

LIBRARY AND INFORMATION SCIENCE

Series Editor: Amanda Spink

Faculty of Education at the Queensland
University of Technology, Australia

Recent and Forthcoming Volumes

Amanda Spink and Jannica Heinstrom
Trends and Research: Europe

Dirk Lewandowski
Web Search Engine Research

Amanda Spink and Diljit Singh
Trends and Research: Asia-Oceania

Amanda Spink and Jannica Heinstrom
New Directions in Information Behaviour

Eileen G. Abels and Deborah P. Klein
Business Information: Needs and Strategies

Leo Egghe
Power Laws in the Information Production Process: Lotkaian Informetrics

Matthew Locke Saxton and John V. Richardson
Understanding Reference Transactions: Turning Art Into a Science

Robert M. Hayes
Models for Library Management, Decision-Making, and Planning

Charles T. Meadow, Bert R. Boyce, and Donald H. Kraft
Text Information Retrieval Systems, Second Edition

Charles T. Meadow
Text Information Retrieval Systems

A. J. Meadows
Communicating Research

V. Frants, J. Shiparo, and V. Votskunkii
Automated Information Retrieval: Theory and Methods

LIBRARY AND INFORMATION SCIENCE

LOOKING FOR INFORMATION: A SURVEY OF RESEARCH ON INFORMATION SEEKING, NEEDS, AND BEHAVIOR

THIRD EDITION

EDITED BY

DONALD O. CASE

*School of Library and Information Science,
College of Communication and Information Studies,
University of Kentucky, Lexington*

Series Editor: Amanda Spink



United Kingdom • North America • Japan
India • Malaysia • China

Chapter 1

Information Behavior: An Introduction

What you *don't* know has power over you; knowing it brings it under your control, and makes it subject to your choice. Ignorance makes real choice impossible.

— Abraham Maslow (1963, p. 116)

Beyond obsessions, curiosity, and creativity, lies a host of motivations *not* to seek information.

— David Johnson (1997, p. 70)

Chapter Outline

1.1. Introduction	3
1.1.1. A Bit of Vocabulary	4
1.1.2. Emphasizing People Rather Than Systems	5
1.1.3. Ten Myths About Information and Information Seeking	7
1.1.4. When, Why, and Where Information Behavior has been Studied	10
1.1.5. The Contexts in Which Information Behavior Is Investigated	13
1.1.6. The Scope of “Information Behavior”	14
1.2. How This Book Is Organized, and How to Use It	15
1.2.1. Organization of the Chapters	15
1.2.2. Which Chapters to Read If...	16

1.1. Introduction

This volume describes common and essential human behaviors: seeking and using information. Noticing a change in the weather, deciding to visit another city, finding out about travel schedules, choosing a departure date, and buying an airline ticket are examples of a range of activities known as “information behavior.” These include accidental encountering, needing, finding, choosing, using, and sometimes even avoiding, information. They are types of behaviors that are basic to human existence.

This introductory chapter describes the scope of the book and its contents. It says briefly what kinds of concepts, questions, and research have

been developed regarding information behavior, and why this topic has attracted attention. I make the case that the nature of this research has changed over several decades, away from an emphasis on institutional sources and searches, and toward a focus on how individuals encounter and make sense of their environment.

The Internet could serve as a metaphor for information behavior and the way our view of it has changed. Think back to a time before the World Wide Web was available. All of the information was out there in individual books, journals, radio and TV programs, offices, filing cabinets, minds, and computers. But because it was divided by source, by location, by person, and by channel, it was not always easily located or examined. Making arrangements for *travel* is one comprehensive example: One could hear the weather forecast on the radio, read about a destination in a travel guide, call hotels to make reservations, telephone an airline to learn departure times and fares, visit a travel agent to pick up a ticket, and so on. In terms of research, each of those needs and transactions have needed to be conducted (and studied) separately. But now it is possible to satisfy all travel-related requests on a single website. Not only have the different channels of communication collapsed down to one, but less goal-oriented behaviors, such as browsing, may play a larger role than ever before. Looking for information becomes more holistic.

The contrast between new and old is even greater when we compare tasks in the office and classroom to their counterparts of 20 years ago. Obscure bits of information — the text of a government regulation, the date of an event, the author of a document — are more easily found in a single “place” — the Web. Both work and education have changed as a result.

In a manner similar to the emergence of the World Wide Web, our view of information behavior has become more integrated and less dictated by sources and institutions. As what we know about these behaviors has grown, so has the vocabulary used to describe it.

1.1.1. A Bit of Vocabulary

In introducing the subject matter of this book I will be using terms like “information,” “information need,” “information seeking,” and “information behavior” without defining them fully until later chapters. For the moment let us assume that there *are* such things as “information” and “information needs” that can be satisfied by “information seeking” or “information behavior.” To tide us over until these concepts are fleshed out, here are some brief definitions:

- *Information* can be any *difference* you perceive, in your environment or within yourself. It is any aspect that you notice in the pattern of reality.

- An *information need* is a recognition that your knowledge is inadequate to satisfy a goal that you have.
- *Information seeking* is a conscious effort to acquire information in response to a need or gap in your knowledge.
- *Information behavior* (hereafter, “IB”) encompasses information seeking as well as the totality of other *unintentional* or *passive* behaviors (such as glimpsing or encountering information), as well as purposive behaviors that do not involve seeking, such as actively *avoiding* information.
- *Information practices*, a term more popular in Europe and Canada than the United States, may be thought of as a synonym for information behavior — although it maintains some differences that will be explored in a later chapter.

The most commonly discussed of these concepts is *information seeking*. It is a behavior so commonplace that it is generally not an object of concern until time pressure makes it so. If we are making a major decision (e.g., buying a house) or finishing a task by a deadline (e.g., writing a report), we might find ourselves in an earnest information seeking mode: talking to others, searching the Web, reading magazines, watching the news, and so on. We may do everything we can to satisfy our desire for input, until either our need is satisfied or we have run out of time. More commonly, it is the latter, as the demand for “information” is usually elastic — there is always more that one could know. After our need is met (or we give up) we return to a more passive state of information seeking, at least as regards the object of our earlier curiosity.

Consider also cases in which the acquisition of information does *not* concern an immediate task like buying or writing something. Our daily life is peppered with instances in which we become interested in learning more about a topic after accidentally encountering some bit of information about it. This sort of curiosity, unmotivated by an immediate goal, is a common aspect of human life.

The situations described above, no matter how familiar to all of us, are much more complex than they may appear on the surface. Information seeking behavior often escapes observation. It is difficult to generalize about behaviors that vary so much across people, situations, and objects of interest, and which often take place inside a person’s head. This book is about the many ways in which information seeking has been defined, explicated, observed, and measured in studies of human behavior.

1.1.2. Emphasizing People Rather Than Systems

Systematic research on information seeking — at least on the use of sources like books or newspapers — dates back nearly a century. In the first three

decades of the twentieth century, studies of information “channels” and “systems” — chiefly libraries and the mass media — accumulated slowly. The 1940s saw the first published reviews of this literature. By the 1960s, such investigations, particularly of the specialized information needs and uses of scientists and engineers, were appearing regularly in a variety of journals and reports.

But much of this older literature was really not about information seeking in the sense in which that concept is discussed in current research. Rather, most of the investigations focused on the *artifacts* and *venues* of information seeking: books, journals, newspapers, radio and television broadcasts, schools, universities, libraries, professional conferences, and the like. What was actually studied were the information *sources* and how they were used, rather than the individual users, their needs (as they saw them), where they went for information, and what kind of results they experienced. Surveys of individuals made such strong assumptions about their needs, motivations, habits, and behaviors that the range of responses they could make was severely constrained; what mattered in these early investigations was how *formal information systems* served the serious (e.g., work, health, or political) information needs of the population studied. Typically this literature was called “information needs and uses” research, or sometimes “user studies” or “audience research.” Choo and Auster (1993) call this tradition “system-centered” research; Vakkari (1999) refers to it as “system oriented”; a host of other commentators have applied similar labels.

It was not until the 1970s that investigations begin to branch out beyond the focus on formal channels and task-oriented needs. The emphasis shifted away from the structured “information system” and toward the person as a finder, creator, interpreter, and user of information. In mass media research the focus shifted to the “*gratifications*” that users experienced, rather than focusing on “*effects*” that messages had on people and how to persuade them to do things. Even studies of formal information systems began to consider a wider range of people, more general needs and problems, and the ways in which those systems often failed to serve their publics. The term “information seeking” — and, later, “sense making” — began to be preferred in describing the kind of phenomena that interested a growing number of scholars.

Some observers (see, e.g., Vakkari, 1999) have stereotyped the concerns of the old versus the new research on information behavior. Table 1.1 contrasts the person and system orientations by posing some examples of research questions that are typical for each.

The right column in Table 1.1 reflects research questions that have motivated thousands of studies — typically institutionally sponsored evaluations of library use, selective dissemination of information (SDI)

Table 1.1: Contrasting examples of information behavior research questions.

	Person oriented	System oriented
Task-oriented studies	<ul style="list-style-type: none"> • How do lawyers make sense of their tasks and environment? 	<ul style="list-style-type: none"> • What kinds of documents do engineers need for their work, and how might the corporate information center supply them?
Nontask-oriented studies	<ul style="list-style-type: none"> • How does a manager learn about job-related information <i>outside of</i> formal organizational channels? • What happens when a voter has too much information about a candidate or an issue? • How do the elderly learn about and cope with problems or opportunities that come up in their daily lives? • Why do TV viewers choose one program over another, and what satisfactions do they achieve in doing so? • Why do people browse in stores when they have no explicit need in or intention to buy? 	<ul style="list-style-type: none"> • How satisfied and successful are student searches of a university library’s Web-based catalog? • How much use do medical doctors make of medical databases? • How does the public use a library for personal pleasure and growth: what they ask for, borrow, and read? • How do we persuade teenagers to act in healthy and responsible ways? What messages about drug abuse do they attend to, in which medium, and why? • Why do people ignore safety warnings on packages and advertisements?

programs, information retrieval systems, interface designs, information campaigns, advertising effectiveness, and the like. A few of these studies will be discussed in this book, almost exclusively the “nontask-oriented” variety. The left column reflects the emphasis of this volume, and hence, the predominate type of examples used within.

1.1.3. Ten Myths About Information and Information Seeking

A key development in the shift toward more user- or person-centered theories and methods were the questions raised in the early 1970s by several researchers, chief among them Professor Brenda Dervin (Ohio State University). A landmark 1976 article by Dervin encapsulated several years of her work by challenging 10 assumptions that had dominated

research on communication and information seeking up to that time. In her article she was concerned chiefly with the everyday information needs of the ordinary, urban resident. However, much of what she says also applies to more formalized needs. Here are the 10 “dubious assumptions” that Dervin (1976a) identified in past writings about information seeking.

1. *Only “objective” information is valuable.* People are rational beings who process data from the environment to analyze alternatives and make optimal decisions. Several problems plague this assumption, including our common tendency to rely on easily available sources of information such as our friends. For most tasks and decisions in life, people tend to settle for the first satisfactory solution to a problem, rather than the best solution.
2. *More information is always better.* Yet too much information leads to overload and thence to deliberate ignoring of inputs. “Having information” is not the same as “being informed,” so increasing the flow of information does not always result in an informed person. Typically there is not a problem getting enough information but rather with interpreting and understanding what information there is — an internal, rather than an external, locus of control.
3. *Objective information can be transmitted out of context.* But people tend to ignore isolated facts when they cannot form a complete picture of them. Individuals yearn to understand how information connects to other facts, beliefs and emotions they have, and how all these affect one another.
4. *Information can only be acquired through formal sources.* This assumption, often made by those in educational institutions, flies in the face of actual behavior. People use formal sources rarely, instead gathering and applying information from informal sources, often friends and family, throughout their lives.
5. *There is relevant information for every need.* The truth is that mere information cannot satisfy many human needs. People may want information in the sense of learning or understanding or entertainment; more commonly they need the physical and psychological necessities of daily life, such as food, shelter, clothing, money, and love. Information cannot substitute for many human needs, nor even facilitate all of them.
6. *Every need situation has a solution.* Institutions such as libraries, medical clinics, and social service agencies are focused on finding solutions to problems. To do so they attempt to map what the client says — the words they use — onto the resources and responses of their system. But sometimes the client is looking for something — a reassurance, an understanding — that does not come in the shape of a

canned response. Nevertheless, the system will usually provide an answer of some type, in its own language and logic, whether it is useful to the client or not.

7. *It is always possible to make information available or accessible.* Formal information systems are limited in what they can accomplish, at least where the vague, ambiguous, and constantly changing needs of the public are concerned. People will continue to come up with their own answers to their own unique, unpredictable questions without resorting to formal information systems.
8. *Functional units of information, such as books or television programs, always fit the needs of individuals.* Information systems such as libraries or broadcasters define themselves in terms of their units of storage or production: in the case of libraries, these are books, journals, audio-visual materials, or websites; in the case of broadcasters, it is programs, ads, or public service announcements. But the “functional units” of the individual are not often these things; rather, they are responses, solutions, instructions, ideas, friendships, and so forth. Thus, client requests for help, action, or resources tend to be reinterpreted by institutions as information needs that can be fulfilled with the units that they provide: books, programs, and the like. The client cannot always effectively use these units of information.
9. *Time and space — individual situations — can be ignored in addressing information seeking and use.* Yet often it is the individual’s definition of the situation that shapes his or her needs as much as the “real” situation itself. If individuals perceive a lack of predictability and control of an outcome, then they worry. The worry itself becomes a need.
10. *People make easy, conflict-free connections between external information and their internal reality.* We tend to assume an ordered universe, in which connections exist between the internal and external. In our research, we tend to ask “what” and “how” rather than “why.” We ask what people read or view, rather than why they do so. We lack understanding about how people inform themselves, how they make connections *over time*, the sense they make of their world between significant events. Dervin said that instead of studying what “information does ... for people” we need to focus on “what people do to information” (p. 333).

Dervin argued that all 10 of these assumptions were flawed in the ways indicated. Of course, she made these statements about *everyday* information needs, not in the context of highly specific, task-oriented needs like scientific or business data for decision making. There are indeed times when people act mostly rational and optimal in their

information seeking and processing. Those situations, too, will be addressed in this volume.

1.1.4. When, Why, and Where Information Behavior has been Studied

As a subject of scholarly attention, information behavior has been studied in many different contexts, with a variety of people and a broad array of motives and goals. All people seek information, yet for some people and in some situations the stakes are much higher. Higher stakes are more likely to create situations that attract research.

To illustrate the kinds of people and situations that have been investigated over the past five decades, first let us consider several hypothetical cases. The examples below are constrained by several assumptions. These assumptions are valuable because they will help us compare situations by creating a “standard” set of reactions. However, each assumption has limitations, which will be noted. The assumptions are that information seeking behavior is highly *rational* (which is not often true), that such behavior is oriented toward making some kind of *decision* (a common, yet flawed, assumption), and that it is possible to make relatively simple judgments about the *value* of our decisions (itself a value judgment to which some people would object).

Consider the relative importance of three types of situations and information needs, as located on a hypothetical continuum (Figure 1.1). This continuum reflects the number of people ultimately affected by the search for information and subsequent decisions based on it: at one end are trivial decisions affecting few people, whereas at the other are important decisions that may affect millions of human lives.

Now, one could argue that it is anthropocentric to use humans as the sole benchmark for judging the *importance* of a decision. We could easily imagine real-world problems that involved other sentient beings (e.g., animals) or nonsentient things (e.g., tropical rain forests). If we think through the implications of such problems, we may notice that we tend to judge their importance by their ultimate impact on our own feelings or

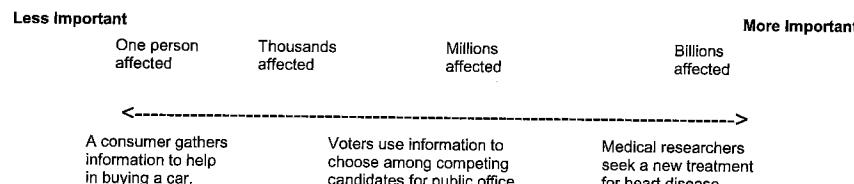


Figure 1.1: A continuum of importance.

well-being as humans. So, while recognizing that we could use other real-world objects as benchmarks, for the purposes of this discussion we will consider “numbers of people affected” as a simple indicator of importance.

First, imagine a person trying to choose between different models of a car that vary by features and price; although there are many publications (both printed and electronic) that offer just such information, the outcome of this search is simply not very consequential. Presumably this situation would fall near the “unimportant” end of our hypothetical continuum because it affects only one consumer.

Second, consider a citizen about to go to the polls, choosing among candidates based on information about their positions and past performance; electing public officials is certainly an important decision, and yet this is just one vote out of many. This situation is more important than the first, but less crucial than others we might imagine.

Third, imagine a biomedical scientist, with years of training and experience, working in an expensive laboratory, developing treatments for heart disease. This scientist must keep abreast of what other researchers are doing in the field, what discoveries have been made, what new equipment and techniques are available, and so forth. (To be more realistic, we could identify just one specific need, such as the answer to the question “What are the effects of dietary fish oil on measures of serum cholesterol?”) Surely the information needs of this person are important, as judged by the investment society has made in the scientist and the potential outcomes of the work. The scientist’s decisions about which research leads to follow could affect millions of people around the globe.

Laying out these hypothetical situations and judging their importance is a precursor to an explanation of why more research has been conducted on some kinds of information seeking and not on other types. We might all agree that the case of the scientist who is working on treatments for heart disease that might affect millions of lives is worthy of study. By studying the information needs of such scientists, and how they go about satisfying those needs, we just might be able to devise a tool or service that would help them reach their research goals a little sooner. In such situations, the potential for public good (and for private profit) is enormous. This is why many of the investigations the reader will encounter in this volume have focused on high-stakes and high-status occupations: research scientists, medical doctors, aerospace engineers, corporate managers, and the like.

And yet many types of information seeking behavior are worthy of study. Sometimes relatively trivial decisions, such as the automobile purchase described above, are the target of expensive investigations due to the cumulative importance of individual decisions. There is an entire industry,

commonly called *market research* that investigates purchase decisions; individual purchases are relatively trivial, yet millions of them add up to significant amounts of money.

We can see a similar logic operating in studies of voting: how a particular individual finds out about issues and candidates may not seem important, but the information-gathering habits of millions of voters may have a crucial impact on a society as a whole. Therefore, there is a sizable literature on political communication, and more specifically on what kinds of information people glean from the mass media. The listening, watching, reading, and learning that takes place in support of buying and voting — and many other daily tasks — is sometimes referred to as “everyday” information seeking. We will learn about that as well in this book.

There is another very important focus of investigation that we have not touched on yet (although the medical study comes close): “basic” research on human information behavior. Ideally, what we would really like to know is how people go about seeking (or avoiding) information in a *generic* way, free of specific contexts like heart disease research or car purchases. Unfortunately, as with other attempts to conduct basic research on human behavior, it is difficult to generalize beyond the specific type of stimulus that prompted the behavior.

