6

Many paths to theory: The creative process in the information sciences

The creative process, by its very nature, is unpredictable and surprising. Nonetheless, one can develop skills that will promote and enhance creativity, and will increase the likelihood of producing fruitful ideas. My present purpose is to say something about this process, based on my own experience. Also, I will describe several research ideas that I think are promising and that I wanted to develop but did not have the time to pursue before my retirement. I hope that readers will draw on both the skills and the ideas presented here to produce further progress in the information sciences.[1]

Part 1: Being open to, and using, ideas

First skill: learn to be open to ideas

First, be open to ideas and research possibilities. It is here, at the beginning, that many people get stopped before they really get started. It helps to recall Sigmund Freud's psychological concepts of the id, the ego, and the

1 At this time of fluidity in disciplinary boundaries, I use several terms to describe the fields to which these ideas apply, depending on the orientation and emphasis of the work described. "Library and information science (LIS)," "information science," and "the information sciences" are all used here. I discuss these field distinctions in several other publications (see, e.g., Bates 1999, 2007a, 2015).

First published as Bates, M. J. (2016). Many paths to theory: The creative process in the information sciences. In D. H. Sonnenwald (Ed.), *Theory development in the information sciences*. Austin: University of Texas Press.

superego in the human mind. The id is the child, the autochthonous root of behavior, unpredictable and seemingly uncontrollable. The ego is the adult manager, seeing you through life safely—judicious and mature. The ego does experiment, but cautiously and thoughtfully. The superego is the controller, the guilt-tripper, the part of your mind telling you to follow social rules and religious edicts.

All of these parts of the mind develop in their own way and time, and work together, more or less well, to produce the acting adult that one becomes. Ideally, when one is doing research, the id produces ideas, the ego manages the ideas productively for the benefit of the individual and society, and the superego insures that the work is done and reported on ethically.

Unfortunately, what happens more often than not is that the superego clamps down on the idea production by the id, along with all the other things in the id that the superego suppresses, as people grow from childhood to adulthood. We have all seen the "uptight" individual, afraid to try anything, super-polite and tightly regulated. In this age of looseness and ease about so many things, we tend to pity these people and wish they would tell that superego of theirs to lighten up. However, other people who are not uptight generally and are quite normal in behavior can still be uptight when it comes to idea generation. It is not so easy sorting out just how you do and do not want that id to be allowed to influence your life. Unfortunately, as we "put away childish things," we often put away as well the wonderful fecundity of the id.

I have observed a number of researchers in my life who, I believe, would never ever allow a stray idea to wander up from their id into their ego as they work. In these people, *everything is controlled*; perhaps a better term would be *locked down*—so tightly that a fleeting creative thought would never be allowed into the thinking mind. The result is a predictability, a dryness, a purged-of-color-and-personality quality to their work. If there is an obvious interpretation, they will find it. They are not just analytical; they are analytical/boring. In most cases analytical is good; analytical/boring is not.

Now, in order to let in ideas from the id, one needs to be confident one can handle the ideas and not make a fool of oneself. This is usually one of the biggest reasons for suppressing that id. But here is where we can apply a really valuable idea, and that is this: keep in mind that when you are coming up with wild, silly, ridiculous ideas, *no one else ever has to know.* You have complete control over what you do and do not say or write. You can try out ideas, play with them, even write about them, but, still, no one else need ever know, if you conclude that these ideas are non-starters.

In other words, do not slam the door on the ideas from the id, but let them in, sit down with them, play with them, and then decide whether you want to do anything further with them, such as taking them out for a visit with the rest of the world.

So how do you become aware of the ideas from the id (or from anywhere in your creative psyche)? At first they may come to you as just fleeting—the sort of thing you ordinarily knock out of your mind without thinking. Well, do not knock them out. Stop and think. Why does this occur to me now? What connection can there be between this goofy thought and the research I was just thinking about? The idea that floats into your mind often constitutes an analogy. People have asked me where I got the idea of "berrypicking" for my paper of that name (Bates, 1989). But moving around through different information sources, getting a bit here, a bit there, reminded me of picking huckleberries in the forest with a boyfriend of mine when I lived in Washington state. I allowed that thought in, instead of dismissing it.

Another thing to keep in mind is that creativity requires fecundity, abundance. In her lifetime, a woman produces several hundred eggs, and their associated menstrual cycles, and men produce billions of sperm, all in order to produce just the one or two or three children you ultimately have. Think of how much you and your spouse produced, in order to have just a few children. Ideas are like that, too. You will have *many* ideas in order to produce the few that you actually concentrate on. There may be a thousand ways to think about a particular issue, but only three ways to solve it. To have any hope of solving it, you need to think of a lot of ideas, not just one or two.

Once you allow ideas in, you should have many; more and more should start coming, once you are open to them. It should not be necessary for you to think that every good researchable topic you come up with must be the result of years of striving to find and shape just one or two ideas. If so, you are doing something wrong—such as not letting those ideas in when they first appear. Generally, research does not do well on a paucity model. After all, research and the development of theoretical concepts require multiple big and little moments of creativity to reach fruition. What may look like a single idea at the end is actually often the result of many successive original thoughts and insights.

And if an idea does not work out, do not beat yourself up about it. You are not *wrong* because an idea does not work out. The idea just did not work out. Most of them do not. Move on. That is another advantage of having lots of ideas. You have plenty to spare.

Second skill: Draw on a variety of research traditions

I often encounter doctoral students who, with a laser-like focus, want to know exactly and only the courses they need to take to get through the doctoral program in the shortest time possible. Given the expense, and the time away from other work and family, this is an entirely understandable sentiment, but, in fact, I do not think this approach is productive for the rest of one's career. Doctoral work should be a time to explore intriguing areas, to take courses that may not be obviously related to your work. This should be a time to trust, and to follow up on, that tickle of interest you feel in anything from whole other disciplines, specialties within disciplines, or simply a research question that has been studied by someone in one or more other fields—or our own. To pursue those interests often involves taking or auditing courses in other departments, or taking not-obviously-relevant courses in your own department, or doing an independent study course. Again, your unconscious mind may be directing you toward research and theory that you may not only have a talent for, but which you may also be able to combine in new and creative ways with other knowledge that you already have.

I emphasize this idea of wide exposure to research areas and topics because I believe that this is one of the most productive ways to gain new insights and identify intriguing new topics for research. Each specialty and discipline necessarily brings a whole set of assumptions and established knowledge with it; these are the "paradigms" (Kuhn, 2012) that we hear so much about, which are particularly important in the social sciences. When you are exposed to several different paradigms in your studying, you may well see important inconsistencies or conflicts between those ideas. Thinking about these conflicts can point to new questions that neither paradigm is addressing. There can also be a valuable synergy between the different approaches. It is not uncommon for two disciplines to address essentially the same topic, but from their separate perspectives—and often without being aware that the other discipline has a long line of research on that very question.

