect when contrasted to more verbose classification headings.

Linearity Is the decline of relevance assessment a linear decline with semantic distance?

Statistical Significance Between Steps The SDM as outlined above merely describes a pattern of declining relevance assessments with semantic distance. What impact on a subject hierarchy would result from requiring statistically significant different steps?

Normative Use of the SDM Can the SDM be employed normatively to test the usability of subject hierarchies by grooming terms until a sample of potential users react with a strong SDM effect?

ACKNOWLEDGEMENTS

I am pleased to acknowledge Allyson Carlyle, Sydney Pierce and Edmund Mignon who read an earlier draft.

REFERENCES

Atkinson, R. (1984). The citation as intertext: Toward a description of the selection process. *Library Resources & Technical Services*, 28, 109-119.

Atkinson, R. (1989). The role of abstraction in bibliography and collection development. *Libri*, 39, 201-216.

Bachman, C. W. (1973). The Programmer as navigator. *Communications of the ACM*, 16, 653-658.

Bertrand-Gastaldy, S., Lanteigne, D., Giroux, L. & David, C. (1995). "Convergent theories: using a multidisciplinary approach to explain indexing results." In T. Kinney (Ed.), *Proceedings of the 58th ASIS annual Meeting: Vol. 32. Forging New Partnerships in Information* (pp. 56-60). Medford, NJ: Information Today.

Broadbent, D. E. & Broadbent, M. H. P. (1978). The allocation of descriptor terms by individuals in a simulated retrieval system. *Ergonomics*, 21, 343-354.

Brooks, T. A. (1995a). People, words, and perceptions: A phenomenological investigation of textuality. *Journal of the American Society for Information Science*, 46, 103-115.

Brooks, T. A. (1997). The relevance aura of bibliographic records. *Information Processing & Management*, 33, 69-80.

Brooks, T. A. (1995b). Topical subject expertise and the Semantic Distance Model of Relevance Assessment. *Journal of Documentation*, 51, 370-387.

Chen, H., Houston, A. L., Sewell, R. R. & Schatz, B. R. (1998). Internet browsing and searching: User evaluations of category map and concept space techniques. *Journal of the American Society for Information Science*, 49, 582-603.

Cole, C. (1994). Operationalizing the notion of information as a subjective construct. *Journal of the American Society for Information Science*, 45, 465-476.

Dervin, B., Nilan, M. (1986). Information needs and users. *Annual Review of Information Science and Technology*, 21, 3-33.

Das Neves, F. (1997). "The Aleph: A tool to spatially represent user knowledge about the WWW docuverse." In Berstein, M.; Carr, L. & Osterbye, K. (Eds.) *Proceedings of the Eighth ACM International Hypertext Conference* (pps. 197 - 207). New York, NY: Association for Computing Machinery.

Furnas G. et. al. (1983). Statistical semantics: Analysis of the potential performance of key-word information systems. *Bell System Technical Journal*, 62, 1,753-1,806.

Green, R. (1995). Topical relevance relationships. 1. Why topic matching fails. *Journal of the American Society for Information Science*, 46, 646-653.

Harter, S. P. (1996). Variations in relevance assessments and the measurement of retrieval effectiveness. *Journal of the American Society for Information Science*, 47, 37 - 49.

Houston, J. E. (Ed.). (1990). Thesaurus of ERIC descriptors. Phoenix, AZ: Oryz Press.

Jacob, E. K. & Albrechtsen, H. (1997). "Constructing reality: The role of dialogue in the development of classificatory structures." In *Knowledge Organization for Information Retrieval: Proceedings of the 6th International Study Conference on Classification Research* (pp. 42-50). The Hague, Netherlands: International Federation for Information and Documentation.

Kreigh, R. J., Pesot, J. F. & Halcomb, C. G. (1990). An evaluation of look-ahead help fields on various types of menu hierarchies. *International Journal of Man-Machine Studies*, 32, 649-661.

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

Mizzaro, S. (1997). Relevance: the whole history. *Journal of the American Society for Information Science*, 48, 810-832.

Pierce, B. J., Parkinson, S. R. & Sisson, N. (1992). Effects of semantic similarity, omission probability and number of alternatives in computer menu search. *International Journal of Man-Machine Studies*, 37, 653-677.

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

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

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

Schvaneveldt, R. W., Durso, F. T., & Mukherji, B. R. (1982). Semantic distance effects in categorization tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 8, 1-15.

Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. Science, 237, 1317-1323.

Shoben, E.J. (1976). The verification of semantic relations in a same-different paradigm: An asymmetry in semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 15, 365-379.

Snowberry, K., Parkinson, S. & Sisson, N. (1985). Effects of help fields on navigating through hierarchical menu structures. *International Journal of Man-Machine Studies*, 22, 479-491.

Spink, A. & Griesdorf, H. (1997). "Partial Relevance Judgments During Interactive Information Retrieval: An Exploratory Study." In C. Schwartz and M. Rorvig (Eds.), *Digital Collections: Implications for Users, Funders, Developers and Maintainers, Proceedings of the 60th Annual Meeting of the American Society for Information Science* (pp. 111 - 122). Medford, NJ: Information Today.

Straw, S. B. (1990). "Challenging communication: Readers reading for actualization." In D. Bogdan and S. B. Straw (Eds.), *Beyond communication: Reading comprehension and criticism* (pp. 67-90). Portsmouth, NH: Boynton/Cook Publishers.

Thorngate, W. & Hotta, M. (1990). Expertise and information retrieval. *Knowledge, Creation, Diffusion, Utilization*, 11, 237-247.

Wallmannsberger, J. (1991). "The Scholar's workplace in the global village: from online services to unlimited docuverses." In Raitt, D. I. (Ed.) *Online Information 91, 15th International Online Information Meeting Proceedings* (pps. 239 - 252). Oxford: Learned Information.