There is some doubt as to how deeply researchers can investigate truly basic human behaviors regarding information. Certainly some psychological investigations of perception, human information processing, and pattern recognition are relevant to basic considerations of information seeking. Such studies deal with the fundamental question “What is information?” and are discussed in Chapter 3 of this volume. Suffice to say that, for this book, I am concerned also with a social element: information seeking is *interpersonal* as well as *intrapersonal*.

In the latter sense, perhaps the closest we come to basic research on information seeking are studies of communication in dyads and small groups in laboratory settings. Social scientists conduct these studies to understand how individuals solicit, process, and interpret data and cues they receive from others. Even in closed laboratory settings, the nature of the information itself may intrude. I will say more on this point as we review specific studies.

Finally, one important distinction that is made in the literature on information seeking is between *formal* and *informal* sources of information. The prototypical formal source is a printed one — a textbook, encyclopedia, or daily newspaper — but may also be exemplified by the words of an acknowledged expert on a subject. Informal sources tend to be friends, colleagues, and family, but in the view of some they could encompass what we learn from popular culture as well: TV programs, songs on the radio, Internet discussion lists, Facebook, tweets, and so forth. I will make use of

this formal versus informal distinction in reviewing some of the findings on information behavior.

1.1.5. The Contexts in Which Information Behavior Is Investigated

The previous section raised the issue of context with considerations such as individual situations, motives for seeking information, the specific activities and kinds of information, the surrounding environment, the types of people, and the size of the social group involved in the investigations. There is no nice, neat, logical delineation of these factors, as human behavior itself is not completely rational or uniform. The examples used later in this volume have been selected with an eye toward the literature that actually *exists* — that is the patterns of studies that have been conducted, particularly since 1990.

There are a variety of approaches that we could use to consider the vast literature on information seeking and related topics. I could, for example, review studies chronologically, showing how they shifted in focus and method. Or investigations could be selected on the basis of the discipline in which they were published, whether in information studies, communication, management, medicine, and so forth. In fact, both *historical* (in this chapter and the final chapter) and *disciplinary* (in most of the other chapters) categorizations are sometimes invoked in this volume, but they take a back seat to three other ways of considering the literature: by theory, methods, and context.

What is meant by “theory” and “methods” may be obvious to most readers, but “context” warrants some further explanation (see Chapter 10 for more details). For the purposes of organizing this book, *context* will be taken to mean the particular combination of person and situation that served to frame an investigation. In Chapters 11 and 12 I review information seeking investigations under three general categories: the *occupation* studied (e.g., manager, doctor, social scientist, and chemist), the *social role* of the persons under investigation (e.g., consumer, voter, student, patient, and television viewer), and the *demographic* groupings (e.g., by age, gender, race, ethnicity, and geography). Although a respondent could easily represent an occupation, a role, and a demographic group at the same time, as well as illustrating the use of any number of information sources, investigators typically choose to frame their research questions and respondent samples in terms of one of these three ways. In Chapter 12, for example, I will describe a study of the “urban poor” (a demographic group) that approached its topic by sampling janitors (an occupation). The bulk of investigations that fall under the heading of “information seeking” have concerned the information needs and uses of a specific occupation, role, or demographic group.

1.1.6. The Scope of “Information Behavior”

Information seeking is a topic that has been written about in thousands of documents from several distinct disciplines. Because almost everything to do with humans is potentially relevant to this topic, I have trimmed the scope of this book to highlight aspects of information behavior that have become more important in recent years.

For starters, there are two things that this book primarily is *not* about. I refer to the time-honored topics of “library use” and “information retrieval.” Both of these (particularly the first) do indeed have strong connections to human information seeking, but each topic has a voluminous literature of its own that is really more about systems (continuing documents or computer files) than it is about people. One could also say that these have received too much emphasis in the “information needs and uses” literature. Generally speaking, the research described here is *not* well representative of pre-1980 information seeking research, which tended to focus on the use of libraries and paper documents or databases; I say little about such studies here. Relatedly I pass by the many thousands of studies on learning and the education of students, even though they do involve issues of information acquisition.

I have also narrowed my review by time period. Nearly a quarter of the publications discussed herein (over 300 items) date from 2007 or later; most of the rest were published during the 1990s and 2000s. Although I make citations to some earlier, seminal discussions and definitions of the concepts discussed in this volume, those are merely included to ensure proper credit and historical perspective. Most of the examples and references in this book are taken from the past three decades of published literature. As is emphasized here and in the concluding chapter, recent investigations of information seeking focus more on the *seeker* and less on the *sources* or *channels* they use, although it is not possible to ignore the latter entirely.

I have chosen to highlight certain aspects that have received too little attention from mainstream investigators of information seeking; among these less-examined topics are the connection between entertainment and information; passive and accidental information acquisition; sharing of information among peers; and ignoring and avoiding information.

My examples are taken chiefly from the disciplines of information studies, communication, psychology, and professional fields like management, business, medicine, and public health. The investigations used to explain typical findings or methods are taken from a variety of contexts. I am aiming for a multidisciplinary understanding of the concept of information seeking. I hope I have succeeded in reaching that goal.

1.2. How This Book Is Organized, and How to Use It

1.2.1. Organization of the Chapters

This book consists of 13 chapters. I like to think of these chapters as grouped into five segments:

- One: Introductions and examples (Chapters 1 and 2)
- Two: Concepts relevant to information behavior (Chapters 3, 4, and 5)
- Three: Models, paradigms, and theories in the study of information behavior (Chapters 6 and 7)
- Four: Methods for studying information behavior (Chapters 8 and 9)
- Five: Research results and reflections (Chapters 10, 11, 12, and 13)

I have begun by mentioning several basic concepts: *information*, *information needs*, *information seeking*, *information behavior*, and *information practices*. Each will be explored in more depth in Chapters 3 and 4. I have sketched out the history and scope of the literature I intend to review.

Chapter 2 is my attempt to give information behavior a human face by exploring five examples from the individual’s point of view. Please do not be misled by the simplistic and everyday nature of these five scenarios. They are here because I believe it is important to recognize that information seeking is something we all do in the course of our everyday existence. It is not a domain of behavior restricted to scientists, engineers, physicians, managers, and the like. We should acknowledge it as a common need before we plunge into the explication of the fuzzy concepts that have tended to plague this research.

Beginning Part Two, Chapter 3 (The Concept of Information) explores the vital notion of *information* and analyzes several problems inherent in its definition. The reader may judge that I say far too much about the definition of information — don’t we all know what it is? But I think it is only fair to acknowledge how much commentary this everyday notion has generated. Readers who are new to this literature would be wise to save Chapter 3 for a later time.

Chapters 4 (Information Needs and Information Seeking) and 5 (Related Concepts) continue the review of basic concepts by offering definitions of many other terms frequently invoked in the information seeking literature, such as *decision making*, *browsing*, *foraging*, *encountering*, *sharing*, *selective exposure*, *avoidance*, *overload*, *information anxiety*, *knowledge gap*, *information poverty*, *pertinence*, *relevance*, *entertainment*, and a variety of spatial metaphors (e.g., *grounds* and *horizons*).

In Part three, Chapter 6 (Models of Information Behavior) and Chapter 7 (Metatheories, Paradigms, and Theories) provide general background about models and theories that have been used to study information seeking. Chapter 8 (The Research Process) contains a brief tutorial about methods of investigation. Then Chapter 9 (Methods: Examples by Type) proceeds to explore methods and techniques commonly used in information seeking studies, providing one or more examples of each approach: laboratory experiment, field experiment, mail survey, e-mail or web-based survey, individual and focus group interview, participant observation, diaries and experience sampling, history, content analysis, meta-analysis, and combinations of these.

In Part four of this book, Chapters 10 through 12 identify 14 commonly researched categories of people, and summarize one or more typical studies for each group. Other relevant studies are mentioned in context. First, Chapter 10 explores the history, size, and structure of the information seeking literature. Chapter 11 then examines findings about *occupations* (e.g., doctors). Chapter 12 also reviews individual studies of information seeking, but this time considering investigations of social *roles* (e.g., consumers) and *demographic groups* (e.g., the elderly). Altogether over 100 investigations are cited in Chapters 10 through 12, and 30 are described in detail.

Finally, Chapter 13 summarizes the approaches and findings of the current literature and suggests avenues for future research. The book concludes with a glossary, a collection of questions for discussion, and a bibliography of almost 1400 works cited in the text.

1.2.2. Which Chapters to Read If...

This book could be used in several different ways, depending on the needs and goals of the reader. For those who simply want a quick review of the recent literature on information seeking, Chapters 10 through 12 could be read on their own.

For methods courses in information behavior, Chapters 6 through 9 (covering theories and methods) could be read together, or in combination with Chapters 10 through 13 (reviewing research results). Methodologists interested in the range of concepts that might be measured in investigations should also read Chapters 4 and 5 on relevant concepts; experienced researchers can skip Chapter 8 on basic methods, as this is intended for neophytes.

Finally, students of information seeking, information behavior, and information needs and uses will want to read the book straight through. This text could also be useful in courses on user-centered design of

information systems, information architecture, and the like. The appendixes include several questions for each chapter, which may be used as the basis for classroom discussions or written exercises.

Now let's begin our exploration by looking at some examples of information seeking from the seeker's perspective.

Chapter 2

Common Examples of Information Behavior

Information-seeking must be one of our most fundamental methods for coping with our environment. The strategies we learn to use in gathering information may turn out to be far more important in the long run than specific pieces of knowledge we may pick up in our formal education and then soon forget as we go about wrestling with our day-to-day problems.

— Lewis Donohew, Leonard Tipton, and Roger Haney (1978, p. 31)

Blood donors often ask, "Will I faint?" Cancer patients ask, "Will I die?" citizens facing everyday situations ask, "How long will it take me to handle this?"

— Brenda Dervin (1992, p. 75)

Chapter Outline

2.1. Six Information Seeking Scenarios	20
2.1.1. Buying Products	20
2.1.2. Finding Information in a Library	24
2.1.3. Healing a Patient	27
2.1.4. Betting on Race Horses	30
2.1.5. Finding the Law	34
2.1.6. "I Want to Know More about Cancer"	36
2.2. Summary	39
Recommended for Further Reading	41

In this chapter we will make the case that searching for information is an important part of being human, and it is something that we do on a regular basis. Out of necessity we will encounter concepts and terminology that will be explained fully in later chapters of this book. Consider this chapter to be a preview of what is to come.

Every day of our lives we engage in some activity that might be called information seeking, though we may not think of it that way at the time. From the moment of our birth we are prompted by our environment and our motivations to seek out information that will help us meet our needs.

This chapter will consider six situations in which information seeking behaviors are in full swing. That is, these are scenarios that face millions of

people (at least in developed nations) each year, in which decisions and choices are made that require a great deal of data, information, and understanding. Each will underscore the complexity of information seeking and the strategies we use to make it simpler.

All of the stories here involve not only the search for information but the choice of which data to retain and consider. Five of the tales can be characterized as decision making as well, a narrower type of behavior that is studied in its own right and is not always considered in studies of information seeking.

Let's first consider a very familiar type of activity: shopping.

2.1. Six Information Seeking Scenarios

2.1.1. Buying Products

Few decisions are more common in developed societies than choosing to purchase a product. In our role as consumer we may buy thousands of items a year, mostly foodstuffs, but also intangibles like services and many hundreds of household items. Of the latter, only a few may be considered major purchase decisions: houses, cars, boats, furniture, and large appliances, among others.

From the consumer's perspective it is the expensive, infrequently purchased items that tend to garner the most thought. However, it is important to recognize that many small purchases over the course of a lifetime — such as toothpaste or soft drinks — also amount to large expenditures. The fact that even our most minor needs eventually amount to a great deal of money accounts for the attention paid to the purchase decision from two different perspectives: market research (intended to aid the producers of goods and services) and product evaluations (intended to aid the consumers of these).

On the production side, a great deal of thought goes into the design and especially the advertising of items for sale. Advertising is intended to present (and sometimes inundate) the consumer with reminders of the product's existence and with persuasive information about it. Hence, the marketing of products attempts to minimize the effort a consumer expends to search for information.

Indeed, from the marketer's perspective, the information that is put out in print, on radio and television, on the Internet, and on billboards would result ideally in an entirely knee-jerk reaction: the consumer sees the ads, then sees the product, and then buys the product. For those who make and sell products, it is better that the consumer does not engage in a lengthy search for information, but simply buys the item as quickly as possible.

Except for those few truly and obviously superior products, the producers are likely to prefer that the consumer does not compare brands at all.

Research on these questions is accomplished by a variety of scientists working in industry and in universities; most have training in psychology or business (or both). When their reports are not proprietary, they may be published in the *Journal of Consumer Research*, the *Journal of Marketing Research*, the *Journal of Advertising Research*, or in more general publications.

Product makers are not the only ones who do research. There are other organizations that, for both profit and public service, provide the research and testing that the consumers do not (and indeed often cannot) do for themselves. In North America an example is the monthly magazine *Consumer Reports*, published by Consumers Union. A nonprofit organization founded in 1936, Consumers Union reviews goods and services and publishes investigative reports intended to help consumers make intelligent purchase decisions. Consumers Union reinforces its independence (unlike some other "consumer guides") by not accepting advertising and not allowing their published opinions to be used in advertising. Other publications of this type include *Consumer Digest* and *Consumers' Research Magazine*, neither of which undertakes the extensive testing programs of *Consumer Reports*.

The publications of *Consumer Reports* offer a prime example of what the consumer needs to know to make an informed purchase. The magazine conducts comparative tests of many brands and styles of a product, presenting the results in simplified tables with accompanying text. No matter whether the cost of the purchase is small (e.g., peanut butter) or large (e.g., a new car), the goal is to reduce the often massive amount of salient information into a few key factors, rated or described in the simplest way possible.

Let's consider a hypothetical review of passenger cars (Table 2.1).

Table 2.1: Comparison of passenger cars test results for sedans.

	Mercedes C-300 \$41,000	Volvo S80 \$37,000	Lexus ES350 \$36,000
Price	\$41,000	\$37,000	\$36,000
Likely reliability	•••	•••	••••
Braking	••••	•••	••••
Ride	••••	•••	•••
Comfort	•••	•••	••••
Controls/displays	••	••	•••
Acceleration, 0–60 mph	7.0 seconds	8.0 seconds	7.4 seconds
Required fuel type	Premium	Regular	Regular
Fuel economy	21 mpg	20 mpg	23 mpg

Imagine a consumer (we'll call her Julie) is shopping for a new car. Like many consumers, she already has some background information regarding cars: their makes, models, styles, cost, popularity, and perhaps a sense of their mechanical reliability. She has seen the models that interest her driving about town, and she has shared opinions about them with friends and relatives.

Having two children and a husband to transport, Julie wants a car of intermediate size and good quality. Three sedans made by Mercedes, Volvo, and Lexus interest her, because she has had at least one friend who owned each one of these models and was pleased with them.

Of course, to completely evaluate a car, one needs to drive it. However, from past experience Julie is wary of dealerships and salespeople, and has decided to go to them only after doing some background research. In fact, she thinks she might ultimately buy the car through a broker, making the deal over the telephone or the Internet after visiting the dealers to test-drive the models.

Julie starts her quest for facts with a review from a consumer magazine. Like many such publications, it contains color pictures of the various models, charts with dozens of facts on each model (e.g., dimensions, fuel consumption, features, and prices), comparisons of cars by type (e.g., the best luxury vehicles), and subject ratings and recommendations. What draws Julie to this particular publication is its reputation for objectivity and frequency-of-repair charts, based on hundreds of thousands of reports from owners of the vehicles reviewed and unique to this magazine.

Julie learns several facts from the magazine that help her differentiate the three cars. The first thing she notices is that the price range for the Mercedes is several thousand dollars more than the other two, apparently because of the manufacturer's prestige and reputation in North America. Acceleration is considerably slower with the Volvo, as it has a smaller engine than the other two. Fuel economy is similar among the three models, yet the Volvo and Lexus can use lower-octane fuel — making the Mercedes the most costly to operate. The Mercedes and Lexus are rated better for braking ability than the Volvo, while the Lexus and Volvo score better on dashboard design. Turning to the frequency-of-repair charts, Julie sees different patterns among the 14 "trouble spots" (electrical, brakes, transmission, etc.) but concludes that the Lexus may be more reliable than the other two. Julie concludes that, in matters other than the above, the three cars are similar.

Julie is leaning toward the purchase of the Volvo, reasoning that its use of ordinary fuel means a lower cost of operation; compounding those savings over many years of ownership makes the Volvo the least expensive of these higher end sedans. Yet she retains some doubt as to whether the Volvo is

truly equivalent in features to the other two. She knows she can get basic price information — dealer costs, sticker price, and the costs of major options — from a variety of sources.

In her office one evening after work, Julie locates the Website for *Kelley Blue Book*. Here she is able to get price quotes for the three cars with similar options; the Mercedes is considerably more expensive than the other two cars in price, yet the engine (3.5 L) is the same as that of the Lexus (3.5 L). Does the engine size really matter so much, Julie asks herself? Probably not, she concludes.

Armed with this information, Julie heads to the Mercedes and Volvo dealers for test-drives, deciding against any further consideration of the Lexus. She likes the Volvo but finds the salespeople at that dealership to be too persistently aggressive. At the Mercedes dealer, in contrast, the salesman subjects her to less talk, and puts her in a positive mood for her spin in the car; she immediately falls in love with the Mercedes she drives, but realizes that it has many more options than the basic version that she has been considering.

As the week goes by, Julie discusses the purchase with several friends and family members. One friend cautions her strongly against the Mercedes on the basis of maintenance problems she encountered with her own Mercedes, an anecdote that Julie finds persuasive. On the other hand, her husband, who up until now has remained silent on the topic, argues strongly for the Mercedes on the basis of its larger engine capacity. "What if we want to tow a trailer some day?" he asks. That is a scenario that Julie had not even thought of, and it causes her to go back to the Internet to find the costs of towing packages for each car. While still convinced that the Volvo would be the more sensible choice, Julie ends up buying a Mercedes through a nationwide car broker. And besides — it came in more attractive colors!

Three common anomalies of information seeking are worth pointing out in this scenario. One, the mysterious influence we call "taste" has a powerful role in the decision process. Two, personal contacts have strong influence, whether they compel agreement (e.g., the advice of a friend or loved one), or disagreement (an overbearing salesperson). Drawing upon basic human emotions, these two influences can overwhelm any collection of facts, no matter how large or persuasive, as noted in the Chapter 1 comments regarding the fallacy of rational decision making. Three, it is useful to keep in mind that affluence and education can make a great difference in both the sources of information available to people and their inclination to pay attention to that information. In this scenario, an affluent (and probably well-educated) shopper both knows about and has easy access to channels of information that a poorer or less-educated person may not be aware of or inclined to use.