Here is a case from my own life history. I wrote about "information search tactics" (Bates, 1979b) and "idea tactics" (Bates, 1979a) after bringing together ideas from several fields. I originally became interested in techniques of information searching when I found myself to be rather incompetent and slow-witted in reference practice exercises in my library education program. A doctoral student serving as a teaching assistant for our reference lab class took fiendish delight in crafting test reference questions whose answers could not be found in any of the "obvious" places. Other students were much quicker in finding answers to these difficult

questions than I was. One day, my friend found answers to twenty-two questions in an hour; I found answers to just two! I was an excellent student in general; why was I so poor on this core librarian task? I soon found that there was little instruction available on this topic in the library literature. Ability to do good reference work seemed to be just assumed by everyone. Note the research opening here: a topic that had not been much addressed in the literature, and yet that was crucial to professional performance as a librarian. Note also that I did not slink away in shame at not being as good as my classmates at this skill; instead I took it as an interesting challenge to find helpful ideas on the matter.

I gathered together everything I could find in the library literature, and then I branched out to psychology, because thinking of other places to look for hard-to-find information was a kind of creative process, and I knew that psychology addressed creativity. I had long been interested in military history and had read books on the subject and liked to watch war movies. So I was aware of the concepts of strategy and tactics. It made sense to me that one should have a strategy for the overall search, and then, as one made various moves to complete the search, one would apply various helpful tactics within the search to increase the chance of success in finding the desired information. So, improbably, I brought military theory together with psychology to address a topic in library and information science. Another word for helpful tactics is "heuristics," a term that cropped up in various fields such as psychology, computer science, and operations research. (I had taken courses in all those areas, by the way, as a doctoral student.) So looking at that concept was productive as well.

The insight about tactics and strategies to promote effective searching provided the framework for my thinking on steps to improve searching. Once I had that idea, I explored all the ways I could think of that would help promote good searching, ending up with several broad categories of tactics: monitoring, file structure, search formulation, term, and so on (over thirty tactics in total). Note that I also drew on my knowledge of librarianship, specifically, the organization of information, in suggesting tactics, for example, that involved experimenting with different terms for a subject—going broader, narrower, and so forth. Thus, several literatures were brought to bear on the topic. These ideas also proved productive for several subsequent articles I wrote (one of which received the Association for Information Science and Technology Best *JASIST* Paper of the Year Award) and led to others in the field doing research on tactics (e.g., Hsieh-Yee, 1993; Xie, 2000).

Incidentally, around the time I was working on these papers, I had an experience that further supports my argument about the value of exposing

yourself to varied fields when you are researching a question. In the late 1970s, I was teaching at the University of Washington, and used to sit in on some of the lectures of visiting speakers in the Philosophy Department. One time, Paul Grice, a very eminent philosopher of language and reasoning, came to speak. What he described in his lecture was his developing work on logical reasoning. He distinguished between formal reasoning and the informal tricks we use to get to an answer to what we are reasoning about. He said he was currently working on the informal reasoning aspect. During the question period, I asked him if he would say that he was working on a conception of the *heuristics of reasoning.* "Well," he said, "'heuristics' is not exactly a term that comes tripping off my tongue." Clearly, he did not know the term. This was a classic example of different fields addressing the same topic under different names. This is not to question in any way Grice's very important and original contributions to this area in philosophy. But might his thinking have changed in any way had he been aware of the several other fields also studying heuristics, each from its own distinctive perspective?

So far, I have presented the idea of drawing on a variety of research traditions as a way to enrich your own work and generate interesting new ideas. But in some cases, the work in related fields is so highly relevant that to *not* address that related work constitutes intellectual failure. A prime example of relevant material typically not referenced in our field is, ironically, the work on relevance by a psychologist and an anthropologist. Dan Sperber and Deirdre Wilson's book *Relevance: Communication and Cognition* (Sperber & Wilson, 1995), originally published in 1986, has generated an enormous amount of interest and debate in many social science fields, as well as in philosophy, and has been cited about 1,500 times, according to the Thomson Reuters Web of Science (wokinfo.com). Only a handful of those citations, however, are from the field of information science, despite the hundreds of articles that address the subject of relevance in information science. Two information scientists who have recognized the significance of the Sperber and Wilson work, and incorporated it into their thinking, are Stephen Harter (1992) and, more recently, Howard White (2007a, 2007b). Otherwise, however, despite occasional references, information science seems to have ignored this high-impact work. Given the recent popularity of "soft," more humanistic approaches to library and information science (LIS), it is puzzling that their subtle approach has not been generally taken up in the field. As long as we in information science like to feel that we "own" the idea of relevance, we had better deal with the many implications of the work on this topic elsewhere in academia. The longer we avoid it, the more behind-the-curve we will appear to be when we finally do take it seriously.

Third skill: Read deeply, not just widely

As I have argued, familiarizing yourself with research traditions in several fields can lead to productive new ideas and approaches in your thinking. However, this suggestion comes with a caveat: if you want to use the ideas drawn from other disciplines in your own work, take the time to really understand what you are reading. As noted earlier, one of the things that makes these encounters between disciplines productive is that the disciplines take different approaches to the same questions. But you cannot just borrow the vocabulary, discuss your own field's work using the new vocabulary, and be successful in making a contribution. The conflicts between the ideas as pursued in the different disciplines often lie at a deeper level than merely the vocabulary. In short, when you study a different approach from another field, you need to understand that approach all the way down, that is, all the way to the underlying philosophical perspective driving that field. Different worldviews and philosophical assumptions separate the various fields. If you want to contribute to the information sciences using those models from other fields, you have to do some serious thinking about how to adapt and integrate those ideas with those of your home field. You need to understand deeply the other field's worldview and how it plays out in the specific topic of interest before you can successfully adapt it to (one of the several) information science paradigms operating (Bates, 2005).

Jenna Hartel's (2010a) work on the use of ethnographic methods in the information sciences represents a recent example of reading deeply and working hard to integrate a methodology widely used in the social sciences with the needs of information-seeking research. Ethnographic techniques are well developed in anthropology and sociology, and have been used by some researchers in the information sciences. What Hartel has contributed to this important methodological philosophy is a deep understanding of information seeking, keeping, and using. The ethnographic methods are not native to the information sciences, and the information orientation is not well known to the other social sciences using ethnographic methods. Her interest in the extensive information gathering and collecting of passionately devoted gourmet cooks surprised some of her sociology professors. As sociologists, they had not ever thought much about *the ethnography of information.* Bringing those two together is a difficult task, requiring extensive reading in philosophy and theory, and years of analytical work. Hartel described her experiences in a series of articles (Hartel, 2010a, 2010b, 2011) that are combinations of research results and expositions of ethnographic methods as applied to the information worlds of the people being observed.