2.1.2. Finding Information in a Library

In the widespread literature that could be included under the rubric of information seeking, there is a genre of empirical work that is larger than any other: studies of people finding information in libraries. Most of the publications in this genre focus on “information as a thing” (Buckland, 1991a), that is, use of books, journals, and other “packages” of information.

So let us imagine another information seeker, this one called Leslie. Leslie is writing a paper for her history class on the 1898 war between Spain and the United States. She has gone to the library to gather background data on the role of the United States president, William McKinley, in the decision to declare war on Cuba. Among her questions are these: Had McKinley favored war from the beginnings of his presidential campaign in 1896? Was McKinley reasonably well informed of the facts regarding unrest in Cuba and Spanish military weakness there and on the high seas? Who were McKinley’s closest advisors and what was their advice to him regarding intervention? So Leslie has gone to a university library to find answers to these questions. The particular library she has chosen contains roughly 3 million books and subscribes to more than 4000 printed journals and has electronic access to many more.

Before we begin following Leslie’s search, let us consider some tendencies of libraries and their users. First, it is important to recognize that all but the smallest libraries can be complex and intimidating. Although libraries make every attempt to place materials on similar topics in close approximation, this goal is elusive. For one thing it is hard to decide what any one document is “about”; for another, library materials are divided into a multitude of forms — books, journals, computer files (typically through a link from the online catalog), video recordings (on DVDs), and loose materials (in file cabinets or archival boxes) to name some major categories.

As if it were not difficult enough to classify the content of these different forms of media, sometimes information on the same topic and medium might still be found in different places; background material on international conflicts could be found on the shelves of the main collection, on nearby shelves reserved for oversized books, in the collection that serves the reference desk, or perhaps in a special collection or archive. For example, biographies of American presidents will be found on the main shelves of a library, but the personal documents to which they refer are likely to be held in the archives of a distant library.

Once one considers the various interactions of (intellectual) content with (physical) form, it can be seen why libraries become difficult places to search systematically, even with experience. Many visitors to a library end their search prematurely when faced with a large building full of millions of items and the imperfect tool of the electronic catalog.

Leslie begins by consulting the electronic catalog, a tool she has used before. Being a regular visitor to this library, she is aware that if she chooses to consider journal articles she will need to consult at least one other electronic index to obtain the titles and full-text of articles that contain relevant information, and that this would entail much more work than a simple search of the library’s holdings. She decides to restrict herself to books about the war and McKinley.

This library offers two different interfaces for searching their holdings. The simpler one is a graphical interface that Leslie has not used before but hopes is less cumbersome than the “traditional” catalog. So first Leslie tries typing the words “war cuba mckinley” into the search box of the graphical interface. Immediately she is presented with a screen divided into two kinds of displays. On the left side of the display is a cluster of words, with lines connecting them. In the middle of this network-like diagram are her own search terms (war cuba mckinley) and branching out from them in all directions are other terms: Cuba, Spain, President, United States, History, Novel, Nations, Battle, McKinley, McKinney ... and a variety of other words that appear to be synonyms or alternate spellings (and meanings!) of the words she typed. On the right of the screen is a list of books, a few of which are indeed relevant to her topic.

Not sure what to do with the cluster of words — many of which seem irrelevant — Leslie decides to retreat to the traditional interface for searching the catalog. Selecting the other interface she types “McKinley” into the “Search:” box and chooses “within Title” from the pull-down list to the right of the box. She is rewarded with a listing of 12 books. Looking at the book titles, however, she is dismayed to see that few of them are about the former president; rather, they are books that begin with the name “McKinley,” whether that is the name of a county or a mountain or a person. Next she receives a list of four books by typing “William McKinley.” Even though that seems like very little material to browse through, she heads for the shelves. Checking the books in the “E711” section of the Library of Congress system, Leslie immediately sees that there are many more books on her topic than were retrieved in her search — dozens more, in fact. Based on the titles alone, Leslie’s first insight is that most of these books are about the life and/or presidency of McKinley, and are likely to contain merely brief descriptions of the war with Spain in 1898.

After browsing a while and picking up two books about the McKinley administration and one biography of the president, she heads to the reference department of the library. There she asks why her search of the electronic catalog was so incomplete. The reference librarian shows her that there are such things as subject headings in the catalog records. Returning to a terminal for the online catalog, Leslie this time selects the “within Subject” choice from the pull-down menu, and types in “McKinley, William.”

This time, 25 books on McKinley are retrieved, even those that do not have his name in the title (Figure 2.1).

Leslie starts to examine the 25 listings, one at a time. The most promising title appears to be *The Spanish-American War and President McKinley*, a book that she just missed because it was at E715. She makes note of several other McKinley books she missed in her browsing session at the shelf, seeing also that two of the more relevant ones are already checked out.

Now that she knows about subject headings, Leslie sees one that could be useful: "Spanish-American War." Searching that phrase as a subject heading gives her 127 titles. Intimidated by the size of this new list, Leslie slowly pages down through the titles until she just cannot read any more. She has jotted down the call numbers of the most promising books, most of which are in the E711–715 range of the Library of Congress system, and a bibliography at Z8561 on the top floor of the library.

Still carrying her initial three selections, Leslie goes back to the E shelves and an hour later has examined 15 other books, selecting just two highly *relevant* ones to check out from the library and leaving behind all three of the first books she chose. She knows that the bibliography in the Z shelves would help her determine whether she has missed anything — this library does not own everything — but the bibliography is two floors above her and she is tired. "This is enough to finish my paper," Leslie says to herself as she heads to the circulation desk.

In this scenario several lessons about information seeking can be observed. Although perhaps two-thirds of adults in the United States and Canada make some use of libraries in a given year, relatively few (mostly

SearchRequest: S = MCKINLEY, WILLIAM Subject Index		University Online Catalog Search Results: 25 Entries Found
MCKINLEY, WILLIAM 1843-1901		
1	COMPLETE LIFE OF WILLIAM MCKINLEY AND STORY OF HIS . . . [1901]	
2	DICKEN TROUTMAN BALKE FAMILY PAPERS [1816] archive-mss	
3	EDWARD HENRY HOBSON PAPERS [1857] archive-mss	
4	FROM MCKINLEY TO HARDING PERSON RECOLLECTIONS OF . . . [1923]	
5	ILLUSTRIOS LIFE OF WILLIAM MCKINLEY OUR MARTYRED . . . [1901]	
6	IN THE DAYS OF MCKINLEY [1959]	
7	LIFE OF WILLIAM MCKINLEY [1916]	
8	LIFE OF WILLIAM MCKINLEY SOLDIER LAWYER STATESMAN [1896]	
9	MAN WHO SHOT MCKINLEY [1970]	
10	MCKINLEY MEMORIAL IN PHILADELPHIA HISTORY OF THE . . . [1909]	
CONTINUED ON NEXT PAGE		
Click on entry to display full record		

Figure 2.1: Portion of an online catalog listing.

students and faculty in universities) search library collections in any degree of depth. Leslie is an untypical user in that she knows how to use a librarian and a catalog; the reluctance of even regular users of libraries to consult these resources is well documented (consider, for example, the commentary by Borgman, 1996).

Leslie is, however, typical in her nonlinear search pattern; her search is not a neat one that moves swiftly from catalog to shelf to circulation desk; rather, there is a back-and-forth movement between the catalog and the shelf, with considerable time taken to examine works and reconsider her query. Typical of library users, Leslie takes some shortcuts (choosing to consider only books, not journal articles), reverses some of her early decisions (leaving behind the initial choices of books), and ultimately ends the search process prematurely by not fetching the bibliography and checking that (presumably comprehensive) guide against her search results.

2.1.3. Healing a Patient

Although our usual focus in information behavior research is the individual, it is important to recognize that much information seeking takes place in groups, and that sharing of information with others is very common. In this scenario we consider the many exchanges that might take place among members of a health care team in a hospital. Because the information seeking plays out over several days and involves many individuals, it is possible only to sketch out some of the dozens of actions and interactions that are likely to occur. Also for simplicity's sake we will consider only a single patient, while recognizing that such a team would certainly be caring for multiple patients at any one time, each at different stages of crisis or healing. To further condense the description, we focus on only four key team members: physician (Dan), day nurse (Marta), night nurse (Gary), and pharmacist (Ming).

The story begins without the team, in the emergency entrance to a rural hospital, where a 22-year-old male has been brought in by an ambulance shortly after midnight. The man was in a car accident and is semi-conscious but shows no other injuries other than contusions. Unfortunately, the man carries no identification and the car was registered in a female name out of state. Thus, he must be admitted without access to a medical history.

The patient is first examined by a nurse, and then by an attending physician. Both of these administer the Glasgow Coma Scale (GCS), a 15-point neurological measurement of consciousness. A score of eight is recorded each time, which could indicate severe brain injury. Then he is taken for CT scan of the head and upper body, and a blood test for presence of alcohol and drugs. Afterwards, the patient is taken to the intensive care

unit (ICU). Here the night nurse, Gary, supervises the movement of the patient to one of the ICU's 15 beds. The patient is soon fitted with an IV to deliver fluids and medications, and devices to monitor blood pressure, heart rate and oxygen levels of the blood. His temperature is taken, and he is given periodic ice packs to reduce swelling. Gary also gives the patient the coma test again, consulting a "flashcard" he keeps for this purpose, as he has trouble recalling the individual indicators. He records a score of 10 on the chart using the conventional breakdown for the Eye, Verbal, and Motor components of the scale: "GCS 10 = E2 V4 M4 at 01:17." The patient is given no medications at this point, as they might interfere with indicators for the scale.

The patient's hospital room contains a number of potential sources of information, the ultimate origin of which, in some cases, is the patient himself. Among these sources are: the monitors attached to the patient (including blood pressure, heart rate, and oxygen level of blood); the CT scan results, viewable at a nearby workstation; electronic and paper versions of the patient records, including orders for medications administered; and a bedside white board for nurses to record any facts that need to be called to the immediate attention of anyone who enters. On this board Gary has noted the latest coma scale score and the words "Patient Name Unknown." In a nearby alcove outside the room is a nursing station, where a number of medical reference works (e.g., the *AACN Procedure Manual for Critical Care*, *Quick Reference to Critical Care*, and *Micromedex*) are available online, along with medical databases like *MedLine* and *Drug Interaction Checker*.

The fact that the patient's name is not yet known is a serious problem, as nothing is understood about medical problems that might have pre-dated the accident. This situation is resolved early in the morning as the responding police officers report to the hospital switchboard that they were finally able to contact the registered owner of the wrecked vehicle and determine that the patient is her son; now they have a name, which is passed upstairs to Gary where he duly records it on the patient's white board. As the morning shift Gary verbally passes on his recorded observations to the day nurse, Marta: the patient is partly conscious now, and has vomited twice during the night; temperature has been only slightly elevated, although blood pressure has been high for someone of this age and weight; his GCS is currently 11.

Marta, with the help of other hospital staff, begins a series of inquiries about the patient's medical records and history, and the status of his health insurance. Emergency medical procedures allow for medical records to be transferred electronically to the treating hospital in such cases. The medical records — which are relatively few in number — become part of the data accessible in the ICU. Late in the morning, the attending physician, Dan,

shows up to examine the patient and be filled in on the situation. Marta calls his attention to several important facts: the patient is not yet fully conscious for more than short periods of time; the blood test showed presence of both alcohol and opiates, which makes the accident a legal matter; the patient has a current prescription for Adderall (for attention deficit disorder); and he has no medical insurance.

Among other treatment decisions Dan and Marta decide to consult one of the unit's pharmacists, Ming, about whether the patient should be given an anti-emetic (anti-vomiting medication), and the implications of administering any other medications, given the patient's medical history. A two-way radio call reveals that Ming is unavailable until later, which is agreeable. Meanwhile, Marta consults the *Drug Interaction Checker* for possible interactions with Adderall, although Dan has told her that the patient could well do without this medication, given the circumstances. She learns that Adderall can be addictive and can also raise blood pressure — factors to keep in mind in this case. She notes this in the patient's electronic chart.

Late in the afternoon the patient is more conscious for a longer time, moaning and complaining of a severe headache. Ming and Dan appear for their consult, and are able to ask the patient questions about his current and past condition, and his medication habits. He claims to have been in good health before the accident (which he does not remember), but denies taking any drugs (prescribed or otherwise), including Adderall. He asks for some Oxycodone, an opioid, for his headache. These responses cause some skepticism among the three medical professionals about the patient's truthfulness.

Outside of the room, the three caregivers discuss what to do. Given that the patient shows possible signs of drug abuse, they are unwilling to prescribe strong pain medications. They decide that small doses of acetaminophen are warranted, as the GCS is now at the high end (12) of the moderate injury portion of the scale. Another CT scan would be advisable if the patient worsens, but not if he improves. They decide to continue monitoring him and checking in again during the evening.

The next day the patient appears better in several respects. Apart from evident bruising, other vital signs are good: lower blood pressure and normal body temperature. His GCS is 13. The patient continues to complain about headaches, but not so loudly as the day before; from training and experience Marta and Dan know that headaches are common for concussion victims for several days after injury. The ICU has an interest in moving patients along as quickly as possible, so that they always have beds free for emergency patients. Two days into the patient's stay the ICU has come close to capacity, and Dan and Marta agree that the patient is stable enough to be discharged to his family. This is done, but not before a

police officer makes arrangements to take a statement before the patient leaves the state. Through a warrant, the police also obtain the results of the blood test administered at the time of admission, however they are not given any other medical records.

Among the features in this scenario are continuing measures of, and searches for, information, careful recoding of any relevant data, and frequent sharing of information among team members. Given the life-and-death nature of the work, missing or inaccurate information is a constant worry, as there are many opportunities for potential adverse events. There is also a concern for control of patient information, due to privacy and legal issues. The collaboration of multiple skilled workers, equipped with both deep knowledge and extensive physical resources, makes a positive outcome much more likely.

2.1.4. Betting on Race Horses

As in the previous scenario, for horse races the seeker makes a choice. Despite the prevalence of horse-racing language in politics (e.g., "the dark horse candidate" and "backing the right horse" in an election), there are some differences between choosing a candidate and picking a horse. Rather than choosing one candidate from among two or three, in a typical horse race several choices might be made from among roughly 5 to 12 horses, and the type of bet that might be made on the horse(s) multiplies the number of possibilities many times over.

The complexity of horse racing leads to a common, sheepish admission at the tracks: "I choose horses based on their name." That is, if a name like "Gambling Everything" makes one laugh and seems to capture the moment, why not bet on that horse? Infrequent visitors to the race course, attending more for fun than the hope of profit, freely admit the unscientific basis for their choices; if the name is especially clever, reminds them of a friend or circumstance in their lives, then that is a sign to bet on the horse. It is not the only simplified system for betting. Some bet on an animal's color, with the gray horse in a race likely to create odds more favorable than the horse's reputation deserves. Occasionally the gender of a horse will determine many bets, especially in a race in which a talented filly vies with stallions. Races that draw horses from other countries may elicit bets based on nationality. Some wager not on the horse but on the record, gender, or nationality of the jockey (and so on).

Why is properly picking a horse so complicated that people will resort to simple strategies like the above? First, it must be obvious that no matter how much data are considered, horse racing itself is not a science and offers many unexpected surprises. The most dramatic example of this is when a

highly ranked horse stumbles in a race, or bumps an opponent in such a way as to be disqualified. Even in the most mundane race, the horse most favored by the bettors wins only one-third of the time (Ainslie, 1986, p. 49).

Serious bettors who gather and use as much information as possible about the horse, the jockey, the trainer, and the track (among other things) are called "handicappers." Handicappers firmly believe that, given enough information, they can swing the odds to their favor in an enterprise where the average bettor loses about 20 cents on the dollar. Not surprisingly, a large publishing industry has grown to serve the information needs of handicappers.

The publications on horse racing are many. Most people who have been to a track are familiar with the racing program sold at the track itself, important for its listings of information about the horses and races to be held on that day at that track. However, a track's own program is just the tip of the iceberg of information about horses. Whatever data a North American track distributes about its own races, that information is overwhelmed by what can be found in the *Daily Racing Form* (published by Triangle Publications, Inc.), which is published every racing day and distributed internationally. The *Daily Racing Form* is in turn dwarfed by the *American Racing Manual*, an encyclopedic by-product of the *Form* that covers an entire year. In addition, there are several (mostly weekly) magazines devoted to racing. Among them are *American Turf Monthly* (advice for handicappers), *The Blood-Horse* (the inside story on horse breeding), and *Turf and Sport Digest* (news for the horsing industry as well as the frequent bettor).

Just what information could merit so many publications? Considering only information about the horse and jockey, there are many items of data to consider. The key category of information is the horse's recent record of racing (the *Form* usually lists the last nine races) and sometimes one to three workouts (trials that were not races). Regarding each race, there are approximately 25 items of information (all on a single line), including the date, length, type, timing, and top three placers of the race, along with the jockey, odds, and weights carried by the horse, and often ending with a subjective evaluation of the horse's performance (e.g., "tired" or "failed to menace"). Elsewhere on the page are summaries of the horse's lifetime earnings by year, and the names of the sire, dam, owner, breeder, and trainer. Altogether, the *Form* typically offers about 250 discrete data items regarding each horse — and there are typically 5–12 horses in each race! Even without considering the other relevant information contained in the form about the race itself (e.g., what types of horses can race and how much they can win) and the track (top times for each of up to 16 different lengths that races can run), we can see that each race offers the potential for consideration of several thousand data items.

But humans do not want to digest several thousand items of information and often simply do not have the time to do so. Therefore each handicapper relies on a selection system that eliminates most of this information; typically they focus on just a few dozen data items about each race and sometimes considerably fewer. Less successful but a lot less work are the systems described above, such as "Always bet on the prettiest horse!"

Let us consider the case of an occasional bettor who would like to become a regular handicapper, a common enough circumstance. Joe, our bettor, goes to the track and buys a copy of the *Racing Form*, along with a program of the day's races. It is his first time with the *Form*, and he is at first overwhelmed by the scope and depth of information he finds there. With the first race just an hour away, he settles into a comfortable spot and starts making notes on what he reads.

Joe has a few rules of thumb in mind to guide his reading. Looking at the listing for the first race, he notes the "morning line" (projected) odds for each of the nine horses that will start that day. Joe can see right away that three of the horses are expected to run at much better odds than the other six; horses one and two are likely to pay \$3 for every \$1 bet (i.e., odds of 3:1), whereas horse three is expected to run at 4:1. All of the other horses have odds of 8:1 or worse, with one poor horse paying 50:1 for a win. For a moment Joe ponders what he could do with \$50 for every dollar he bets, then shakes himself back to reality with a reminder that such horses very rarely win; he also reminds himself that he has to check the TV monitors at race time to see how much the odds have changed as bets are wagered; in parimutuel betting the odds are based on the actual pattern of bets made, rather than the estimates of the *Racing Form*'s handicappers.