Fourth skill: Relate your work appropriately to that of others

One of the standard requirements of an academic article is that the author should show how the work fits in with other research and thinking in the area. This is not an idle requirement of fussy journal editors; it is done for a reason. Work that is published should not only be new, but also be *shown* to be new. One does this by describing prior work and explaining how the current work relates to the prior work and yet advances beyond it. The literature review in an article (1) *contextualizes the research* presented in the article, so the reader can see where the work fits with other work in the field; (2) demonstrates that the author is *current with other work in the area,* and so is not rediscovering things already known; (3) *affords recognition* to those authors for the earlier work; and (4) *claims its own new territory.*

So the literature review, which may be seen superficially as a boring prelude to the real action of the article where the research is described, is actually a socially very significant part of the article, and needs to be thought about carefully.

I bring this up here in a chapter on creativity in the information sciences because the creativity that produces new ideas or methods can come into conflict in a variety of ways with providing recognition to others for their work. How much credit is due to my thinking, and how much credit is due to others for their thinking?

Why is recognition important? The currency of business is the bottom line—the amount of money made. No matter how clever your product or service, if you do not make money on it, you are not a successful businessperson. The currency of science, on the other hand, is *recognition* for work done and discoveries made. People go into research to discover new things and advance science or scholarship. The mark of success, therefore, is to have other talented people, who also know that research area well, recognize and esteem your work, and value your presence in the field. Esteem is shown, for academics, through successful publication in good venues, through promotions, through awards, through naming things after their discoverer, and so on. It is all about recognition for one's talents by one's respected colleagues.

I got a lesson in this importance early in my career. I was writing the article to be later titled "Rigorous Systematic Bibliography" (Bates, 1976). After the article was, for the most part, finished, I was reviewing it and realized suddenly that I had drawn throughout the article on the work of Patrick Wilson, one of my professors at UC Berkeley, where I got my doctorate. In effect, the article represented a way to operationalize, or put into practice, the much more theoretical ideas to be found in Wilson's book *Two Kinds of Power* (Wilson, 1968). I had so thoroughly absorbed Wilson's ideas

that I did not recognize that I was working through methods of applying his ideas to the practice of bibliography in the article. I had not submitted the article yet, and, somewhat embarrassed, I rewrote it to include full and repeated references to his work. After a bit, I realized that I had gone too far, and was not highlighting what was original in my own contribution in the article, so I rewrote it again, with what finally felt like a good and accurate balance.

The relationship between our own work and that of others can be a very touchy matter, and it is better to recognize that, and work it through, rather than deny what is happening. One makes a good-faith effort at reviewing the earlier literature, and one cites it when it is relevant. I suspect that the intense hostile animus that shows up in some journal article reviews arises from the reviewer feeling that you have somehow violated, wittingly or unwittingly, his or her intellectual turf. Perhaps I take this matter of giving credit where credit is due a little too seriously, but I have seen many occasions where people did not deal well with the challenge of getting the credit right.

One time a senior professor wrote to me enthusiastically about how much he liked one of my articles and how influential it had been for him. Later, he wrote a whole book on the subject, and I looked to see the role my work played in it. He cited the reference to my paper seriously incorrectly, and, to read the relevant section, you would think my work had virtually no effect on his, and was at best peripheral to his work. In another case, I was asked to review an article for a double-blind journal (i.e., neither author nor reviewer knows the identity of the other) in which the author had used almost the identical research design and research question as I had in a major research project, but had not related it to my study, or referenced it. I pointed to the articles that I had published out of my study, and asked that they be appropriately referenced and related to the results of the work in the article being reviewed. Later, after the article had been published, I saw that the senior professor who wrote it now cited my article but still did not reference or discuss any of the parts that were actually overlapping. The reader would have no idea of the similarities.

I have read many articles that used my work but did not reference it or that referenced my popular article "The Design of Browsing and Berrypicking Techniques" (Bates, 1989) instead of referencing the article of mine that actually dealt with the content of the referencing article. By referencing an irrelevant article of a person, recognition appears to have been given, when, in fact, it has not. The new work presented in the article is not connected at all to the actual relevant prior work that should have been discussed in the article. One reviews prior work in order, among

other things, to compare the results and discuss any differences. By citing a different article than the relevant one, one avoids addressing altogether the actual relationship between the current and cited research. This happened to me so many times that I once sat down, with the help of the Thomson Reuters Web of Science, to review every reference to my work in the literature in order to tally the actual frequency of this problem. However, not wanting to be the skunk at the garden party and offend my professional colleagues, I gave up on this endeavor.

It can be disappointing to discover that someone else has been there before you—but they did get there, and you owe it to them to acknowledge it. Further, if the works differ in results, some discussion of this is in order; such discussion and debate is what science is supposed to be about. If your own work is any good, you will also be contributing work of value; it does not undermine your own contribution to recognize those who have gone before.

Let us turn now to those promising ideas to be put forward. I will have more to say about being creative in information science research as I go along.

Part 2: Promising ideas

In this section, I discuss a number of ideas that I think have promise, and that I am unlikely to be able to pursue in my lifetime. I encourage students and established researchers to take up these ideas and see where they lead. Keep in mind that these are *ideas*; they are not fully researched or substantiated. There will be gaps, unreferenced literature, and so on. Actual in-depth work with these ideas might lead to different conclusions than I anticipate. In short, much remains to be done, but most of these should have the potential to be good dissertation or other research project topics.

Area 1: Information seeking at the human and machine interface

Here I will describe just three of many topics that could be mentioned as promising ideas to follow up in the area of information seeking at the human and machine interface. However, to introduce those ideas, I need to pick a fight—just a little one—with my colleagues, in this book and elsewhere, from the field of human-computer interaction (HCI). I believe that the full contribution that information science can make to this area is generally not recognized outside the field, particularly by researchers in HCI. Many in HCI come from a psychology background, and when they address the experience of people at a computer interface, they bring a psychologist's desire

to discover fundamental principles about human psychology in carrying out the work of using computers. The objective is to understand the whole HCI experience; it is assumed that the most fundamental discoveries will apply across most or all application areas. From that standpoint, *information seeking* on computers is just a single application area, and therefore not of much intrinsic interest to HCI researchers.

I would argue, in reply, that information seeking is a much larger area than generally recognized and has distinctive, important features that need to be understood and designed for in order to produce the best HCI in information-seeking contexts. General HCI principles are not enough for optimal results.

Gregory Bateson (1968) made an interesting distinction between what he called "value seeking" and "information seeking" (p. 178–179). In value seeking, a person has an idea in mind and goes out into the world to shape that world so as to produce a result that matches the idea in mind. If one wants bacon and eggs for breakfast, one does certain things in the world with pigs, chickens, and a stove, with the end result that one has a plate of bacon and eggs. In information seeking, on the other hand, one goes out into the world to discover things so as to create the idea in mind to match or reflect what is in the world. This is a simplified distinction, obviously, but a powerful one. The nature of our actions in the world will be very different when we have a plan in mind to impose on the world, versus when we are open to the world imposing some part of its character or shape on our own minds. If we move to a part of the world where bacon and eggs are not to be found, and we are hungry, we face a very different type of challenge in our effort to learn about things that might satisfy our hunger.

Information seeking is not just an application field, like engineering, or retail services; rather, it is the other side of action. It is about gaining knowledge, and there are many characteristics of that behavior that distinguish it from value seeking. First, information seeking makes one vulnerable. Information, by many definitions, is surprise. Not all surprises are fun or desirable. Therefore, the seeker opens him- or herself up to risk, to the potential need to reorganize or reorient his or her hard-earned knowledge. The consequences of this risk ripple throughout the behavior we call information seeking. The behavior ranges all the way from avoiding information to actively seeking it out in cases where our lives are on the line or our passions are engaged with a fascinating subject.

Second, by definition, *you are seeking something that you do not know.* How, then, does one specify the sought information? Whole courses in information science are devoted to this question. Not only the designs of classifications and indexing vocabularies are at issue, but also courses

on the interaction between people and information. We know quite a bit about what people do to their own queries to make them understandable to information systems and to information professionals. These propensities have huge consequences for the design of the information-seeking interface in information systems.

Some years back in the Department of Information Studies at the University of California, Los Angeles (UCLA), one of my colleagues taught a course on HCI in information systems, and I taught a course on user-centered design of information systems. By then, information systems were overwhelmingly online, so one might expect quite a bit of overlap in the contents of the two courses. But there was, in fact, virtually no overlap, because I emphasized the design issues that were specific to information seeking/searching, and she taught about HCI in general. *There is a distinct body of knowledge in information science around information seeking on computer interfaces.* Approaches assuming that general HCI knowledge applies without an understanding of the information-seeking part are underperforming in the quality of design for this central human process. My objective in the following sections is to describe some HCI areas involving information searching and retrieving that have interested me, and that I believe would be of value to pursue further.

Topic 1: Design for real browsing What is generally called a "browsing capability" in online systems is nothing of the sort. Overwhelmingly, as currently manifested, this capability in online systems consists of being able to *scan* down the page, with the help of the scrolling function. That systematic, top-to-bottom or left-to-right scanning behavior is not browsing. Think about standing in a bookstore or at a magazine stand, or shopping in a bazaar, for that matter; it is not about systematic scanning. Your eyes dart all around. You glimpse here, then there, then way over there. Things catch your eye. You look at one, then at another. If something is really interesting, you pause and take a serious look at it, then select it, or move on. Rarely do you run your finger along the books, or items in the bazaar, in a systematic way, studying one, then the one right next to it, then the one right next to that, and so on. The eye darts around in browsing; it does not scan, as I have argued in detail (Bates, 2007b). Things catch your eye because you first do a gross glance that does not analyze the visual scene in detail; then you put your closer attention to the things that pass your crude filter (Wolfe, 1994).

I have argued that "browsing is seen to consist of a series of four steps, iterated indefinitely until the end of a browsing episode: (1) glimpsing a field of vision, (2) selecting or sampling a physical or informational object

within the field of vision, (3) examining the object, (4) acquiring the object (conceptually and/or physically) or abandoning it. Not all of these elements need be present in every browsing episode, though multiple glimpses are seen to be the minimum to constitute the act" (Bates, 2007b). Thus, design for browsing in online systems would necessarily be very different from the current provision of the capability to scan. First, the screen needs to be large enough to allow the eye to take glimpses, attending to one part, then to another part. Second, there should be a variety of options available to the user—and not just on pull-down menus, which require choosing to pull down the menu! Instead, the many options should be available *at the same time* on the screen—just as a scene glimpsed by our forebears in the forest contained many points of interest. Icons scattered on the screen, representing different types of search capability, could simultaneously present to the user a rich array of options for searching. A single search box, à la Google, has been viewed as the sine qua non for searching—the easiest, most simplified approach. But what if we found a way to make the search itself interesting? Let people browse through different capabilities, and different classes of metadata or taxonomies. There are many ways to structure such a system; the problem is that *we have not ever taken seriously the desirability of designing for true browsing.*

Topic 2: Interface design specific to searching For a very long time, there has been a pervasive assumption among computer scientists that an information system search interface can simply be superimposed on any body of searchable data. Set up the information so that various elements can be searched, then superimpose some sort of search engine—most any kind will do—and you have your information system. But, as I argued in "The Cascade of Interactions in the Digital Library Interface" (Bates, 2002), there are, in fact, many layers of design needed beneath and in front of the interface that culminate in design imperatives for the resulting interface. The information itself—its content, structure, medium, types of indexing and metadata, and so on—influences how one can best find desired results within the body of information, and therefore how search should be designed in the interface for the user, as well as behind the scenes in the computer. Likewise, everything on the searcher side of the interface—subject area of interest, type of query, level of skill as a searcher, and so on—interacts with the interface design in a productive or unproductive way. Bad design at just one of these several layers can block the effective functioning of all the other layers (Bates, 2002).

Here is an example contrasting two situations. In situation A, you have a database of biological reports addressing the various concerns of

biologists, from studying animals in their natural habitat to using parts of nature for research in other areas, such as agriculture, animal breeding, and environmental concerns. In situation B, you have a database of articles published in the humanities literature on the several fields encompassed by that term—national literatures, philosophy, religion, and so forth.

Extensive research and practical experience have demonstrated that faceted vocabularies provide the most appropriate means for indexing humanities articles (for the explanation, see Bates, Wilde, and Siegfried, 1993). The Getty Research Institute (associated with the Getty Museum) consequently committed itself to creating faceted vocabularies, using them to index its extensive database production. In contrast, scientific databases do best with the one- to three-word phrases known as descriptors, which have been prompted by classical indexing theory and embodied in the technical standards developed for them (National Information Standards Organization, 2005). With faceted indexing, the whole query is composed of terms drawn from each of several distinct facets, while with scientific literature, the search is composed of descriptors presented to the system in Boolean combinations, either implicitly or explicitly. If searchers are to take advantage of these differences, the interfaces must be designed differently for the two systems—a single simple search box, in particular, will not do.

Examples of this sort could be proliferated indefinitely. If the one database contains technical reports with a certain standard structure in introductory matter (author, affiliation, abstract, etc.), and the other database consists of humanities articles using typical bibliographic rules for humanities articles (e.g., University of Chicago Press rules rather than American Psychological Association rules), then there are different ways of coding and representing these bibliographic entities that can either promote successful search or get in the way of it. And I have not even mentioned other types of information, such as video, image, multi-lingual, and other variant forms of data.