Joe decides to concentrate on what are clearly the best three horses in the race and on just three types of information: the horse's performance in recent races, the jockey's record, and something called the "speed rating." He notes that horse one, Entropy, won one of his last nine races, and has placed second (i.e., "placed") in three others. Horse two, Uncertainty, has also won one of his last nine races, and has placed second in two others. The number three horse, Signal, has not won a single race, yet has managed to come in second three times, and has come in third (i.e., "showed") two times as well; furthermore, all of Signal's second-places were in the last three races, indicating an improving performance.

Not seeing a clear pattern in these data, Joe decides to look more closely at each race for further clues. He considers an important question: has the horse won *at this distance* before? Joe knows that distance is an important factor and that an earlier win at the same distance is a strong indicator of potential success today. Races can be a variety of different distances, and many horses excel at the shorter or the longer distances, but not both; in addition, some races are run on the turf (i.e., the grass strip inside of the dirt

track), and some horses specialize in such a surface. From the top of the *Racing Form* Joe sees that this is a 1-mile race on the dirt track, and recalls that it has not rained this week; a track muddy from recent rains would require consideration of yet other factors, and Joe feels like he has too much to think about already.

Our bettor happily reads that Uncertainty's win and places have all taken place on the turf, and that may well allow him to eliminate that horse from consideration. His comfort is lessened when he realizes that Entropy's victory was in a race of 6 furlongs (three-fourths of a mile) whereas Signal's near-misses have all been in races of 1 mile. So neither horse has won at this race before, but Signal looks better prepared for this long race.

Next Joe considers the speed rating, which is an objective metric that compares the horse's performance with a rating of 100, meaning that the horse's running time (in a particular race) equaled the track record for that particular distance. Entropy, Joe sees, has a speed rating of 80 for his winning race, indicating a performance 4 seconds slower than the track's record; his places earned similar ratings. Although not usually a dirt racer, Uncertainty is undoubtedly a faster horse at an average rating of 84 for his three best races. Signal begins to look worse with speed ratings between 77 and 81.

So now the picture is looking even more muddy. Joe turns to a separate table in the *Racing Form* that lists the performance records of each jockey, over many races and with different horses. Here, at last, is clear-cut information: Entropy's jockey has won 17% of his races, while the other two jockeys hover near 8%. Joe jots this down on a little table he has been sketching on a notepad (Table 2.2). This cinches it: Entropy is to be the horse Joe bets to win.

It is now 32 minutes to race time and Joe remembers that he needs to check out the latest information on the track TV monitors. Up to this moment he has been dealing with information that is at least hours, and in some cases months or years, old. It is with a shock that Joe reads on the monitor that Entropy has been scratched from the race, and that Signal has had a change of jockey. Another look at the table in the racing form tells him that the new jockey has a better record — 11% wins — than Signal's old jockey. Joe looks

Table 2.2: Joe's notes on the three top horses.

Horse	Odds	Wins	Places	Track	Speed	Jockey won (%)
Entropy	3:1	1	3	Dirt	80	17
Uncertainty	3:1	1	2	Dirt	84	8
Signal	4:1	0	3	Turf	79	8

at the latest odds and realizes that the turf horse, Uncertainty, is now the crowd's favorite at even odds (1:1) while Signal has crept up to 5:2. "Do the other bettors know something I don't?" Joe wonders.

Joe considers spending more time with the *Racing Form*. He knows that it contains information on the horse's recent condition, class, parentage, and so forth, but he is feeling undecided and a bit pressed for time; he needs to place a bet within the next 20 minutes. It is then that Joe recollects a potentially valuable source of information that appears in no published document: the horse's condition at this moment. Walking quickly over to the rear of the racing stands, in a few minutes Joe finds himself at the paddock, where the horses for the first race are being paraded for a large number of bettors crowded around the rails. After a few minutes of craning his neck and standing on his toes, Joe manages clear glimpses of both Uncertainty and Signal; the former animal appears nervous while Signal appears calm and strong. With his mind made up, Joe rushes to the nearest betting window where, after 5 minutes of waiting he bets \$10 on Signal to win at the latest odds of 2:1.

Ten minutes later the finish line is crossed first by Channel, a horse that Joe had not even considered. Trying to salvage a lesson from this first race, Joe notices that in the last 15 minutes of betting, Channel's odds moved from 6:1 up to 3:1 — a sign that many bettors had begun to favor the horse. "I would have noticed that trend if I had been paying more attention to the monitor," Joe says to himself, and then with a start realizes that he has much less time to do the next analysis: Race Two will begin in only 25 minutes.

We can notice a few principles in this case study: the searcher has much too much data to systematically analyze in the brief time before each race. Hence Joe attempts to simplify the "search space" for a solution by ignoring many categories of information he deems less relevant. Unlike many other search processes that result in some kind of decision, in this case the most important data — the betting odds — keep changing constantly, right up to 2 minutes before the race. Like the opportunity for first-hand observation of additional information (the horse's condition), the need to "monitor the monitor" can be a distraction from further analysis.

2.1.5. Finding the Law

George is spending his Saturday in the law library. A lawyer in a small Kentucky law firm, George has been practicing just 5 months (problem 1) and is overwhelmed by his work (problem 2), which is chiefly tax law (problem 3). Nevertheless, he promised his sister, Edna, to write a legal memorandum on her alimony case even though he feels that he does not

have the experience, time, or background to do a perfect job of it. The hard fact is that Edna cannot afford a lawyer, is starting to have trouble paying her rent, and besides, she is his sister. So even though the course George took in family law is only a dull memory now, today he joins dozens of other lawyers and students in a quest to find the law.

George knows the basic facts of the case all too well. Last year when Edna divorced her husband Fred, the court ordered him to make monthly alimony payments. Three months ago Fred took a new job in another town, moved to an apartment, gave up his phone, and stopped making payments. Fred's only direct communication with Edna was a phone message warning that he would have to skip her next payment to afford the deposit on his new place, and that he would "make it up to her later." The first few weeks Edna was annoyed yet somewhat sympathetic, because she was the one who filed for divorce. But after a month went by with neither the next nor the previous alimony payment, Edna felt foolish that she had not acted more quickly; to make matters worse, Fred has proved to be impossible to contact.

Edna knows that she has legal recourse through the court, but also knows that it could be a struggle to document the facts and recover lost payments. So she turned to George first. George is determined to settle this dispute as quickly as possible, even if it means he has to pay another attorney to follow up his work later; unfortunately he has precious little time over the next few weeks to devote to it. Though George believes that Edna's case is a straightforward one, he wants to check two issues in particular: first, whether a recent raise in Edna's modest salary could reduce alimony or impede collection of what is owed to her; and second, whether Edna's hesitancy to act when the first payment never arrived might allow her husband to argue that she had agreed to the stoppage of payments.

George starts with the index to the *Corpus Juris Secundum* (referred to as the CJS) and locates the section on Divorce in this encyclopedia-like reference tool. He finds, two subheadings down, a section on "arrears"; he spends some time reading this section of the main body of CJS, and also in a newer *Supplement*. As well as reminding himself of the key points of the law in this area, he locates a citation to a recent case decided in Kentucky. George jots down the citation to the Kentucky case, a parallel citation, and a frequently cited decision from New York State. Next George turns to the *Quick Index* to the state volumes of the *American Law Reports* and finds an entry (Figure 2.2) to a discussion of "arrearages" under the heading "Alimony — Delinquent or overdue payments." After reading the discussion in *American Law Reports*, and noting references to yet other cases, George decides to read the cases themselves.

George looks up the Kentucky case in *Kentucky Decisions*, which reports cases for his state, and another in the *New York Reports, Second Series*. For his last task of the day, George uses a public terminal and his office's LEXIS

ALIMONY—Cont'd

Delinquent or overdue payments
debt, right of spouse to set off debt owed
by other spouse against accrued spousal
or child support payments, 11 ALR5th 259

laches or acquiescence as defense, so as
to bar recovery of arrearages of permanent
alimony or child support, 5 ALR4th 1015
visitation, withholding visitation rights for
failure to make alimony or support payments,
65 ALR4th 1155

Figure 2.2: Excerpt from the *American Law Reports*.

account number to search the online version of *Shepard's Citations* to check on the continuing validity of the cases he examined; he finds the most relevant cases are still valid.

George has spent 3 hours in the library and is tired. But at least he feels grounded in the relevant law and has checked the case updates. Tomorrow George will draft a memorandum on Edna's situation, and the following week he will discuss it with a colleague who practices family law. He knows that it may be hard to get money out of Fred if he continues to avoid responsibility, but George is certain that, should Fred attempt legal resistance, Edna's case would be most compelling to a court.

Given Brenda Dervin's emphasis on "everyday" information seeking (see Chapter 1), most of her 10 points are not particularly relevant to a highly structured search of formal information sources by an experienced attorney. However, it is worth noting that, although George certainly found "the law" governing his sister's situation, her problems are not over yet! There is much work to be done before the records of law and previous decisions may result in some kind of action in Edna's favor. As Dervin implies, few problems are immediately solved by the discovery of relevant information. And as some psychologists and criminologists have documented, court decisions are not always as rational as we might hope, either.

2.1.6. "I Want to Know More about Cancer"

Let's consider an entirely different kind of desire for information, one in which there is no decision or choice to be made, and in which it is difficult to determine exactly what the "need" is. "Curiosity" is the label that we

might apply to the situation described below; according to Webster's *New World Dictionary*, curiosity is a word that is used to indicate a general "desire to know," sometimes "about things that do not necessarily concern one."

It is certainly hard to say what makes us curious about a subject. In a famous article about why and how people ask questions, Robert Taylor (1968) wrote about "visceral needs" of mysterious origins that make themselves known only through a vague uneasiness about not knowing something. The visceral need remains "unarticulated" until we verbalize it to ourselves or someone else. In the process of trying to state what it is we want to know, the information need usually comes out in an imperfect and unsatisfactory statement ("compromised," in Taylor's words). Brenda Dervin's (1983a) investigations into "sense making" employ a similar concept: the information *gap*. That is, until we recognize the existence of a gap in our knowledge — often signaled by a mild anxiety and/or a need to act — we are not motivated to search for information. However, whether we ask questions, read books, or take another kind of action to find something out, it is important to recognize that information often comes to us, fortuitously, in the course of our normal lives. The *serendipity* factor — the seemingly accidental discovery of relevant information — operates more often than we might expect.

Our searcher this time is named Maria. Maria, who is in her early 30s, was never particularly concerned with matters of personal health until a favorite cousin discovered that a firm, red lump on her arm was cancerous. Through several weeks of a successful treatment plan, Maria kept in frequent telephone contact with her worried cousin, who lived in a distant state.

Maria had heard the dreaded word "cancer" her entire life. When she was a teenager her grandfather had died of lung cancer, but other than him, nobody close to her had ever been diagnosed with cancer. She knew that many old people died of cancerous growths. Cancer, like death itself, was something that Maria would rather not think about.

Not long after her cousin's discovery, Maria came down with a sore throat and visited a nearby medical clinic. While in the waiting room she noticed a brochure, "What You Need to Know about Skin Cancer," published by the National Cancer Institute of the National Institutes of Health. She took the brochure home to read and was surprised to learn that almost half of all mature adults are likely to have had skin cancer. She had thought cancers were pretty rare, except among the elderly. At least Maria felt confident now that she knew how to spot skin cancer herself and what to do to avoid it. And, for a time, that was all she wanted to know about cancers.

Some weeks later, coming across the brochure on skin cancer in her living room, Maria found herself curious. What *was* cancer, exactly, and what

caused it? The brochure did not say much about the underlying nature and causes of carcinoma, but it listed a toll-free number for further information (1-800-4-CANCER), which she called to request other brochures in the National Cancer Institute series. Later she had coffee with a friend who was a nurse and asked some very basic questions about cancer. Maria did not understand quite everything that she was told by her friend, but understood enough to know that she wanted to know more. *Why* she needed to know more, she was not exactly sure, but maybe it could be useful in defending herself against future illness.

One day Maria bought some skin cream made by the Avon company. With the skin cream came a list of other Avon products, which mentioned that company's "Breast Cancer Awareness Crusade" and listed a Website. Out of curiosity Maria used her home computer to reach Avon's Web page, where she found, along with answers to frequently asked questions about breast cancer, some fascinating narratives by women who had survived the illness. She had never thought about using the Internet for this sort of information; somehow it made it easier for her to read about the scary topic of cancer.

Maria noticed that the Avon site did not refer to any other related sites. She decided to do a search on the word "cancer" and was bewildered by the number of sites (over 200 million) that contained a reference to the word. Obviously, she had to be more specific. Her nurse friend had mentioned the Mayo Clinic as a good source of information; searching for the words "Mayo Clinic" lead to a site called the "Mayo Breast Cancer Center" that included many pages of clearly written material on cancer in its various forms. She printed some pages, and skimmed or read others.

Finding more information on cancer became a kind of challenge. Through the Mayo site she found a link to a page sponsored by the National Institutes of Health. Maria felt like she might be able to trust this information more, because it was provided by a national government. However, she found the site somewhat overwhelming — so much information was about government projects or about research projects — and she did not find the sort of common-language explanations she was looking for. She noticed a link to the National Cancer Institute's Website and there found the same material that she had read in the brochures they had sent her. At this point she decided to call it a day and stop searching.

Maria's interest in cancer did not stop here. At various times she spent hours browsing the personal health sections of a local bookstore and the public library in her town. She talked to her friends about "all this stuff I've learned on the Internet" and became known as somebody who liked to talk about health matters. In turn, Maria learned a great deal listening to the experiences of her friends — their worries about staying healthy, and their stories about relatives who had cancer and heart disease. It seemed like every

conversation Maria had with others about health sent her back to the Internet to answer a question, to learn about a disease, or simply to surf.

Maria's case represents two of the softer dimensions of information seeking: a tendency to avoid certain information until one is ready to deal with it, balanced against a curiosity motivated by deeply held feelings. In situations like Maria's, the urge to find facts and hear personal stories may satisfy some emotional need to be reassured, to be comforted, to *connect* with others. In this sense there is no final answer that will end information seeking — it is the project of a lifetime.

2.2. Summary

We have seen, through six fictitious case studies, the playing out of information seeking in different contexts. Common to them all has been the need to deal with (potentially, at least) great volumes of information, much of it complex. In all cases the searchers have attempted to lessen their cognitive load by jettisoning some types of information, taking a shortcut to a state of satisfaction or decision.

The results of most searches, although incomplete and perhaps even resulting in failure (witness Joe's lost bet), were, at the time, *good enough* to satisfy the needs of the seeker, a type of behavior that is called *satisficing*. With the exception of the ICU team (a group of professionals with considerable resources) the seekers of information did *not* make every possible attempt to attain the most complete, accurate, and detailed information available (*optimizing*) but rather gathered just enough data, opinions, and impressions to feel satisfied with the process. When a person reaches such a stage, he or she may end the task with a feeling of closure.

These six scenarios were chosen to provide readers with contrasting elements across different information seeking activities. These contrasts are highlighted in Table 2.3, in which the scenarios are ranked in order of *time pressure*. The primary *motivation* in each of the six scenarios varies widely; two searches are prompted by assignments given to the seeker by other persons, and the rest are personally chosen by the seeker out of self-interest. *Sources of information* used by each person are more homogeneous. Most seekers use a mixture of formal (e.g., printed publications or electronic sources) and informal information (e.g., the opinions of friends); the major exceptions are the three assigned tasks, in which documentation of official sources is important.

Time pressure is determined by how quickly a search for information must be concluded: minutes, hours, days, weeks, or months. The *degree of thoroughness* roughly summarizes the ratio of data-actually-gathered-and-considered to all-data-potentially-relevant in a search. The *time pressure* and

Table 2.3: Comparison of six case studies.

Seeker(s) and situation	Main motivation	Sources of information	Time pressure	Degree of thoroughness
Julie: car purchase Leslie: library research	Optimize functionality and value Class assignment; earn credit/grade	Friends, Web pages, salespeople Online catalogs, books, journals, professional advice (on how to search)	Low (months) Moderate (weeks)	Low Moderate
Hospital ICU team members: caring for an accident victim	Work assignment; desire to help others	Observation of patient, paper and electronic records, monitoring devices, medical manuals, hospital employees Special journals, observation, intuition Special databases and publications, professional advice	Very high (hours or days, based on patient improvement) Very high (minutes) High (days)	High
Joe: horse race wager George: legal research	Desire for thrill; to win money Work assignment; help relatives	Special journals, observation, intuition Special databases and publications, professional advice	Very high (minutes) High (days)	High High
Maria: information on cancers	Curiosity; preemptive information search	Web pages, books, brochures, friends, experts	None (lifetime)	Moderate

required degree of thoroughness faced by the seekers of information vary the most, as can be seen by examining the more extreme cases.

Maria's interest in cancer-related information comes out of a personal curiosity that may last a lifetime; in her case there is no identifiable time pressure at present. In an optional purchase such as Julie's, the decision could be postponed indefinitely, making the search for information protracted; as well, the number of variables is large and the importance of subjective judgment (e.g., taste) is so important that a relatively low proportion of potentially relevant data are considered in the final decision. In the case of Joe's racing wager, a decision must be made within minutes, and the amount of information of potential relevance is so large that only selective use of information is possible. Finally, George's legal research requires not only very fast turnaround (days) but absolute accuracy as well; therefore, George must be especially careful to search all of the relevant sources and to do so thoroughly. The other cases lie in between these extremes regarding time and thoroughness.

The cases we have examined represent quite common scenarios: writing a term paper, buying a product, satisfying our curiosity, and so forth. There is one sense in which they are *not* representative, however: most of the academic research conducted on information seeking examines scenarios more like those of George (work in a professional or occupational setting) and Leslie (searching formal sources, particularly in libraries).

In fact, a high percentage of the situations examined in traditional information seeking research have involved either laboratory scientists, engineers in private firms, or university research faculty (in a broad array of disciplines). Because we will be reviewing such cases in future chapters, there is no need to emphasize such complicated research settings at this point in our discussion. The human factors bearing on information seeking are similar with all questions and in all fields of inquiry.

Now that we have grounded the topic in familiar human situations, we will next turn our attention to a basic question: their corresponding definitions and corollary concepts. What is "information?" What is "information seeking?"

Recommended for Further Reading

Aspray, W., & Hayes, B. (Eds.) (1956). *Everyday information: The evolution of information seeking in America*. Cambridge, MA: MIT Press.

This edited book features examples of how individuals have pursued their daily information needs over several decades, including chapters on car buyers, airline travelers and sports fans.

42 Recommended for Further Reading

Connaway, L. S., Dickey, T. J., & Radford, M. L. (2011). "If it is too inconvenient I'm not going after it": Convenience as a critical factor in information-seeking behaviors. *Library & Information Science Research*, 33(3), 179–190.

A review of two recent studies of library users, noting that they frequently cite convenience as a factor in how and where they search for information, however their definition of convenience may change over time, depending on their situation.

MacIntosh-Murray, A. (2001). Scanning and vicarious learning from adverse events in health care. *Information Research*, 7(1). Retrieved from <http://InformationR.net/ir/7-1/paper113.html>

This review of studies of information failures in medical practice discusses how health care organizations can systematically learn to avoid such problems.

PART 2

CONCEPTS RELEVANT TO INFORMATION BEHAVIOR

Chapter 3

The Concept of Information

Information seems to be everywhere. We talk of its being encoded in the genes ... Disseminated by media of communication ... Exchanged in conversation ... Contained in all sorts of things Libraries are overflowing with it, institutions are bogged down by it, and people are overloaded with it ... [yet] no one seems to know exactly what information is.

— Christopher Fox (1983, p. 3)

Information, usually seen as the precondition of debate, is better understood as its by-product. When we get into arguments that focus and engage our attention, we become avid seekers of relevant information. Otherwise we take in information passively — if we take it in at all.

— Christopher Lasch (1995, p. 162)

Chapter Outline

3.1. Searching for a Definition of Information	46
3.1.1. Explicating “Information”	47
3.1.2. The Concept of Information	48
3.1.3. Typologies of Information Concepts	49
3.2. Definitions of Information and their Problems	51
3.2.1. The Influential and Restrictive “Information Theory”	52
3.2.2. Five Problematic Issues in Defining Information	56
3.2.3. Utility as a Requirement	57
3.2.4. Physicality as a Requirement	61
3.2.5. Structure/Process as a Requirement	62
3.2.6. Intentionality as a Requirement	66
3.2.7. Truth as a Requirement	67
3.3. Must there be a Universal Definition of Information?	68
3.4. Distinctions among Information, Knowledge, and Data	72
3.5. Summary	73
Recommended for Further Reading	75

3.1. Searching for a Definition of Information

“Information” is a fairly old English word, making an early appearance in one of Chaucer’s tales sometime between 1372 and 1386 (Schemert, 1993a, p. 177); Capurro and Hjørland (2002) take its origin back to Latin and Greek terms of the pre-Christian era. One would think that hundreds of years of usage would tend to settle a word and result in a consensus on its meaning. This has not been the case with the term “information.” Especially in the last five decades, as the various phenomena that people call information began to be objects of empirical study, meanings of the word have proliferated. Schrader (1983, p. 99) goes so far as to complain about “the multiplicity of vague, contradictory, and sometimes bizarre notions of the nature of the term ‘information’.”

One of the problems of studying any phenomenon — or merely talking about a thing — is reaching an agreement on what to call it. Words are ambiguous, the same string of characters often having multiple meanings. Each meaning may identify a distinct concept in the way that the noun “port” can refer to a fortified wine, the left side of a ship, or a gateway or opening for passage (a harbor, a modem port, a valve port, etc.). The case of the word “information” is much more complex, as it has been used to denote various overlapping concepts, rather than neatly distinct phenomena as is the situation with “port.”

Unless otherwise stipulated, in this book “information” will be taken to mean any difference that makes a difference to a conscious, human mind (Bateson, 1972, p. 453). In other words, information is whatever appears significant to a human being, whether originating from an external environment or a (psychologically) internal world. This definition was chosen by the anthropologist Gregory Bateson, after he had struggled for two decades with the inadequacies of mathematical definitions of information. A perceived difference, according to Bateson, is a basic “unit of mind” that can be inferred through study of both humans and animals.

Other authors have employed definitions of information that are similar to Bateson’s. A popular version originated with communication scholar Gerald Miller (1969): information is any stimulus we recognize in our environment. Others (e.g., Bates, 2005b; Dervin, 1976a; Higgins, 1999; Johnson, 1997; Rogers, 1986) have defined information as the recognition of patterns in the world around us.

Obviously, the characterization of information as a difference implies a very broad definition for a common word that has been defined in several distinct ways — with virtually all other definitions implying more restrictions on meaning. That is, many authors have used other words to define a concept that they have called “information,” some of them incorporating specific requirements, such as information must always be

true or useful, or it must be embodied in a form or object, or it must be intentionally transmitted, and so forth.

The reason for adopting this broad (some would say vague) definition is because this book reviews a great many studies from several disciplines and needs to cover a variety of more restrictive concepts. The reason some other definitions are less appropriate will be made clearer in the remainder of this chapter, which first discusses the nature of conceptual explication and then examines a variety of definitions for “information,” each of which identifies a somewhat different concept than the rest.

3.1.1. Explicating “Information”

To discuss and study any concept, we first need to define it. In the process, we may identify and define other ideas that are related to (and sometimes derived from) the concept under study. In the case of information, two related concepts include “knowledge” and “data.” “Information behavior,” “information seeking,” “information source,” and “information use” are among several other, higher-order concepts that build upon the concept of information.

Social scientists call the process of defining a concept “explication.” As discussed by Chaffee (1991), explication is the intellectual process that relates theory to research, which links a focal concept to the ways in which it is studied. For researchers who aspire to direct observation of phenomena, explication eventually results in an operational definition of a concept, a set of procedures used to observe and measure instances of a concept. In this chapter, we deal only with the initial stages of explication — reviewing and analyzing existing definitions — and leave issues of measurement for later chapters. Lambert and Loiselle (2007) offer examples of the later stages of explication.

The process of explication often starts with a word for which we have only a general meaning. At this stage, we have only a nominal, or dictionary, definition for a term — that is, a word is defined by other words. Explication continues by examining what has been written about the concept; we review the publications about it, with an eye toward how different authors have defined and used the concept. In doing this, we may not only find multiple definitions for the term, but discover that some authors have studied the same concept but called it by a different name.

The next step in explication is to analyze the meaning of a term by one or both of two approaches. The first possible approach is a top-down procedure. We distill the discussions of many authors to their abstract core: what is the heart of what they say about the concept? In the case of information, a core idea may be that it is a message expressed in some

medium, and/or that it has the potential of altering a person's consciousness. As Chaffee (1991, pp. 26–27) points out, finding a single, central meaning is unusual, particularly when distinct research literatures are examined.

In the second alternative, the bottom-up approach to meaning analysis, we list all of the subsidiary concepts that make up the focal concept. For "information," we might attempt to list exhaustively all the possible forms that information could take — a Web page, a book, a radio broadcast, a conversation, a handwritten note, e-mail, and so on. This is a massive task that would be subject to change over time as new forms of information appear or are identified. Listing all examples of information has been one way that researchers have guided the observation of the concept.

Whichever means of analysis is chosen, the eventual result is a more abstract definition than one defined by near synonyms. The definition may be expressed as a series of critical distinctions between the focal concept and related concepts; for example, what is the difference between information and data? Or, instead, it may simply identify attributes that serve to identify something as an instance of the concept; for example, a book is an instance of information because it contains symbols that can, or are intended to, inform someone.

The remainder of this chapter will consider, through literature review and distillation, the various definitions of information and their key distinctions.

3.1.2. The Concept of Information

Ordinarily, we both use and hear the word "information" without much concern for its definition; we know what we mean when we use the word. At first glance, the *Oxford English Dictionary* definition seems adequate: "(1) The action of informing. The action of telling or fact of being told of something. (2) That of which one is apprised or told; intelligence, news." This nominal definition reveals at least one important distinction: the term may be used to indicate either a process (informing) or a kind of message (news).

Further distinctions lay buried in the nominal definition, as a series of publications have made obvious. One explication of the term (Wellisch, 1972) uncovered eight distinct definitions of information without any common elements. Not long afterward, Wersig and Neveling (1975) identified 17 unique definitions, which they grouped into six broad categories. Summarizing 30 years of commentary, Levitan (1980) declared that 29 different concepts had been associated with the term information. A review by Schemert (1993b) includes a selection of 22 definitions

written between 1968 and 1989. More recently, Frické (2009, p. 139) points out a dozen characterizations of information, while allowing that there are "many more." How has the concept of information been used such that so many definitions have resulted? The empirical work of Shenton and Hayter (2006) demonstrate that it is not only scholars who are unclear on the concept, ordinary people are confused about what the word means, too.

The central difficulty is that the word "information" has been used to denote several different concepts. The adoption of the term by multiple disciplines is part, but not all, of the problem. The same term has been used to refer to, among other phenomena, sensory stimuli, mental representations, problem solving, decision making, an aspect of human thinking and learning, states of mind, the process of communication, judgments about the relevance of information to information needs, the content of subject specialties, recorded knowledge, and particular objects that carry information such as documents. It is no surprise that scholars have struggled to come up with a formulation that promises to condense most of these meanings into one universal principle or attribute.

3.1.3. Typologies of Information Concepts

Let's first examine four parallel attempts to identify different "families" of information definitions. Two attempts to distinguish types of information concepts, one from 1976 and the other from 1992, illustrate how periodically we revisit the problem of defining information. Two articles by Dervin (1976a, 1977) set the stage for the development of the sense-making school of thought (see Chapters 4 and 7) regarding information seeking. In her articles, Dervin posited three types of information, based on the writings of philosopher Popper (1972):

1. Objective, external information is that which describes reality (but never completely so).
2. Subjective, internal information represents our picture or cognitive map of reality, the structures we impute onto reality.
3. Sense-making information reflects the procedures and behaviors that allow us to "move" between external and internal information to understand the world, and usually to act on that understanding as well.

Dervin argues that to look at information in such a way has several advantages. For example, it acknowledges that legitimate inputs may come from inside us, rather than viewing the only important information as arising from external sources. In a similar way, this view does not privilege

formal information systems (e.g., books) over informal sources (e.g., friends, relatives, or coworkers); consulting the latter is a much more common approach to understanding than are the former channels.

In a manner reminiscent of Dervin, Ruben (1992, pp. 22–24) places information conceptualizations into three “orders.” The first-order information captures information as “environmental artifacts and representations; environmental data, stimuli, messages, or cues.” This environmental sense of information consists of “stimuli, messages, or cues, waiting to be attended to.” The second-order information is that which is “internalized, individualized appropriations and representations.” Here Ruben identifies information as something that is “transformed and configured for use by a living system,” internal representations that include “semantic networks, personal constructs, images, rules or mind.” And the third-order information is that which is “socially constructed, negotiated, validated, sanctioned and/or privileged appropriations, representations, and artifacts.” The third-order information, then, is the social context of information.

Dervin’s and Ruben’s types of information are parallel but not identical, especially in the terms and examples they use to describe their third category, which for Dervin is decidedly intrapersonal, abstract, and process-oriented. For Ruben, the social context is external, is socially constructed, and may encompass physical objects like books (which seem to fall under Dervin’s objective category).

Two other typologies, both from the 1990s, are also somewhat parallel, but they bear only a modest resemblance to the Dervin and Ruben schemes. Buckland’s (1991a) widely cited typology portrays uses of the term “information” as falling into three categories. The first category is information-as-process, which refers to the act of informing, the communication of information, and how a person’s state of knowledge is changed. A second sense of information is information-as-knowledge, a usage of the term denoting that the information is perceived in the first category (i.e., the knowledge communicated). The final sense of the term is information-as-thing, in which “objects, such as data and documents … are referred to as ‘information’ because they are regarded as being informative.”

Buckland takes great pains to explain the difficulties inherent in such a typology, pointing out the intangible nature of the first two categories (which makes them difficult to observe), the issue of intentionality (some definitions of information take for granted an intention to communicate), and the problem that any object in the world might potentially be informative (“if everything is information, then being information is nothing special”). He concludes that it is essential to investigate information-as-process, even though information-as-thing cannot be dismissed as a focus of study.

The second typology is similar to that of Buckland but breaks out his two categories of information-as-thing and information-as-knowledge into three overlapping conceptions of information. Altogether, McCreadie and Rice (1999, pp. 47–58) identify four distinct “conceptualizations,” the first of which is information as a resource or commodity. Under this conceptualization, information is something that can be “produced, purchased, replicated, distributed, sold, manipulated, passed along, controlled,” such as a message that travels from sender to receiver, with or without some kind of payment in exchange.

The second type of information is characterized as data in the environment, that is, “objects, artifacts, sounds, smells, events” that may be perceived in the environment. This category takes into account the potential for unintentional communication of information, such as when one observes and interprets natural phenomena.

McCreadie and Rice’s third type of information concept is that expressed as a representation of knowledge, such as in “documents, books, periodicals.” Finally, their fourth type of information is as a part of the communication process. That is, information is meanings that are created as people go about their lives and try to make sense of their world.

While at first glance it may seem that these latter two typologies characterize types of information in similar ways, there are several differences, particularly in the distinctions they make between representation, thing, and (in the case of McCreadie and Rice) resource. McCreadie and Rice use “documents, books, periodicals” among their examples of representations, whereas Buckland uses those as examples of “things.” It seems that McCreadie and Rice are trying to make finer distinctions than Buckland regarding possible embodiments of information.

Ultimately, the typologies of Dervin, Ruben, Buckland, and McCreadie and Rice are each distinct from one another in several ways. At least the distinctions these authors make are useful in illustrating the many ways one could parse the attributes of the information concept.

3.2. Definitions of Information and their Problems

The typologies discussed earlier in this chapter fall short of providing specific definitions of “information.” Rather, their intention is to show that there are distinct usages of the term rather than a single universal usage.

Nevertheless, many authors have attempted to create a general definition of information that at least would be adequate for some areas of investigation. As we shall see, many scholars have incorporated into their

definitions specific and powerful assumptions regarding the nature of information.

We begin our examination with the most influential definition of information, one that developed a half-century ago for the study of signal transmission in broadcasting and telephony. Now it might seem odd that a definition for such a mundane concept as information should come to us from a highly specialized field as telephone engineering. In fact one writer (Tor Nørrestrand, 1991; English translation, 1998) jokes that:

there are plenty of grounds for a conspiracy theory of the most devious kind: that the notion of information was invented and developed by engineers from big private corporations who then made a profitable business out of having the rest of us talk about truth, beauty, meaning and wisdom — on the phone. (p. 96)

In Nørrestrand's view, this development was unfortunate because it shifted our attention away from the more important elements involved in information — the senders and receivers of messages — and toward characteristics of the carrier.

3.2.1. The Influential and Restrictive “Information Theory”

The first widely recognized attempt to define information, the misnamed “Information Theory” (properly called “The Mathematical Theory of Communication”), is still frequently invoked to describe the nature of information. The popularity of information theory cannot be overemphasized: a review of two decades (Zunde, 1984) listed over 400 selected citations to this theory; by now, the number of references to information theory surely runs into thousands.

Fifty years ago, the works of Shannon and Weaver (1949) on communication of messages gave rise to a popular conception of information. Shannon, an engineer at Bell Labs, was concerned with the fidelity of telecommunications signals, such as those sent over radio waves, and the determination of the effective capacities of telecommunication channels. It was Shannon who came up with a model of communication as a process of signal transmission. His model of communication became the basis for applying measures to parts of messages based on the statistical

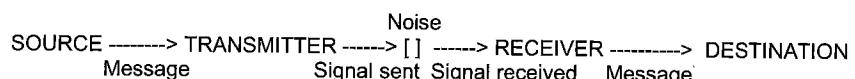


Figure 3.1: Shannon's model.

probability of their appearance — a technique that led to improvements in signal transmission because it helped to predict the likelihood of errors and decide how to correct them, such as the sending of redundant portions of a message.

In Shannon's famous diagram (Figure 3.1), the source and destination of a message were seen as being at the opposite ends of a chain, linked by a message converted by a transmitter into a signal sent over some kind of channel to the receiver, which converts the signal back to a message for delivery to the destination. The channel was acted on by sources of noise, which could disrupt or distort the message.

Along with the diagram came both a definition and measure for the concept of information, as it is encoded in a message. Shannon's definition of information was based on the notion of entropy, a measure of the degree of disorganization in a system that reflected a tendency for any state of affairs to lose order and become more random. In signal transmission, noise is the vehicle for the effects of entropy, that is, noise degrades the signal to some degree. Messages are organized exchanges (e.g., grammatical sentences) based on selections from an agreed-upon set of signals (phonemes, words, letters, etc.). The requirement that the message elements are selected from a fixed universe of possible elements has led some scholars to refer to it as a theory of “selective information” to contrast it with theories of “semantic information,” that is, theories concerned with the meaning of messages.

The effects of entropy lead to more randomness in messages that, in turn, leads to higher levels of uncertainty. In Shannon's view, these higher levels of uncertainty imply the potential for more information in the message. At the opposite end of the entropy scale would be messages that are highly organized — and thus familiar to the receiver — but which tend to carry little “new” information. Shannon's theorems dealt with statistical probabilities associated with the selection of signals from a well-defined set. However, subsequent applications of Shannon's work tended to interpret the theory in terms of uncertainty reduction for the receiver of the signals — what Ritchie (2003) calls “epistemological probability.”

Such a definition of information is somewhat counterintuitive because we tend to associate information with certainty, rather than uncertainty (Miller, 1983a). In fact, Shannon had been advised by computer scientist John von Neumann to call his concept “entropy” rather than “information” (or “uncertainty,” another near synonym) because entropy was a term less likely to be confused with everyday meanings associated with the word “information” (Campbell, 1982, p. 32; Machlup & Mansfield, 1983, p. 48). Indeed, these opposite or “negative” forms of Shannon's definition appeared in the writings of physicist Leo Szilard in 1929 and philosopher Charles Pierce in 1878 (Morowitz, 1991).

To demonstrate how easy it is to misunderstand Shannon's notion of uncertainty when we apply it to human communication, Miller (1983a, p. 495) provides the example of the sentences "Rex is a dog" and "Rex is a mammal." The latter sentence contains terms less likely to appear in everyday usage, so according to Shannon's measure it would carry more "information" that is, a rarer, more surprising, message. But the term "dog" is more specific than "mammal" (which could be a dog, a bat, a dolphin, or many other creatures); semantically, therefore, we would judge that "Rex is a dog" carries more information, reversing the logic of Shannon's measure. It is all too easy to misinterpret Shannon's definition of information outside the realm of signal transmission (Losee, 1997).

Common misunderstandings of Shannon's information theory are partly attributable to his coauthors and advocates. Warren Weaver, a physicist, was invited to write an introduction to two journal articles by Shannon that the University of Illinois was publishing under the title "The Mathematical Theory of Communication." In his introduction to Shannon's work, Weaver speculated on how Shannon's model of signal transmission might be applied to human communication. Weaver acknowledged in a later publication that the model could be taken too literally: "Information must not be confused with meaning," Weaver said (1949, p. 8). Thus, he anticipated that some scholars would attempt to extend the theory to the subjective interpretation of signals by humans.

Nevertheless, Weaver's broad-ranging analogies and speculations became conflated with the very limited theorems devised by Shannon, much to the latter's chagrin; according to Ritchie (1986, 1991), it is Weaver's extrapolation of Shannon's model to which most writings refer, not to Shannon's original explanation and theorems. To confuse matters further, Shannon himself was inconsistent in his use of the terms uncertainty and entropy (Cole, 1993).