When designing for search, one must design the interface all the way down to the actual content itself. A single, standard search interface, superimposed on the huge actual variety of types of information and information organizational schemes, can just about be guaranteed to sub-optimize. Furthermore, there is a huge variety of types of queries characteristically associated with different types of needed information. For example, art databases are searched not only by art history scholars but also by designers and artists looking for inspiration in the images they find, and by schoolchildren for their assignments. The kinds of search interface design features that each of these groups could most benefit from

differ. Different capabilities are needed for each. To optimize information search, all these various design layers need to be recognized, understood, and designed for in an interface that nonetheless feels simple and natural to the end user. This is a high standard, and one that has been mostly ignored outside of information science. There are abundant opportunities here, with the right funding and sufficient imagination, to see how complex much of information searching actually is.

Topic 3: Question and answering—the 55 percent rule Up through the 1990s, there was considerable interest and research in library and information science on what was called the "55 percent rule." Researchers addressed the question of whether the answers provided at library reference desks were accurate. This was usually done by sending testers to library reference desks or by calling library reference departments to ask typical reference questions of the librarians, then checking the accuracy of the answers provided. To the researchers' horror, in study after study, the accuracy rate came out, surprisingly, to a fairly consistent figure of around 55 percent. How could this be? Responding to reference queries is the bread and butter of a substantial portion of professional librarians, the so-called public services staff, and to compile an accuracy rate so low was concerning, to say the least. Why did study after study get results around 55 percent?

I first got interested in this question when asked to review a book by Frances Benham and Ronald Powell (1987) that reported results of two separate such studies. The bottom-line results, the accuracy rates for answering questions in the two studies, were 52.73 and 58.73 percent—remarkably similar, considering variations in the sampling of the two studies (Benham & Powell, 1987, p. 136). (There is a large literature on this subject, which will not be reviewed here.)

Something about the consistency of these results across many studies troubled me, however. If training in the reference interview and teaching reference librarians about the typical reference resources available to provide answers to these questions really were the chief influencing factor in performance, then why was there not *more* variation? Surely, smaller libraries with poor reference collections and, probably, staff with less or no professional training should produce poorer results, and the larger, better staffed and stocked libraries should produce much better results. Yet the patterns were very similar.

Then I learned about the US Internal Revenue Service call sites, where people can call the agency to inquire about aspects of the tax laws. Studies had been done over several years, and there was some controversy about the

nature of the sampling and test questions; these questions were resolved through a thorough review of methodology. By 1989, with good, verified methodology, the test call survey report stated: "IRS' overall. . . results for the 1989 tax filing season showed that IRS telephone assistors responded correctly 62.8 percent of the time to the survey's tax law test questions" (US General Accounting Office, 1990, p. 1).

That result rang a bell for me. *If this completely different context could produce a result so similar to that for libraries, then the problem was likely due to something other than simply inadequate library resources or professional training.* As I explored the literature around this question, I came to suspect that the problem concerned any information question-asking situation that was at all complex. In other words, I was forming the hypothesis that in *any* situation where there is a lot of detail or context necessary to fully understand the issue, there will be a 55 percent general rate of accuracy. Indeed, another of the reports on the IRS tests states: "For questions that required IRS assistors to probe callers for more information in order to sufficiently understand the question, the accuracy rate was 56 percent compared to 90 percent for questions where probing was not required" (US General Accounting Office, 1987, p. 2). *I hypothesized that this is a general human communication problem, not a library-specific problem*.

This rate is a source of embarrassment wherever it crops up, because it seems so low. I suspect, however, that the error rate is high because it is impractical for both questioner and assistor or librarian to carry out the necessary amount of probing needed to insure that the question is actually understood in all relevant respects, and can therefore be answered correctly. The questioner has in mind a vast amount of context surrounding the question, and, often, has no way of knowing which particular elements of that context the assistor needs to know in order to produce a correct answer. On the other hand, the assistor does not know that context, and may not realize that some key invalidating characteristic of the context might, if the assistor knew about it, change the assistor's response to accommodate it. There is probably no easy answer to this situation, because to fully insure that all factors have, in fact, been taken into account might require a very long interaction. Most of the time—say, about 55 percent of the time—that extra long interaction would be wasted or unnecessary, because the initial response would be correct, but the rest of the time, a much higher investment might be needed to provide higher accuracy. (This is not to suggest, however, that more complex questions may not, in themselves, and apart from context, be more subject to error, on the part of both assistor and requester, just due to their complexity.)

My research assistant and I did quite a bit of literature searching around this question, looking at informational interactions in law and medicine, as well as in general. There is a whole subliterature about inaccuracy and failure to communicate in physician-patient interactions alone! I planned to devote part of the summer of 1994 to explicating this whole issue of question-asking accuracy. Unfortunately, I was not able to follow up, because that spring the campus administration proposed shutting down our school, and two miserable years followed, wherein I served as department chair as we fought back to keep the program alive, though, ultimately, as a part of a new, combined entity: the Graduate School of Education and Information Studies.

On the basis of what I had found in the literature, and on what, to me, seemed the uniting themes of that research, I planned to further research the hypothesis about information-seeking interaction that summer, and would probably conclude by making the argument stated above. I felt that one particularly telling argument would be this: notice that in cases before a court, where absolute accuracy in all respects is the goal of the questioning of witnesses, a long string of questions is posed to the witness, probing, in numbing detail, every aspect of the situation being discussed. "Was the stop sign visible from where you stopped? Were there any leaves or tree branches blocking your line of sight to the sign? Was the overhead streetlight shining on it?" And on and on and on. The practice of law has demonstrated the need to go far beyond the basic initial interaction in order to get all the relevant facts, to get the details needed to push the answer beyond the point of being just 55 percent accurate. Indeed, even when this probing is done, there may be still much more to the story than comes out in court (Finnegan, 1994).

The relatively poor accuracy results were embarrassing for librarians and for the IRS. I think it is important to realize that in dense, complex question-asking situations, this accuracy problem is probably built into the nature of human knowledge and interaction. We have no hope of improving the librarian success rate without attributing the problem to its likely true sources in general human communication, rather than to library training and resources alone. This topic needs to be further explored, argued, tested. . . .

Area 2: Information organization: Event indexing

At its very heart, all indexing theory is built around the core idea that we index *nouns, or conceptual objects.* As the technical standard for thesaurus

development states, an "indexing term" is "the representation of a concept in an indexing language, generally in the form of a noun or noun phrase" (National Information Standards Organization, 2005, p. 6).

I wish to challenge that core assumption to produce a result that can still fit within the larger framework of indexing but that might enrich the possibilities available now. Indexing is generally described as an effort to identify and flag the "subjects" of a document. Those subjects are, implicitly, the topics "covered" or "discussed" in the document. As I write, workers for the San Francisco Bay Area Rapid Transit, or BART, are striking to get more favorable contract terms. Conventionally, a newspaper index would index articles on this event as "strikes and lockouts," "labor unions," or "collective bargaining." These are all noun phrases and certainly represent the topics addressed in the articles, as per our usual understanding of indexing.