Despite its flaws, Shannon's simple depiction of signal transmission as linear, one-way process was seen by many scholars as an adequate model of human communication. Additional interpretations (e.g., Berlo, 1960) resulted in inevitable simplification and distortion of the model. David Berlo's famous "Source-Message-Channel-Receiver" model (Rogers, 1986, pp. 86–90) dropped the "signal" component of Shannon's model. Conflating the concepts of message and signal ignored an important distinction between meanings (messages) and their encodings (signals).

For several decades, various simplified versions of Shannon's model became the basis for studying the exchange of messages among people (Rogers, 1994). As Shera and Cleveland (1977) put it:

Everybody tried to get into the act, hopeful that Shannon's magical formula would unlock countless information secrets and give a

quantitative measure for laying a scientific theoretical foundation for practically every major field lacking one. Unfortunately this overextension was generally an intellectual get-rich-quick scheme and, in the long run, most of the hopefuls fell to the wayside. (p. 261)

A 1974 article by James Watt and Robert Krull can serve as an example of how some researchers applied Shannon's concepts. In a study of television viewing habits, Watt and Krull (1974) noted that other researchers had no classification system in common for program contents. Some would use categorizations like "news, mystery drama, situation comedy, quiz-audience," while others used "documentary, crime-detective, comedy-variety, game shows" to cover the same content. Obviously, these variations posed problems for researchers trying to test for effects of television viewing, given that the results of independent investigations were not directly comparable.

Instead of subjective classifications of content, Watt and Krull argued that "structural or form characteristics of the program may also have an effect on the audience"; therefore, they proposed a "content-free measure of television program form" that applied "information theory entropy terms" to features that appeared on television screens (1974, pp. 44–45). In a study of adolescent viewing, they developed several formulas for measuring various aspects of a broadcast. For instance, "verbal time entropy is defined as the degree of randomness of the time of audible behavior on the part of characters in a program"; a formula measured this in terms of a negative, logarithmic function of the time that a series of television characters produced audible sound. I do not need to reproduce Watt and Krull's various formulas to suggest that although they did indeed carefully measure "nonrandom viewing patterns," their results are difficult to interpret in terms of what normally concerns us about television viewing and its effects. Our judgments about television programs are almost always based on the content of the program, whereas their measures deliberately ignored the content. Nevertheless, Ritchie (1991) describes Watt and Krull's work as among the more successful applications of Shannon's entropy measure, along with Finn's (1985, 1986) studies of unpredictability in news articles.

Eventually, the Shannon and Weaver model came to be seen as inadequate for expressing many of the important features of human communication. As early as 1969, Donald MacKay complained that:

communication engineers have not developed a concept of information at all. They have developed a theory dealing with only one particular feature or aspect of messages "carrying" information — their unexpectedness or surprise value. (pp. 56–57)

Nørretranders (1991/1998, p. 96) observes that Shannon's view of information equated it with "something completely meaningless, something closely related to disorder quite unlike what the rest of us understand by the everyday word 'information' — meaning, content, overview, order." It is what Brier (1992) calls a "mechanistic concept of information," which reduces human cognition to the level of computer processing. Shannon's so-called "information theory" simply did not adequately reflect the way in which people interpret and assess the "meaning" of messages. As the Canadian sociologist Klapp (1982) concluded:

Meaning, being subjective, and referring to synthetic or holistic properties that cannot be reduced to the sum of parts, might be called a higher sort of information that does not come easily, let alone inevitably, from a growing heap of mere information. (p. 58)

3.2.2. Five Problematic Issues in Defining Information

Despite its popularity, Shannon and Weaver's implied definition of information contains several assumptions and requirements that differ from the ways we usually think about and experience similar phenomena in everyday life. Their definition is useful only in a very limited sense.

The Shannon and Weaver definition is not alone in posing such problems. Most information concepts contain assumptions regarding five issues that often turn out to be problematic when we try to apply their definitions. The five types of assumptions are about the following:

Utility: Does information, in order to be information, have to have some kind of effect, some sort of usefulness for humans? If not, what would be the point of talking about it? In particular, must information reduce uncertainty about something? If information does not reduce uncertainty, must it be useful in some other way — for example, in providing entertainment, emotional uplift, or some other kind of useful stimulation?

Physicality: Must information always take on some physical form, such as a book, the sound waves of human speech, or a natural object that embodies some kind of data? Is it even proper to discuss what people know, or believe, as being information? A related, and perhaps distinct, issue is whether information (or at least its effects) must be directly observable. If the effects are not directly observable, then how can it be the subject of scholarly study? This latter question bears on Belkin's seventh and eighth requirements, which he calls methodological, and is addressed in Chapter 6.

Structure/Process: Must information be structured in some way? That is, must it be composed of elements in fixed relations to one another, or in some way consist of a complex "whole," such as an image? Or is information a process, some kind of function, a series of steps — a sort of recipe?

Intentionality: When studying information, is it necessary to assume that someone (or something) intends to communicate it to another entity? Or is some information simply out there in the environment, to be perceived and interpreted by a sentient organism? For instance, we can imagine circumstances in which information is not communicated with a *purpose* (i.e., a utility) in mind; a glance at threatening clouds informs us that rain is imminent, but in this case the generator of the message is an aspect of the natural world and has no intentions.

Truth: Must information, in order to be information, be true? Is it improper to call something information if it is demonstrably false? If so, then we need another term for that which is untrue, such as misinformation.

Let's examine each of these issues in turn and consider what various authors have written about them.

3.2.3. Utility as a Requirement

As Dretske (1990) says (in "Putting Information to Work," his essay in a volume about language and cognition), the concept of information isn't fruitful if it doesn't account for an effect of some kind:

Information isn't much good if it doesn't do anything. ... a difference that doesn't make a difference isn't really a difference at all. ... I mean a causal difference, a difference in the kinds of effects it has. (p. 112)

Dretske's concern is the relationship of "information" to causation and hence explanation of behavior; he doesn't think that information is the same thing as a "belief" or a "meaning." We can accept Dretske's claim at a superficial level: it is not worth discussing a concept that makes no "difference" in the world. Yet this leaves many questions unanswered: what, how, and when are these effects happening?

First, it is necessary to say something about the relationship between information (or knowledge) and power (see Braman, 1989, 2006, for a fuller discussion of this topic). Suffice it to say that information can have powerful effects on humans, but usually does not. Yes, it is easy to imagine that, if we

knew in advance which team would win the world cup or which stock would increase in value, we could easily turn that information into a large sum of money — which in turn could be used in various, powerful ways — but the accurate prediction of future events is rare. Of course, if we know other people's secrets, we may be able to compel them to do our bidding by threatening to disclose the secrets — an unusual and perverted kind of power. And it is true that having specialized knowledge (e.g., of medicine) may grant one certain privileges and enable one to attract money. But such "formal knowledge" is closely bound to issues of performance, competence, and (especially) institutions and social relations. As Friedson (1986) says, knowledge must have agency in order to exercise power. Many discussions in the "information is power" vein tend to underplay the complex social relations (e.g., organizations like universities, governments, and corporations) or material objects (e.g., properties or weapons) typically involved in accumulating, maintaining, and using power.

In contrast, most of the "information" in our possession has little value or effect in and of itself. Most of what we know results only in subtle changes or uses in the real world. Some examples of relatively "powerless" information are the current time and temperature, the names of capitals of the world's nations, or what the person next to me just said. Nevertheless, Kari (2007) offers some evidence that information can have effects of both a psychological and a physical nature.

A more specific utility of information in which there has been more scientific interest and evidence is the ability of information to reduce uncertainty. From the 1950s through the 1970s, definitions of information proliferated, most of them incorporating the uncertainty aspect of Shannon's model. A review by Bouazza (1989) reflects the majority view regarding the role of uncertainty in these definitions: "The most cited and perhaps the most useful definition of information is 'that which reduces uncertainty'" (p. 145).

In the interest of brevity, we ignore definitions that try to preserve major portions of Shannon's definition (e.g., those discussed by Artandi, 1973; Fairthorne, 1975), as they are too narrow to be compatible with common-sense notions of information. The most common drawback of these early attempts is that they insist on defining information in terms of uncertainty reduction — typically in the execution of a choice or decision. For example, Wersig and Neveling (1975) declare:

The basic term "information" can only be understood if it is defined in relation to ... the information needs of people involved in social labour ... Either as reduction of uncertainty caused by communicated data. Or as data used for reducing uncertainty. (p. 138)

Like many such definitions, Wersig and Neveling's definition implies that information must be useful ("involved in social labour") and intentional ("communicated").

Similarly, Everett Rogers defines information in terms of reducing uncertainty in a decision task, as "patterned matter-energy that affects the probabilities of alternatives available to an individual making a decision" (Rogers, 1986, p. 85). Other writers leave out the assumption that a task is being performed, but they still cling to the uncertainty component, in which information is whatever "removes the doubt, restricts the uncertainty, reduces the ignorance, curtails the variance" (Nauta, 1972, p. 179).

The emphasis on uncertainty continues to current times. For example, one book (Klir, 1996, p. vii) on the nature of information begins with the words: "The concept of information, which is the subject of this book, is intimately connected with the concept of uncertainty ... information in a given context is obtained by a cognitive agent whenever relevant uncertainty is reduced." This is a typical approach for defining information among economists. Hirshleifer (1973), for example, discusses the probability of signals but draws a contrast with the way that probability is applied (p. 31):

Uncertainty is summarized by the dispersion of individuals' subjective probability [or belief] distributions over possible states of the world. Information, for our purposes, consists of events tending to change these probability distributions. A rather different concept of "information" is employed in communications and statistical theory, according to which a dispersed probability distribution is called less "informative" than a concentrated one. This latter concept uses the term "information" merely as a negative measure of uncertainty.

It is important to note once again that while this writer refers indirectly to Shannon's work, he is actually discussing epistemological probability — likelihood estimates in the "real world" of human communication.

The idea that information must be useful to be information has been undermined in several critiques, most notably by Fox (1983) and Losee (1997). The latter observes that:

a good definition or theory of information ... should bear some resemblance to the natural language notion of information but need not adhere to it when the natural language definition loses its generality and explanatory power. This happens when the common language definition of information, for example, becomes conflated with the notion of *useful* information, that is, information is understood to be in all cases useful. For those accepting this concept of information, if it is not useful, it is not information.

Requiring that all information be useful limits the domain of discussions about information to cognitive processes that can “use” something; it excludes the information carried by a subatomic particle which is not sensed by a cognitive process. We try to avoid excluding information phenomena. (p. 257)

Neither must information (intentionally communicated or otherwise) automatically reduce uncertainty. To use an example from Fox (1983), suppose that I tell you something but you do not believe me. I have not reduced your uncertainty. Or suppose I tell you that the stock market crashed this morning. Last you heard, the market was going up, so perhaps I have reduced your uncertainty about the overall direction of the market (assuming you have been wondering about it), and yet I have probably created uncertainty within you regarding collateral knowledge (e.g., the values of your individual stock holdings). As Fox points out, there are other scenarios in which communicating information might actually increase uncertainty, rather than decrease it.

Some other disciplines that initially exploited uncertainty reduction have recently questioned the ubiquity of such a psychological drive. For instance, in the interpersonal communication literature, we have the emergence of problematic integration theory (Babrow, 1992) and uncertainty management theory (Babrow, Kasch, & Ford, 1998). Both of these theories question the assumption that humans always strive to reduce uncertainty. While most discussions emphasize the negative effects of uncertainty, some other scholars (e.g., Brashers, Goldsmith, & Hsieh, 2002; Huber & Sorrentino, 1996; Sorrentino & Roney, 2000) identify positive psychological effects of uncertainty — for instance, it may provide “hope.”

Uncertainty management theory, in particular, highlights how people sometimes deliberately increase uncertainty. Uncertainty management theory holds that uncertainty is experienced “not simply as an uncomfortable tension demanding reduction” (Bradac, 2001, p. 463) but as feelings and cognitions that can be managed in other ways as well; these may include “seeking instead ambiguity and even confusion” (Bradac, 2001, p. 471). This is because “individuals may use uncertainty as a tool ... sometimes this cognitive state will be cultivated, rather than eradicated” (Bradac, 2001, p. 464).

Two examples of the deliberate increasing of uncertainty in interpersonal conversation are apparent when a physician must deliver a threatening diagnosis to a patient. One party to the dyad, the physician, might choose to provide an uncertainty-increasing message when he or she believes the patient is certain of bad news (Ford, Babrow, & Stohl, 1996); the patient, in turn, might avoid information in order to maintain uncertainty, or even seek out uncertainty-increasing information (Brashers et al., 2000). In both cases,

increased uncertainty might actually provide some increase in comfort for the patient, even though in a way that might compromise his or her treatment. Affifi and Weiner (2004) give other examples in the health care context whereby individuals deliberately “avoid relevant information” (p. 182).

Other scholars have also debated the effects of new information on uncertainty. Lang, Newhagen, and Reeves (1996) imply that cognitive capacity limits may prevent new information from reducing uncertainty. Berlo (1977) notes that information always reduces uncertainty in the “now,” but in the long term it may have the opposite effect. Yovits and Foulk (1985) conducted an empirical study in which they tested the assumption that information always reduces uncertainty; they found that sometimes new information made their subjects less sure that their evaluations of a problem were correct. Similar conclusions are reached by Kellermann and Reynolds (1990) and Robertson (1980).

Although uncertainty is not satisfactory as a basis for defining information itself, it is nevertheless an important concept for information seeking. Even though information can be encountered in a passive way, actively acquiring information implies recognition of uncertainty or anomalies at some level. Kuhlthau (1993b) makes good arguments for considering uncertainty as a beginning stage in the process of finding information, and Yoon and Nilan (1999) demonstrate that one cannot study uncertainty without considering what informants already know (i.e., certainty).

What about other effects of information? That is, if information does not always reduce uncertainty, must it have some other utility to be considered “information”? Some authors include in their definitions of information almost any kind of stimulation that humans find useful. Examples of such stimulation could include sound (from music or from a waterfall), sight (the words of a novel, the images of a painting, a photograph, or a film), or touch (the feel of warm sunshine or cool water on skin). Each of these sensations “tells us” something. In Chapter 5, and obliquely below under “Truth,” I argue that such stimuli include potentially useful information. Therefore, if information must have an effect, it should extend to outcomes beyond “reducing uncertainty.”

3.2.4. Physicality as a Requirement

Everyone acknowledges that information can have a physical form (e.g., see Buckland, 1991a; McCreadie & Rice, 1999), but few explicitly argue that it must. Indeed, many scholars take pains to state that a more useful conceptualization of information is as a phenomenon that exists apart from

physical media; that is, that we should not think about information as primarily something found in human-created messages like printed texts.

However, as Krippendorff (1984, 2009) has pointed out, at some level, information must have a physical form. He argues in favor of information defined as “a change in an observer’s state of uncertainty caused by some event in his world” (*sic*, 1984, p. 49). This, of course, places Krippendorff’s definition in the “uncertainty requirement” camp we have already discussed. But the interesting aspect is how he analogizes information to energy:

First, neither energy nor information exists in a vacuum. Both are embodied in material processes to which one must refer. Just as one can speak of energy only in conjunction with some specific resource, fuel or storage capacity ... so one can speak of information only in conjunction with a physically identifiable source, a message or a situation as described by an observer, and relative to what he already knows. Second, energy and information are measures of work Information is a measure of the (intellectual) work required to distinguish, to a degree better than chance, among a set of initially uncertain possibilities. (1984, pp. 49–50)

Certainly information in the sense of thoughts has a physical dimension: the electrical impulses of a human nervous system. In any event, it does not pose much of a restriction to insist that information must have a physical component, in the sense that energy does. In Krippendorff’s case, it is much more of a restriction to require that information reduce uncertainty — as his Shannonesque conceptualization makes clear.

3.2.5. Structure/Process as a Requirement

Other families of definitions avoid the uncertainty concept through use of analogy — typically to a structure or process — and sometimes require intentionality to do so. Frické (2009, p. 139) calls this characterization of information “a state or structure transformer.”

Boulding (1956) used the analogy of an image or a “picture in our head” in his popular characterization of messages and meaning. In Boulding’s view (1956, p. 7), “the meaning of a message is the change which it produces in the image” — that is, the image of reality (or a portion of it) that exists in someone’s mind. This conception of information is similar to the one offered by Bateson (1972, p. 452) at the start of this chapter, defining “change” as “difference which occurs across time.”

Expanding on Boulding’s analogy, Pratt (1977) defines information as an event: “That which occurs within the mind-upon-absorption-of-a-message”

(p. 215). That is, information (or an “informative event”) is what we call a change in one’s mental image. For Pratt, then, information is the event that changes someone’s image of reality. Like those of Boulding and Bateson, Pratt’s definition has ties to the “internal information” described by Dervin and Ruben earlier in this chapter.

The “image” metaphor is evoked by Donohew and Tipton (1973) as well. But, for them, the idea of “image” is not something like a holistic picture, but rather a complex mental structure of parts and subparts:

An individual’s “image of reality” is divided into three parts. First are the goals, beliefs, and knowledges [sic] which an individual has compiled as a result of his lifetime of experiences. These cognitive “objects” are defined as any concepts, issues, material objects, or ideas which exist psychologically for a person ... The second part of an individual’s image or reality is the concept of self. This includes an evaluation of his ability to cope with various situations ... The third part of the image of reality is an information-handling “set” developed out of past experiences. The “set” probably controls the selection of information used by the individual to cope with the environment. Here we are talking about an individual’s information-seeking and processing “styles.” (pp. 246–247)

Similarly, MacKay argued that information must be considered in the context of a hierarchy of goals (Cornelius, 2002, p. 413).

Other authors have conceptualized information as a structure or organization of experience and sensory data, for example, Thompson (1968) and Belkin and Robertson (1976). Following Thompson’s definition, Belkin and Robertson (1976) state that “information is that which is capable of transforming structure” (p. 198) — in other words, it changes the knowledge state of the recipient. A parallel characterization comes from MacKay (1969) in which information is “that which does logical work on the organism’s orientation” (p. 95).

Belkin (1978) notes that characterizing information as something that transforms knowledge structures has its problems, but it relates well to information as it has been defined in a variety of disciplines. Given that information-as-process assumes that a process has an effect on some entity — as an alteration of a mental image or the creation of meaning in a human mind — the process and structure views of information are analogically similar. In some definitions of this type, individual authors have added one or more restrictions to serve their purposes. Belkin and Robertson (1976), for example, are concerned with document retrieval systems and therefore assume that messages are intentional and that messages are represented by texts: “a collection of signs purposefully

structured by a sender with the intention of changing the image structure of a recipient" (p. 201).