But a strike is also an *event* that takes place in time and spools out through time. A little bit of reflection will suggest that while an event can be considered a topic of discussion, if we really want to understand the event, and be able to search for it, should we not develop a way of indexing that is more true to the nature of that event?

Here is one way to think about it. Compare *narrative* and *expository* writing. A *narrative* is a story that takes place through time. It may be fiction or nonfiction. *Exposition* of a topic elaborates and explains that topic. *Standard indexing theory is designed for exposition, not narrative.* Indeed, the question of how to index fiction has been raised many times, and there are no easy answers regarding either the ease of indexing fiction or the value of doing so (Beghtol, 1994; Pejtersen, 1978; Saarti, 1999).

However, here I want to stick to questions of indexing nonfiction, as in newspaper and magazine indexing. Compare topic indexing and event indexing (see table 2.1). In the *topic-indexing* condition, we are addressing a document or document portion devoted to expounding on a topic. The bulk of the text involves description and explanation. If the text is reasonably coherent, it can be considered to be addressing one or more particular topics. Classical indexing attempts to identify those topics and provide good, consistent descriptions of them through the development, control, and application of appropriate index terms.

Now, what do we have in the *event-indexing* condition? First, we have narrative, which is the telling of a story of some sort, spooling out in time. Furthermore, the narrative describes, above all, a *situation.* A situation is a complex of conditions, circumstances, events, and actors existing at a point in time, and often, through time. Within the context of a situation, one or more events may occur. An event is an occurrence, a complex of

actors and circumstances that change one condition to another within a noticeably short time frame.

In the example of the Bay Area transit strike, the *situation* is that of contract negotiations between the labor union and BART management. The strike is an *event* that occurs during the contract negotiations.

Now, suppose we are a newspaper needing to index the ongoing events associated with the contract negotiations. We create a template specifically for situations and events. We use the classical aspects for event description that have been identified in journalism: who, what, where, when, why, how. *When* should include duration, that is, beginning and ending times.

Thus, within the larger *situation,* we code for various *events* that occur within it (table 2.2). In indexing, the newspaper creates a situation name and assigns a code number to that situation. At the beginning of contract negotiations, an indexer fills in the names of the parties—the union and the management—notes the dates, location, and so on. Then, as the negotiations continue through various events, the coding and names automatically populate over to the next event, except where the event itself changes things, and then those features are changed by indexers.

Since events are about *things that happen, that is, action,* we should explore what indexing might look like if *verbs* were actually indexed, too.

TABLE 2.1. *Contrasting topic and event indexing*

TOPIC INDEXING	EVENT INDEXING
Exposition	Narrative
Topic	Situation
Description/explanation	Event

TABLE 2.2. *Situation and event coding for indexing*

SITUATION: BART CONTRACT NEGOTIATIONS 2013			CODE: 54025
EVENT	INITIATE TALKS	GO ON STRIKE	RETURN TO TALKS
Who			
What			
Where			
When			
Why			
How			

So, verb indexing terms might be words like "initiate [talks]," "strike," "settle," "conclude [talks]," and so on. Alternatively, a separate set of index terms that are nominalized verbs could be carved out to be used specifically to index events. These would be terms like "initiation of talks," "strike," "settlement," "conclusion [of talks]," and so forth. For a newspaper, or for legal, law enforcement, or medical records, having a separate action category of verbs or of verb-based nouns might be useful for identifying contents and enabling discovery by searchers. For example, an ongoing lawsuit is between two parties, and the case name and parties involved will be scattered throughout the records of the suit. But if one is looking for a particular action—say, to discover when a request was made to enjoin someone from acting—then a search on "enjoin" or "injunction" might be helpful. Such event-indexing terms might be extracted from existing thesauri or created anew as specifically verb index terms to draw attention to the sequence of actions that characterizes events and situations that spool out through time. In a certain sense, events have been masked, or submerged, within the broader scope of classical subject indexing. The approach suggested here highlights the distinctive features of events and situations, that is, of *narratives,* as distinct from *exposition.*

Area 3: Information densities

The term "information densities" actually covers a number of possible areas to pursue. I will develop them here as well as I am able, given that I have not actually pursued this area in depth.

Topic 1: Information whomps In March 1977, Marilyn Levine published an article in the *Journal of the American Society for Information Science* called "The Informative Act and Its Aftermath: Toward a Predictive Science of Information." Levine argued that important events in a society have an impact that can be quantified in (among other ways) the number of books written on that subject or event. The event has an emotional, economic, political (etc.) impact, and people respond by adapting to the event, discussing it, and reorganizing their thinking and their lives in order to accommodate the event and move on. One of the important ways that human beings do all these things is to produce new information expressing their reactions, ideas, and solutions. They then share these ideas with the rest of society through publications, which promotes the absorption and integration of the collective experience into society and people's thinking. Today, unlike in 1977, we see some of this reaction process through the immediate production of "tweets" and other forms of brief communication. But the process of

reacting occurs in a deeper and more thorough way as well through the publication of books, articles, and other written communications.

Levine compared the resignation of President Nixon to the resignations of a New York City mayor and of several state governors. She calculated the information "whomps," or the collective impact of the news, as a multiple of the number of people affected in the country, state, or city by the resignation times the level of stress associated with the resignation and the single hard bit of information, namely, that the resignation had occurred. Obviously the resignation of a US president affects more people than the resignation of a governor or mayor. She then looked at the number of books that appeared in subsequent years in the then-standard listing, the *Cumulative Book Index* (H. W. Wilson Company, 1898–1999; the index ceased publication in 1999), and she found that the number of books subsequently appearing about each of the events was roughly proportional with the calculated number of whomps. In other words, societies reflect the impact of important events through publication; the higher the impact, the more publications that result.

Now, I hasten to point out that, in my opinion, there are many methodological and theoretical problems with Levine's article. The initial literature review is superficial and does a poor job of linking together the various theories and bodies of research that she cites. Further, there are methodological issues with how she counts whomps. One might argue that a resignation of a state governor affects more than the citizens of that state, and challenge her basis for calculating the amount of hard information and the degree of stress associated with each type of event. In short, there are a lot of issues with this article, and I doubt that it would be published today in this form. But I am glad that the editor (Art Elias) published the article, because the core idea is a powerful one, and deserving of more attention.