Cole (1994) notes a dilemma that accompanies the assumption that information changes a cognitive structure: if "new" information can modify knowledge structures, then "old" or "expected" information must not modify knowledge structures in the same way, or at least at the same time. New, or "pure," information must be extremely rare because such information is completely unanticipated, and there is a natural tendency not to recognize, see, or perceive that which is unanticipated. Information, then, has the quality of being unexpected and expected, old and new, at the same time. Therefore, for information to be unexpected and expected, old and new, at the same time, information must enter the perceptual system in at least a two-stage process.

More recently, Losee (1997) has advanced a general and coherent definition of information. He attempts to resolve some of the conflicts between definitions of belief, knowledge, information, and misinformation by viewing information generically as processes that produce outputs. The processes, or functions, may be invoked by humans, or machines, or other entities. The inputs into the process can be perceived from the environment or retrieved from human memory. The output of the process (e.g., the value taken on by a variable) is informative about both the original inputs and the process that produced the output from those inputs. Thus, Losee defines information as "the values of characteristics in the processes' output" (p. 256).

In Losee's view, by examining any output we can usually infer something about the process that created it. Examining a tree, for example, informs us about its origins, soil, moisture, and growth process. Or a cake, which is created through a procedure that includes ingredients, instructions, and heating, may be inspected to determine some, but not all, of the ingredients and the process by which it was created.

Defining information in terms of the output of a process, Losee concludes, moves beyond discipline-specific definitions (such as the tendency for decision theorists to define information in terms of uncertainty reduction) and provides a link between various studies of information. Yet, while Losee's definition sidesteps some of the criticisms of Fox regarding information-as-process, it substitutes the vagueness of "process" with a mysterious "function" that takes input and returns a value to be attached to variables (or "characterizations"). This in turn begs the question of how variables emerge in the first place and what determines the nature of the functions. These are not necessarily questions that Losee is responsible for answering, but rather problems to be faced by researchers trying to apply his definition of information. The chapter by Hörrz (1996) is yet another commentary on the notion of information as structure or process.

Fox (1983) criticizes all structure-based definitions as failing to provide a clear definition of "structure" itself. Fox also makes compelling arguments that information cannot be considered either an event or a process, although he also provides counterexamples supporting the process view of information.

Fox himself favors defining information as a type of "telling," as represented in propositions. It would be impossible to fairly convey Fox's arguments in any shorthand version, presented as they are in 213 pages. It may suffice to say that this is the most extended dissertation on the subject thus far. Fox summarizes his conclusions in this way:

Information need not be true, though misinformation must be false; information need not be believed by anyone; information need not originate with a reliable informant, but it must originate with someone in an appropriate position to know. Ontologically, information is propositions [the identification of which] depends on contextual factors. (pp. 212–213)

Fox admits that his conclusions leave several issues unresolved, including the:

... crucial notion of the amount of information carried by a set of sentences remains unanalyzed. The notion of informativeness remains unanalyzed. ... The details of how meaning determines propositional content as a function of context is not well understood. (p. 213)

In the intervening years since Fox's book, Nørrestrand (1991/1998) has published a text that addresses the relation of context to content. Nørrestrand introduces the term exformation to describe the ways in which messages may refer to a "mass of information" that is "not present" and "explicitly discarded" but nevertheless is understood to be relevant by the receiver and is used in construing the meaning of a message (p. 92).

Nørrestrand provides two examples to illustrate the concept of exformation, first involving an extremely short message and the second involving no message at all. The first example is Victor Hugo's famous query to his publisher regarding the appearance of his latest novel, *Les Misérables*, in 1862. On vacation and out of touch with news about public reaction to his work, Hugo mailed a letter consisting of a single character: "?" His publisher replied, simply, "!" Without prearrangement, both parties understood these exchanges to mean something like the question "How is my book selling?" and the response, "Surprisingly well!"

Nørrestrand's second example corresponds to the saying "no news is good news." When parents do not receive a phone call from their son

away at college, they assume that he is OK and that things are fine. Information has been conveyed without sending a message at all. This example echoes Cole's (1994) observation about "new" information versus "old" or "expected" information: can "no news" be viewed as merely preserving the original knowledge structure or does it still modify the structure, but in a different way?

3.2.6. Intentionality as a Requirement

The manner in which Fox characterizes information has a critical limitation. His analysis is based, by necessity, on propositions expressed in the form of sentences. ("In this work I deal only with information carried by sentences," Fox, 1983, p. 7). Fox notes that his propositions represent "what is asserted to be the case by (someone who writes or utters)" (p. 77). One problem with this constraint is that it returns us to the assumption of a message intentionally sent by a sender to a receiver.

We could call this type of intentionality the "communication assumption" — that information necessarily involves communication, and hence, intention to communicate. Bowers and Bradac (1982) see the presence of intentionality as a key dividing point among rival definitions of "communication." Their examination of 27 metatheoretical discussions of communication finds that 18 of their authors hold that intentionality is a requirement for communication to exist. Few of those theorists have an unusual definition for "intention"; most mean the concept in its usual sense: a "purposeful activity [that] must be explained by 'in order to' as well as 'because' statements" (Bowers & Bradac, 1982, p. 7).

Although the restriction of intentionality may hold true for what is the most important sense of information — the exchange of information between humans (e.g., see Buckland, 1998) — it does not apply to all senses in which we use the word. Information may originate outside natural language propositions, for example, as signs occurring in our environment. Whether we are viewing the natural world (e.g., trees, animals, rocks) or the human-made world (e.g., what people are wearing and doing, or a printed sign that says "exit"), we can take in stimuli that have meaning. The only way to retain the notion of intentionality is to assume that it can refer to either a "sender" ("someone who writes or utters") or a "receiver" (the viewer of the world), but does not necessarily involve both ends of a communication process.

If we believe that people must intend to receive in order to take in information, then information is, in this more limited sense, intentional. Intentionality solely on the part of the receiver was suggested by Westley and Maclean in 1957, and in a discussion of news-seeking behavior by

Westley and Barrow (1959). The latter described "the need of the selecting receiver to be oriented in his extended environment" (p. 431); this assumption would take in the kind of "viewing" that I discussed previously. Theirs was a rather radical conception of communication, because it did not assume that a sender's intentions were involved; hence, Bowers and Bradac count it among the "nonintentional" definitions of communication. However, it is an intentional view of information behavior. This is a different view from that of Stonier (1990, p. 21), who asserts that:

... information exists. It does not need to be perceived to exist. It does not need to be understood to exist.

3.2.7. Truth as a Requirement

Losee (1997) also considers the notion of "misinformation." He notes that information can have various flaws, including inaccuracy, incompleteness, lack of justification, and intent to deceive. Do we need a special label for information that is so flawed as to be untrue?

Traditionally, philosophers have made a distinction of this type regarding knowledge, the common stance being that knowledge is "justified true belief." In this definition, belief is taken to be "the most elementary of our opinions ... characterized by two qualities: ... either true or false ... arrived at either rationally or nonrationally" (Cherwitz & Hikins, 1986, p. 31). "Justified" means that the believer has sufficient, relevant evidence that his or her belief is true. The "justified true belief" definition of knowledge (sometimes called "strong knowledge") has been criticized (e.g., Gettier, 1963), leading to the concept of "weak knowledge": beliefs or views that are true, yet lack justification (Frické, 2009). Frické (p. 140) says that "... weak personal or public knowledge, in its recorded form, is suitable as a view of information."

Patrick Wilson's quote at the beginning of the fifth chapter suggests that the truth or falsity of information is something that we can ignore in discussing information in the abstract sense. For one thing, it could be argued that we rarely know for sure if something (a statement or perception) is true or not; even if a "fact" is demonstrated to be true at this moment, it may be possibly proven wrong a few moments later. Fox (1983, p. 212) and Derr (1985, p. 496) also hold that information need not be true, based upon analyses of usage of the term in ordinary discourse. Buckland (1991b) concludes that

... the question of whether specific bits of knowledge are true is not central to our concerns. We adopt the position that the process of

becoming informed is a matter of changing beliefs. Whether these beliefs are held or denied by others and whether they are compatible with some a priori or fundamental assertion need not detain us. (p. 43)

Some would disagree with this point of view and instead argue that a true-false distinction is worth keeping in defining “information.” Frické (1997), for example, argues that information should be “truthlike” in order to “fit the world” so that we can “succeed in our interactions with the world” (p. 888). Dretske (1981, 1983) also states that information must tell us truly about a state of affairs, such that we can learn from it; he concludes that false information and misinformation cannot be considered to be varieties of information, but rather distinct concepts. These points of view are shared by Floridi (2005), who argues that for semantic information to be meaningful, it must be at least “contingently truthful,” otherwise it is mere “pseudo-information.” Finally, Budd (2011, p. 56) suggests that information “cannot be defined unless within the context of meaning and truth” and that information consists of “meaningful communicative action that aims at truth claims and conditions” (p. 70).

However, we are concerned here with a broad view of information phenomena that fits both real life and empirical studies of real life, not with establishing a philosophical distinction. Studies of information seeking provide many examples in which people value information that they know not to be entirely true. For purposes of this text, then, we will generally ignore any distinction between the truth or falsity of information, unless such is the focus or finding of a given investigation.

3.3. Must there be a Universal Definition of Information?

All the definitions we have examined have taken a stance on one or more of these issues. The distinctions and disagreements among reviewers of definitions are too many to resolve; in short, there is as yet no single, widely accepted definition for the concept of information, even among the few thousand scholars who have written about information seeking. At least among recent reviews, however, there has been some agreement on the types of definitions of information that exist.

To summarize the chapter thus far, we can see that there have been many attempts to characterize information, some of them quite broad (e.g., the image/event/structure/process definitions), whereas others have been very narrowly focused (e.g., the view of information as a selection of signals from a well-defined set of symbols, the reception of which may reduce uncertainty for the receiver of the signals).

Narrow definitions assume one or more of the restrictions discussed earlier in the chapter. They hold that information must be useful, that its transmission is intentional, or that it must be represented in a recordable medium (in written or spoken language, or images), and/or that information must be true (or at least easily verifiable). Shannon’s model contains examples of all of these assumptions. The vast majority of the early definitions and investigations of information seeking include at least the first three of these assumptions.

Before we go any further, it is important to note that defining information in an absolute and final sense is not entirely necessary for the study of information phenomena to proceed. As Artandi (1973) and others have pointed out, all we need are useful conceptualizations of information. Belkin (1978) makes this point most effectively when he says:

... we are not concerned with definitions of information, but rather with concepts of information. The distinction is that a definition presumably says what the phenomenon defined is, whereas a concept is a way of looking at, or interpreting, the phenomenon ... by accepting the idea of a concept one becomes free to look for a useful concept, rather than a universally true definition of information. (p. 58)

If we wish to keep talking about “information,” we may have to give up on a universal definition. Hempel (1952) notes that there are some terms in any conceptual scheme that are so basic that they need not be fully explicated. Hempel calls these basic concepts “primitive terms.” Primitive terms are simply accepted as they are commonly understood. Chaffee (1991, p. 7) provides the example of the concept of a “person,” or a “human.” Perhaps there are some fields in which what we mean by “human” needs to be carefully defined and is subject to debate — in zoology, for example. But, for most purposes of study, we do not need to explain that particular notion. It is when we deal with concepts that build upon the notion of humans — family, community, society — that we are in need of careful definitions for those concepts.

Information can be, and has been, treated as a primitive term as well. Some writers believe this approach is problematic — see Frohmann (2004, p. 86) and Nunberg (1996, p. 110). Yet it has been a common practice in the IB literature; commenting on a variety of recent studies, Vakkari (1997) notes that:

... one of the striking features in many studies was the use of the central concepts, like information, knowledge, information need, seeking, and use as primitive concepts, i.e., without definition. (p. 460)

Can we reconcile the various definitions of “information” with one another? It does not seem so, and perhaps it is not necessary. Although there is scholarly disagreement over the “most rigorous,” or “most easily quantifiable,” or “most productive,” or “most parsimonious” meaning of “information,” these debates have done little to promote a fuller understanding of the concept among a community of scholars. In fact, if anything such discourse has resulted in a fracturing of scholarly effort in studying the phenomenon of information, it has resulted in too many definitions that defy comparison and that provide no common basis for understanding. Allen (1969) and Dervin (1977) discussed using the concept of “communication” instead (see the insightful discussion by Frohmann, 2004, pp. 53–67). Similarly, Furner (2004) argues that there are so many productive substitutes for the concept of “information” that we could do without it all together:

... philosophers of language have modeled the phenomena fundamental to human communication in ways that do not require us to commit to a separate concept of “information.” Indeed, we can conclude that such a concept is unnecessary for information studies. Once the concepts of interest have been labeled with conventional names such as “data,” “meaning,” “communication,” “relevance,” etc., there is nothing left (so it may be argued) to which to apply the term “information.” (p. 428)

Instead, let us treat “information” as a primitive term, as a phenomenon that we all recognize when we see it in its various forms (Fox, 1983, p. 16). Information would then be treated as “anything that exists psychologically for a person” (Carter, 1965; Chaffee, 1991, p. 9). We have only to look around us to establish the fact that information exists in the form of physical objects (what Buckland (1991a) calls “information-as-thing”), and hundreds of studies have documented that people believe that information exists as a psychological object as well — a disembodied result of “becoming informed.”

Fox (1983) observes that the “ordinary notion of information” is one through which “information scientists apparently do succeed in communicating with one another quite effectively regarding information and related concepts” (p. 5). Nunberg (1996, p. 110), as well, notes, “‘Information’ is able to perform the work it does precisely because it fuzzes the boundaries between several genetically distinct categories of experience.”

Allowing a broad definition of information poses problems for operationalization and measurement of concepts, as shall be seen as individual studies are reviewed in later chapters. Yet to argue for any tighter definition of

information would be to limit the scope of this book, which is intended to review a broad spectrum of investigations having to do with information seeking and sense-making. Therefore, in this text, I allow for any definition of information, however, vague or difficult to study. Where more restricted definitions of information apply in the review of individual theories, methods, or studies, they will be made explicit.

It should be noted that the definition “any difference that makes a difference” places at least one important restriction on the scope of information: it rules out the possibility of information existing independently of a knowing mind. For the purpose of this chapter, we assume that a conscious brain must be engaged at some point for information to be said to exist. Otherwise, we are back to the unhelpful stance that “everything is information.”

Two examples, the first suggested by Fox and the second by Buckland, will help to make the restriction clear. First, Fox provides the example of someone who keeps a secret diary that no one else is ever allowed to read. Some definitions of information would imply that, since the content of the diary was never communicated, it cannot be considered information. (Fox uses this straw man to defeat the requirement that a message must be transferred — i.e., received — to qualify as information.) Of course, a diary is a clear example of information-as-thing, and the symbols written in the diary are an expression of a human mind — it is some kind of message, even if never received by any other than its creator. So, yes, a secret diary (an unviewed, human-created record) can safely be considered information.

In the second example, trees could be viewed as carrying information in the form of their growth pattern of rings, which among other things tell us about the amount of rain that fell in a past season. Even if no person has viewed that information, is it not still information? For our purposes, the answer is “no.” It is nothing more than wood until someone both encounters and makes some sense of it. So, if a tree falls in the forest and there is no one there to see it, then it conveys no information.

To conclude this portion of the chapter, it bears emphasizing that in this text we will consider only human information behavior, and the making or meaning or sense, and therefore we interpret information as requiring the involvement of a human mind. (It could be easily argued that animals use information, but they are simply outside the scope of this book; see Bates (2005b) for a distinction between animal and human use of information.) As will be seen in later chapters, a broad conceptualization of information is in keeping with the way the term has been employed in studies of information needs, uses, seeking, and sense-making.

3.4. Distinctions among Information, Knowledge, and Data

A side issue as regards information seeking research is worth noting. Much attention has also been granted to defining the concepts of “data” and “knowledge.” Machlup (1983) examines the issue of whether information is synonymous with data and knowledge, noting that there has been a tradition to treat the three as a hierarchy, with data at the bottom and knowledge at the top. Raber (2003) raises a troubling question about such an arrangement when he says: “But at what moment and how does information become knowledge?”

Machlup (1983) holds that historical usage of the three terms does not fully justify the notion that information is data that has been processed and/or organized. The origins of the term in the Latin *dare*, “to give,” along with the history of its usage, imply that the word data (“the givens”) can be assumptions, facts, measurements, and so forth, expressed in either words or numbers. As Machlup points out (p. 647), many writers claim that data are a “raw” type of information, whereas a few others see information as a type of data. Machlup concludes that there is neither precedent nor need to establish a hierarchy between the two words.

The common notion that knowledge is information that has been sifted, organized, and understood by a human brain is on firmer ground. Brown and Duguid (2000) complain that the two concepts are unfortunately conflated:

People are increasingly eager that their perfectly respectable cache of information be given the cachet of knowledge. Such redefinitions surreptitiously extend the overlapping area where knowledge and information appear as interchangeable terms. Nevertheless ... there do appear to be some generally accepted distinctions between knowledge and information For example, it sounds right to ask, “Where is that information?” but odd to ask, “Where’s that knowledge?” (p. 2)

Machlup (1983) makes the useful point that “information is acquired by being told, whereas knowledge can be acquired by thinking” (p. 644). Through our inner experience of thought, we can form new knowledge without taking in new information from the external environment. Information implies transfer, says Machlup, while knowledge is a state (“knowing”). Knowledge and information are therefore not usually the same, except that “information in the sense of that which is being told may be the same as knowledge in the sense of that which is known, but need not be the same” (p. 644). Hayes (1993) makes a somewhat different point when he says that “knowledge is internal; it cannot be received but must be internally created” (p. 5). Savolainen (2009a) discusses similar issues in distinguishing between knowing and acting in practice.

Other authors sometimes raise the issue of truth, discussed earlier, in the knowledge-information distinction back to the truth issue discussed earlier. Raber (2003, p. 8) notes: “The relationship between knowledge and truth is especially problematic,” whereas Frické (2009) points out that data can be flawed and thus untrue. Yet Dretske (1981, p. 45) makes no distinction between the two in regards to a truth requirement: “Information is what is capable of yielding knowledge, and since knowledge requires truth, information requires it also.”

Decades of arguments about distinctions among the words data, information, and knowledge (and sometimes “wisdom,” too) have not prevented the continued use of terminological hierarchies. Frické (2009) and Rowley (2007) each examine examples of the commonly used DIKW (data-information-knowledge-wisdom) pyramid in which wisdom appears at the top, with knowledge beneath, supported by information and then data lying at the base. Both authors find the underlying assumptions of this structure to be problematic. Echoing arguments made above, Frické (p. 140) advocates a view that “... makes knowledge and information synonymous. Knowledge and information collapse into each other.” Saab and Riss (2011) also demonstrate that different levels in the hierarchy are intertwined in the process of making meaning (or sense) of patterns of stimuli.