The discussion and disputation of ideas and events is core to any literate society. It seems to be a reasonable hypothesis that the amount of writing and discussion around certain events or issues would be roughly proportional to their importance to our social discourse. Why is there no bibliometric science today focusing on the measurement and discussion of the number of information whomps associated with various events, or debates about why a particular event produced fewer or more publications than one might have expected? Surely, this kind of measurement would be of interest to historians, social scientists, and information scientists. (We may be returning to these questions through another route, for example, the growing attention recently to discovering the significance of various trends in social media. See, e.g., Mike Thelwall's work on Twitter [Thelwall, Buckley, & Paltoglou, 2011; Thelwall et al., 2010].) Thousands of articles have

been produced on scientific reputations as measured by number of citations. Though our academic egos are no doubt involved in such questions, surely the study of publication rates as reflections of societal upheaval, stress, and progress is of at least as great importance?

The situation with Levine's idea illustrates another point about creativity in research. Yes, the ideas could have been better developed, and perhaps the colloquialism "whomp" put some people off. But this scrappy little diamond in the rough of an idea is still a diamond—waiting (for the past thirty-five years!) for someone to take it seriously and develop it well and thoughtfully. The lesson: learn to differentiate the several parts of an issue or topic, and select out and develop the good parts. Do not just reject the whole package because some parts of it make you uncomfortable. Derek J. de Solla Price (1975, 1986) studied some basic statistics about publication in the history of science and came up with a remarkable range of important new knowledge about the nature of science and scientific publication patterns. Levine had the great insight to introduce the concept of whomps; whomps are still waiting for their Price for full development.

One final point: I am sure that many people would reject this topic out of hand because of the age of the article. In some contexts, materials that are more than a few months or years old are derisively rejected as being hopelessly out of touch. That is a common way that people unnecessarily limit themselves. In fact, some older things are out of date, and other older things are highly relevant, and can still be the stimulus for further creative developments. One must develop discrimination and selectivity, not just wave something away because it is not currently trendy. *Look for the intrinsic value, not just the current fashion, then reshape the material for the present.* Given the current social media context, and the constant attention to "trending" topics, Levine's idea seems even more relevant—and possibly more easily measured—today than it was at the time of original publication. Even with the demise of the *Cumulative Book Index*, one can still measure the societal impact of various issues and events in book publication, as well as in many other media. Take the earlier idea and adapt it to the current context. Bingo! You have a new area of research.

Topic 2: Information investment As many in the information sciences know, Fritz Machlup (1972) argued that a very large portion of all human economic activity centers on the production and use of information. The book blurb on the publisher's website for his book *The Production and Distribution of Knowledge in the United States* summarizes it nicely: "Machlup's cool appraisal of the data showed that the knowledge industry accounted for nearly 29 percent of the U.S. gross national product, and that 43 percent

of the civilian labor force consisted of knowledge transmitters or full-time knowledge receivers. Indeed, the proportion of the labor force involved in the knowledge economy increased from 11 to 32 percent between 1900 and 1959—a monumental shift" (Princeton University Press, 2013).

Robert Hayes, professor and former dean at the Graduate School of Library and Information Science at UCLA, produced some research in the early 1980s that I have long felt has much more potential than has actually been developed. A much more recent discussion of this area can be found in Koenig and Manzari's (2010) encyclopedia entry entitled "Productivity Impacts of Libraries and Information Services." Hayes and Erickson (1982), according to Koenig and Manzari, used

> the Cobb-Douglas production function to estimate the value added by information services. In the basic Cobb-Douglas formula, the value of goals [*sic:* goods?] and services sold is calculated to be the product of a constant times the values of different inputs, labor, capital, and so forth, each raised to a different power (exponent). The exponents are solved for by seeing which exponents best fit a number of separate cases. In the Hayes and Erickson formulation, the value added V in manufacturing industries is a function of labor L, capital K, purchase of information services I, and purchase of other intermediate goods and services X. (Koenig & Manzari, 2010, p. 4311)

Hayes and Erickson found that those industries that thrived best over the period he studied were those, such as pharmaceuticals, that had very much larger investments in information relative to the other Cobb-Douglas factors than other industries had. Koenig and Manzari note that Yale Braunstein continued the work, making some modifications in the formula, yet still came out with much the same result—that there is "substantial underinvestment in the purchase of information" (Koenig & Manzari, 2010, p. 4311).

I leave it to economists to explicate these ideas—but if the premise holds true that there is huge underinvestment in information in the economy, this should have huge implications for many disciplines, including ours. The argument should be researched and developed further, and the results published in the economics literature, as well as the information sciences literatures.

Topic 3: Distilled information There is yet another sense in which we seldom discuss information, yet that may have interesting implications for

theory in the information sciences as well as, ultimately, for practice. This sense concerns the idea of information as a distilled product of thought and communication.

Books, journal articles, newspaper articles, blog entries—even micro-messages such as "tweets"—are all different forms of work products. Often, they have serious economic value or economic implications. Someone put some brainpower into the creation of them. Furthermore, having created the text, images, and so on, human beings have engaged in various forms of economic production to make them available, whether investing in a printing press and paper or in a computer server to hold this work product ready for access on the Internet.

Both the intellectual investment and the physical investment made so these products will be available represent the information economy in the Machlup and Hayes and Erickson senses. We invest a substantial portion of both our cognitive and physical energies as a society in creating, distributing, and maintaining (in libraries, on the web) these intellectual products. The quantity and quality of the work that goes into these information products varies tremendously, of course. But the same thing about quality and quantity of work could be said of retail clothing, drugs, or many other products of human activity.

But let us focus particularly on these intellectual products. An academic book, for example, may represent a decade, or a lifetime, of research, study, analysis, and writing. We may think of the resulting book as a distillation of all that massive amount of work by an individual. The encyclopedia that I recently edited with Mary Niles Maack (Bates & Maack, 2010) had about seven hundred authors, each of whom spent some substantial chunk of time in the preparation and writing of journal-article-length entries. I was retired and spent four years full time as editor-in-chief of the encyclopedia, while Maack spent many weeks and months as co-editor, taken from her otherwise hectic schedule as a professor, to produce the resulting seven-volume, 5,742-page encyclopedia. On the usual rule of thumb of two thousand hours of work for an individual for a year, the editing alone took five or more person-years. If each of the 565 articles took a month of person-time, which may be an underestimate, forty-seven or more years of person-time went into the creation of the encyclopedia entries. The resulting seven-volume set, taking up a little more than a cubic foot of volume, represents over fifty full-time person-years of human thought, writing, and editing effort.

In fact, each book or other information product can be thought of as such a distillation. *In the network of human relationships and social activity, each document is a node of distillation, a point of intensification of human*

labor and intellection. We have developed means of condensing much thought into small packages, and we store those nodes of information all around us. Picture one of those economic atlases that shows the countries of the world, with clear plastic layers that may be laid over the map, each layer representing something of economic interest, such as agricultural production, manufacturing production, and so on. Lay over that map an additional layer that shows the location of information stores. If this is done, then server farms, libraries, bookstores, websites, and many more resources all become points of great intensification of information in the economic map of society.

In my view, the huge amount of energy invested in bibliometric studies of various kinds should include the study of the distribution of these information densities in society, followed by the study of the amount of use that is made of them, where they are underutilized according to formulas such as Cobb-Douglas, and so on.

Concluding thoughts

The role of information in the economy and in social relations is so integral to all we do that it is like the air we breathe. Sometimes we do not see it at all because we take it so much for granted. But it has great meaning and social impact in terms of the mental and physical processing and organizing of our world. There are so many more ways to think about and study the role of information than we have fully engaged in; we have the prospect of huge disciplinary development in the information sciences if we can only start to see the (informational) air we breathe.

REFERENCES

Bates, M.J. (1976). Rigorous systematic bibliography. *RQ, 16*(1), 7–26.

Bates, M.J. (1979a). Idea tactics. *Journal of the American Society for Information Science, 30*(5), 280–289.

Bates, M.J. (1979b). Information search tactics. *Journal of the American Society for Information Science, 30*(4), 205–214.

Bates, M.J. (1989). The design of browsing and berrypicking techniques for the online search interface. *Online Review, 13*(5), 407–424.

Bates, M.J. (1999). The invisible substrate of information science. *Journal of the American Society for Information Science, 50*(12), 1043–1050.

Bates, M.J. (2002). The cascade of interactions in the digital library interface. *Information Processing & Management, 38*(3), 381–400.

Bates, M.J. (2005). An introduction to metatheories, theories, and models. In K.E. Fisher, S. Erdelez, & L. McKechnie (Eds.), *Theories of Information Behavior* (pp. 1–24). Medford, NJ: American Society for Information Science and Technology.

Bates, M.J. (2007a). Defining the information disciplines in encyclopedia development. *Information Research, 12*(4), paper colis29. Retrieved from http://InformationR.net/ir/12-4/colis/colis29.html.

Bates, M.J. (2007b). What is browsing—really? A model drawing from behavioural science research. *Information Research, 12*(4), paper 330. Retrieved from http://InformationR.net/ir/12-4/paper330.html.

Bates, M.J. (2015). The information professions: Knowledge, memory, heritage. *Information Research, 20*(1), paper 655. Retrieved from http://InformationR.net/ir/20-1/paper655.html.

Bates, M.J., & Maack, M.N. (Eds.). (2010). *Encyclopedia of Library and Information Sciences* (3rd ed., 7 Vols). New York: CRC Press.

Bates, M.J., Wilde, D.N., & Siegfried, S. (1993). An analysis of search terminology used by humanities scholars: The Getty Online Searching Project report no. 1. *The Library Quarterly, 63*(1), 1–39.

Bateson, G. (1968). Information and codification: A philosophical approach. In J. Ruesch and G. Bateson (Eds.), *Communication: The social matrix of psychiatry* (pp. 168–211). New York: Norton.

Beghtol, C. (1994). *The classification of fiction: The development of a system based on theoretical principles.* Metuchen, NJ: Scarecrow Press.

Benham, F., & Powell, R.R. (1987). *Success in answering reference questions: Two studies.* Metuchen, NJ: Scarecrow Press.

Finnegan, W. (1994). Doubt. *New Yorker, 49*(48), 48–67.

Hartel, J. (2010a). Managing documents at home for serious leisure: A case study of the hobby of gourmet cooking. *Journal of Documentation, 66*(6), 847–874.

Hartel, J. (2010b). Time as a framework for information science: Insights from the hobby of gourmet cooking. *Information Research, 15*(4), paper colis715. Retrieved from http://InformationR.net/ir/15-4/colis715.html.

Hartel, J. (2011). Visual approaches and photography for the study of immediate information space. *Journal of the American Society for Information Science and Technology, 62*(11), 2214–2224.

Harter, S.P. (1992). Psychological relevance and information science. *Journal of the American Society for Information Science, 43*(9), 602–615.

Hayes, R.M., & Erickson, T. (1982). Added value as a function of purchases of information services. *Information Society, 1*(4), 307–338.

Hsieh-Yee, I. (1993). Effects of search experience and subject knowledge on the search tactics of novice and experienced searchers. *Journal of the American Society for Information Science, 44*(3), 161–174.

Koenig, M., & Manzari, L. (2010). Productivity impacts of libraries and information services. In M.J. Bates and M.N. Maack (Eds.), *Encyclopedia of library and information sciences* (3rd ed., pp. 4305–4314). New York: CRC Press.

Kuhn, T.S. (2012). *The structure of scientific revolutions* (4th ed.). Chicago, IL: University of Chicago Press.

Levine, M.M. (1977). The informative act and its aftermath: Toward a predictive science of information. *Journal of the American Society for Information Science, 28*(2), 101–106.

Machlup, F. (1972). *The production and distribution of knowledge in the United States.* Princeton, NJ: Princeton University Press.

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

Pejtersen, A.M. (1978). Fiction and library classification. *Scandinavian Public Library Quarterly, 11*(1), 5–12.
Price, D.J. de Solla. (1975). *Science since Babylon.* New Haven, CT: Yale University Press.
Price, D.J. de Solla. (1986). *Little science, big science—and beyond.* New York: Columbia University Press.
Saarti, J. (1999). Fiction indexing and the development of fiction thesauri. *Journal of Librarianship and Information Science, 31*(2), 85–92.
Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition* (2nd ed.). Malden, MA: Blackwell.
Thelwall, M., Buckley, K., & Paltoglou, G. (2011). Sentiment in Twitter events. *Journal of the American Society for Information Science and Technology, 62*(2), 406–418.
Thelwall, M., Buckley, K., Paltoglou, G., Cai, D., & Kappas, A. (2010). Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology, 61*(12), 2544–2558.
US General Accounting Office. (1987). *Tax administration: Accessibility, timeliness, and accuracy of IRS' telephone assistance program* (GAO/GGD-88-17). Retrieved from http://www.gao.gov/assets/210/209820.pdf.
US General Accounting Office. (1990). *Tax administration: Monitoring the accuracy and administration of IRS' 1989 test call survey* (GAO/GGD-90-37). Retrieved from http://www.gao.gov/assets/220/211963.pdf.
White, H. D. (2007a). Combining bibliometrics, information retrieval, and relevance theory, Part 1: First examples of a synthesis. *Journal of the American Society for Information Science and Technology, 58*(4), 536–559.
White, H. D. (2007b). Combining bibliometrics, information retrieval, and relevance theory, Part 2: Some implications for information science. *Journal of the American Society for Information Science and Technology, 58*(4), 583–605.
Wilson, P. (1968). *Two kinds of power: An essay on bibliographical control.* Berkeley, CA: University of California Press.
Wolfe, J.M. (1994). Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review, 1*(2), 202–238.
Xie, H. (2000). Shifts of interactive intentions and information-seeking strategies in interactive information retrieval. *Journal of the American Society for Information Science, 51*(9), 841–857.