In this book, the usage of the terms data, information, and knowledge will generally be used synonymously, because they are usually not clearly delineated in studies of information behavior. Knowledge, however, is strictly a phenomenon of the human mind, whereas data and information are often represented by tangible, physical objects. That information usually has a physical manifestation has often been the key consideration in past studies of information seeking. The way that information seeking is typically approached under the new paradigm, although, is in the sense of knowledge — something in someone’s mind — not primarily as a physical object.

The usage of data, information, and knowledge outlined above represents a necessary simplification of the many definitions and examples that have been discussed in dozens of scholarly works. However, the fine distinctions made among data, information, and knowledge are of little value in most studies of information seeking. This book will treat information as a broad concept, encompassing instances that would be considered unusual by some scholars.

3.5. Summary

This chapter has explored the central concept employed in studying information seeking: information. We have seen that there are widespread disagreements about what would constitute a general definition of information. Most of these disagreements concern the issues of truth,

physicality, intentionality, uncertainty, and utility. The most common types of definitions that have emerged assume that information is something that either reduces uncertainty or changes one's image of reality. In this chapter, I provide examples that suggest that a truly universal concept of information would need to fulfill at least the following requirements:

1. allow for common-sense notions of information used in everyday discourse;
2. allow for unintentional origins of information (e.g., observations of the natural world) as well as for purposeful communication among people;
3. allow for internally generated information (e.g., memories, constructions) as well as externally generated information (e.g., reading a text);
4. allow for types of information beyond that needed for "solving a problem" or "making a decision";
5. admit the importance of informal sources (e.g., friends) as well as formal sources (e.g., data or documents); and
6. involve the human mind, either in the creation, perception, or interpretation of information; to leave out such a requirement is to declare that anything is information and that would leave us with no focus for our investigations.

I have considered numerous distinctions made over the years, but I argue in favor of treating information as a primitive concept that is so basic to human understanding that it does not require a tight definition. To the extent that information needs a definition, it must be a broad one, such as "any difference that makes a difference" (implying a change to the structure of a human mind), or Bates' (2005b) second definition of information as "some pattern of organization of matter and energy that has been given meaning by a living being"). Such characterizations, vague although they are, would allow us to consider what many authors have said about information seeking without having to worry about whether they restricted their observations to phenomenon that must be true, observable, physical, intentional, and so forth.

In the next chapter, I will build on the initial discussion of information to define information needs. Following that, Chapter 5 ventures farther afield to consider more peripheral concepts and behaviors related to information seeking. A review of these other concepts is important in addressing several vexing questions about information-related behavior:

- Why do people seek information?
- What makes information relevant?
- Can information be found without intentionally searching for it?
- Is it possible to have too much information?

- Why do people sometimes avoid information?
- How does information differ from entertainment?

These and other issues are taken up in Chapters 4 and 5.

Recommended for Further Reading

Boulding, K. (1956). *The image: Knowledge in life and society*. Ann Arbor, MI: University of Michigan Press.

An oft-cited book by the unconventional thinker Kenneth Boulding. Boulding uses the analogy of an "image" to discuss how we come to know our world.

Capurro, R., & Hjørland, B. (2002). The concept of information. In B. Cronin (Ed.), *Annual review of information science and technology* (Vol. 37, pp. 343–411). Medford, NJ: Information Today.

A comprehensive discussion of the history of definitions of information.

Frické, M. (2009). The knowledge pyramid: A critique of the DIKW hierarchy. *Journal of Information Science*, 35(2), 131–142.

A highly readable discussion, from the standpoint of philosophy, of definitions and relations among the concepts of data, information, knowledge, and wisdom, touching on the issues of truth, physicality, utility, and structure.

Frohmann, B. (2004). *Deflating information: From science studies to documentation*. Toronto: University of Toronto Press.

While focused on the discourse surrounding the concept of "scientific information," the initial chapters of Frohmann's book give a fascinating account of how information seeking, use, and other behaviors have been studied over several decades.

Schemert, J. R. (1993). Communication and information. In J. R. Schemert & B. Ruben (Eds.), *Information and behavior* (Vol. 4, pp. 3–33). New Brunswick, NJ: Transaction Books.

Jorge Schemert examines a variety of definitions of information. A companion article in the same volume supplies the history of the word itself.

Thayer, L. (1987). How does information "inform"? In B. D. Ruben (Ed.), *Information and behavior* (Vol. 2, pp. 13–26). New Brunswick, NJ: Transaction Books.

Lee Thayer's amusing essay considers several problematic aspects of information as the term is commonly used. He convincingly makes the point that information is always "from the perspective of some observer."

Chapter 4

Information Needs and Information Seeking

Need for information consists of the process of perceiving a difference between an ideal state of knowledge and the actual state of knowledge.

— Lidwien van de Wijngaert (1999, p. 463)

Information seeking is the behavior that is the directly observable evidence of information needs and the only basis upon which to judge both the nature of the need and its satisfaction.

— Bryce Allen (1996, p. 56)

Chapter Outline

4.1. The Motivational Puzzle	77
4.1.1. What Is a “Need?”	78
4.1.2. Needs versus Demands	80
4.2. Four Scholars Ponder Information Needs	81
4.2.1. Seeking Answers	81
4.2.2. Reducing Uncertainty	82
4.2.3. Making Sense	84
4.2.4. The Spectrum of Motivations	85
4.3. The Trouble with Information Needs	87
4.4. Information Seeking and Information Behavior	89
4.5. Summary	91
Recommended for Further Reading	92

4.1. The Motivational Puzzle

Not only has a definition of “information” proved difficult to establish, describing exactly how it influences human behavior has also been controversial. Krikelas (1983) suggests that there are at least as many definitions of subsidiary concepts — including *information need* — as there are for information itself.

What do we mean when we say that people “need” information? This concept is the next most fundamental, building on a primitive notion of

"information." In this chapter I will consider how various scholars have defined the concept of *need* and *information need*.

4.1.1. What Is a "Need?"

It is fitting to begin with a definition of what we mean by a human "need," because it is upon this hook that most writers hang the motivations for information seeking. "Needs" are typically characterized as an "inner motivational state" (Grunig, 1989, p. 209) that brings about thought and action. Other "inner states" may include, for example, wanting, believing, doubting, fearing, or expecting (Liebnau & Backhouse, 1990; Searle, 1983).

The distinctions made among varieties of "need" can be bewildering. An essay by Andrew Green (1990) describes debates over the nature of "needs" that have taken place among political philosophers and social policy advocates. Green identifies four general conclusions about the concept of need (pp. 65–67). First, a need is always *instrumental*: it involves reaching a desired goal. If I "need to know" the chemical composition of heroin, it is typically because I desire to accomplish something with that information. That "something" may be to answer a test question, to write about narcotics for a class assignment, or simply to satisfy my curiosity. It is also the case that my need in those examples is based on some preexisting need: to pass a class, to get a degree, to be a knowledgeable person, and so forth. The key factor is that knowing it will put me at, or closer to, an end state that I want to achieve.

Second, according to Green, "needs are usually contestable. In this they differ from wants." That is, if I say that I *want* to know the chemical formula for heroin, you could hardly argue with me about this odd desire. However, if I say I *need* to know what heroin consists of, you might ask me why; if I replied that I need it to write an essay on drug addiction for an English course, you could perhaps reasonably argue that I do not really "need" to know that fact to write a good essay on addiction (or to demonstrate writing skills, either). A good exploration of Green's first two points are found in Beautyman and Shenton (2009).

Third, need is related to the concept of *necessity* in such a way as to carry, at times, more moral weight. That is, we use phrases like "human need" or "basic need" to refer to goal states (e.g., to be safe or to be loved, in the view of Abraham Maslow) that everyone agrees are good. Doyal and Gough (1984) say that basic human needs include "health, autonomy, learning, production, reproduction, communication and political authority"; Lederer, Galtung, and Antal (1980) suggest that hypothesizing any needs beyond "primary" ones like food and shelter is problematic. Distinctions among primary versus secondary needs have led some information seeking scholars

(e.g., Wilson, 1981) to argue that information is clearly a secondary, rather than a basic, need.

On the other hand, for a half-century the psychological literature has treated "the need for cognition," as though it were a basic, rather than a secondary need (e.g., Cacioppo & Petty, 1982; Cohen, Stotland, & Wolfe, 1955; Henning & Vorderer, 2001). Experiments with sensory deprivation point to pathologies that can result from lack of stimuli (Zubek, 1969). Psychologists like Henry Murray (1938) and Abraham Maslow (1963), as well as economist Manfred Max-Neef (1992), believed that people have a need to understand, and perhaps even to share, what their senses tell them. For example, among the basic needs identified by Murray (1938) were three involving information: "exposition" (the need to provide information to, and educate, others); "harm avoidance" (the need to avoid pain, whether physical or mental); and "infavoidance" (the need to conceal weakness and failure). Max-Neef's (1992) needs include those for learning and curiosity, and for leisure and imagination.

To *deny* a life-sustaining need (e.g., for medical care) would be morally wrong. Yet even regarding such a basic need we might make distinctions; if I said to you that I *needed* narcotics, that statement may be true in the sense that drugs are *necessary* to accomplish my goal — to satisfy my addiction. But it would probably not motivate you to help me satisfy those needs. You might judge my felt need to imply such a "bad" purpose that, even while acknowledging the truth of my need, you would feel comfortable in denying me your help. Perhaps you would rationalize by saying that I really need something else such as a drug treatment program.

This line of thinking leads to Green's fourth point: that need is not necessarily a state of mind, and it is possible to be unaware of one's *true* needs. For example, I may *think* I need to scan every psychology journal in the library to find information about recovered memory syndrome. But an experienced librarian might judge that what I *really* need to do is to search *Psychology Abstracts* on the Internet. Thus, a need may be unrecognized or unacknowledged, undesired, or simply misunderstood, by the individual who has it (Derr, 1983).

So this leads us back to the distinction between needs and wants. Others have needs that we may judge to be "merely" desires, not needs. Obviously it is more difficult to find evidence of needs than it is of wants, because wants more typically result in observable behaviors. We can also ask people what they want, but people may not be able to articulate their needs so easily — this is certainly true if they are not even aware of their needs.

Green suggests that, within the study of information behavior, most attempts to define "need" faded away after some initial attempts made during the 1960s and 1970s. According to his view, most subsequent writers have taken for granted whatever definitions had been proposed up to the

1970s, such as those discussed by Brittain (1970), Line (1974), Menzel (1966a), and Roberts (1975). One notable exception is Richard Derr (1983, 1985) who made worthwhile attempts to define both information and information need during the 1980s. A new book by Charles Cole (2012) promises to bring attention back to the important issue of need, as have recent investigations like those of Lundh (2010) and Lu and Yuan (2011).

4.1.2. Needs versus Demands

So what then is an “information need?” Forsythe, Buchanan, Osheroff, and Miller (1992) point out that

no explicit consensus exists in the literature regarding the meaning of the central concept of “information need.”... In effect, “information need,” has been defined according to the particular interests and expertise of various authors. (p. 182)

“Information need” is often described simply, and somewhat circularly, as a cause of information seeking. Ikoja-Odongo and Mostert (2006, p. 147), for example, say that “an information need is a requirement that drives people into information seeking.” It is problematic, however, to simply assume that actions one takes are solely caused by an earlier, internal “need.” Thus, the relationship between the initial need, and what we do about it (e.g., make a demand of an information source), is critical.

Brittain’s (1970) book ventures into the realm of need by first noting that most research up to that point had instead concerned *demand* — the requests made of an information system, such as a library or database. Data regarding demands were readily available from (or at least easily collectible in) information agencies that supported studies of their users. Demands are relatively easy to measure. Investigations of information demands supported the goals of libraries and vendors to improve their services. So, according to Brittain, “most studies which have purported to be of information needs have in fact been of information uses or, at best, demands” (1970, p. 3).

Another writer, John O’Connor (1968), has suggested three “possible meanings” of information need: (1) a “negotiated” (and thus, refined) version of the *initial* question or demand stated by the inquirer; (2) whatever information provided that actually “helps” the work of the inquirer; or (3) giving the inquirer documents that he or she judges to be “pertinent” on the basis of a comparison with their internal need. O’Connor finds problems with all of these possible meanings, as they may involve differing standards of judgment. That is, different people (e.g., the inquirer, the provider, and groups of colleagues) use varying criteria, at different points in time

(e.g., immediate vs. long-term effects). O’Connor’s concern with the relativity of judgment was echoed by Line (1974, p. 87), who ponders, “who is to say what is ‘necessary’ for himself or others?” Or, as Michael Ignatieff put it:

there are few presumptions in human relations more dangerous than the idea that one knows what another human being needs better than they do themselves. (1984, p. 11)

Few investigations of information seeking delve very deeply into the issue of what human “needs” really are. Not many even question the notion of “information needs.” Rather, most writers assume that information needs exist and are relatively unproblematic. When information seeking researchers *do* refer to more fundamental discussions of how information needs arise, they typically cite one or more of four authors in doing so: Robert Taylor, Nicholas Belkin, Carol Kuhlthau, and Brenda Dervin.

4.2. Four Scholars Ponder Information Needs

4.2.1. Seeking Answers

The earliest of the popular depictions of how information needs arise is that of information scientist Robert Taylor. Taylor discussed the origins of information needs in 1962, and his 1968 article holds the distinction of being one of the most frequently cited items in the information seeking literature (Edwards, 2005; Palmquist, 2005). Taylor’s characterization of the origins of information needs provides a particularly useful frame of reference for the concepts discussed in the next chapter.

Taylor focuses on how and why people come to ask questions at library reference desks. He describes a series of four stages or levels that began with a “conscious or even unconscious need for information ... a vague sort of dissatisfaction ... probably inexpressible in linguistic terms” (1968, p. 182). Taylor calls this unexpressed need for information the *visceral* need. The next level a person reaches is “a conscious mental description ... an ambiguous and rambling statement” which sometimes results in talking to another person about it.

At this point the inquirer may be able to construct a *formalized* (“qualified and rational”) statement of the need. However, the person is not aware whether the need could be answered in that form by any available person or information system. In the fourth and final stage, “the question is recast in anticipation of what the files can deliver.” This *compromised* need may be a question asked of a librarian, or a search statement entered into an

Visceral need ---> Conscious need ---> Formalized need ---> Compromised need

Figure 4.1: Taylor's typology of information needs.

information retrieval system. At this point the question also reflects the kinds and forms of data that may be available (such as books, images, or tabular data) and the ways in which they are organized or indexed. Essentially, the final stage is a *compromise* between how the requester originally envisions the query and how the query must be restated to match the language used by the source.

In summary, Taylor says that our chain of cognition and communication often proceeds as shown in Figure 4.1.

The implications of this conceptualization of information need are several. A perception of need may differ greatly from its ultimate expression in words. There may be “unconscious needs.” Recognition of uncertainty does not always lead to action. And central to the entire process is the ability to communicate one’s thoughts, to “negotiate” questions and answers.

Let’s look at an example of how this typology can relate to a real-life situation. A patron asks a librarian, “Do you have books on philosophy?” It may be that the inquirer in this case actually wants to understand the traditional distinction philosophers have made between “truth” and “truthfulness.” She may start with a broader request that she hopes will orient the librarian to what is coming: a more specific statement about her needs. A helpful response to such a question would not be one of the obvious ones (“yes,” “no,” or “I don’t know”) but rather a request for clarification (“What is it you’d like to find out?”). The inquirer might then follow with a more specific request (“I need a dictionary that would explain the difference between...”), or instead might settle for directions to the philosophy section of a library. Thus, Taylor’s typology helps to explain why people seeking help in libraries may ask questions that are overly general.

4.2.2. Reducing Uncertainty

The notion of *information as uncertainty reduction* is one that dates back to at least to the nineteenth century, according to Morowitz (1991). In the late 1940s Shannon and Weaver popularized this connection between information and uncertainty, although in a counterintuitive way (see Chapter 3). By the 1970s, reducing uncertainty was firmly cemented in scholarly dialogue about motivations for information seeking.

For example, in 1973 Charles Atkin (1973) offered a definition of *information need* as “a function of extrinsic uncertainty produced by a perceived discrepancy between the individual’s current level of certainty about important environmental objects and a criterion state that he seeks to achieve” (p. 206). The “environmental objects” in his definition refer to people, things, events, or ideas that possess psychological importance for the individual. In Atkin’s view, humans sense differences between what they know and what they want to know as regards a salient “thing” in their mental universe. Thus, they constantly compare current levels of knowledge against goal states that they wish to reach (presumably “perfect knowledge” of all those things that concern us), and react by seeking information whenever they sense uncertainty.

Most of Atkin’s work was aimed at understanding use of the mass media. A more sustained research project concerning uncertainty reduction evolved in the field of interpersonal communication. For over 30 years, Charles Berger (1987, 1997, 2002), Berger and Calabrese (1975), James Bradac (2001), and Berger and Bradac (1982) have investigated the role of uncertainty in interpersonal relations, demonstrating how people seek information strategically when faced with uncertainty in conversation and social situations. Dale Brashers (2001) and Brashers et al. (2002) have portrayed information seeking as one of several strategies that one may use to “manage uncertainty” in one’s environment. Uncertainty reduction continues to be fruitfully explored as a goal in face-to-face interactions.

More central to information behavior (IB) research is the work of Nicholas Belkin (1978, 2005) and his collaborators (Belkin & Robertson, 1976; Belkin, Oddy & Brooks, 1982; Belkin & Vickery, 1985), and Carol Kuhlthau (1988a, 1988b, 1991, 1993a, 1993b, 1997, 2004a, 2004b, 2005) and her co-authors (Kuhlthau & Tama, 2001; Cole & Kuhlthau, 2000) in advancing a view of information as tied to uncertainty.

In terms of its vocabulary and sources, Belkin’s writings most reflect Taylor’s concept of “visceral need.” To Belkin the basic motivator of information seeking is an “anomalous state of knowledge” (ASK). An ASK exists when a person recognizes that there is an anomaly (i.e., a gap or uncertainty) in their state of knowledge regarding a situation or topic. Faced with an ASK, individuals may attempt to address their uncertainty by requesting or consulting information. The person will then judge whether the anomaly has been resolved; if it is not resolved, another ASK may be generated, or the motivation to address it may be exhausted.

To illustrate these last two stages of ASK, consider our earlier request for philosophy books. If indeed the hypothetical inquirer is directed to shelves filled with hundreds of philosophy texts, she may very well decide to give up and ignore her anomalous state of knowledge. Alternatively, she might be handed a dictionary of philosophy that contains a discussion of the concept of “truth”;

reading what the dictionary says, she decides whether the text resolves her uncertainty, or whether she needs further explanation. Her anomaly might well change in the process of seeking: now she realizes that it is the nature of “belief” that she does not understand. Again she must decide whether to continue until her information needs are met, or be satisfied with what she knows now. In a sense, an inquirer always “gives up” eventually, because there is always more that could be known regarding a topic. The question of “when” is determined by available resources and the inquirer’s level of motivation.

Much of Carol Kuhlthau’s research is based on psychologist George Kelly’s (1963) theory of learning as a process of testing constructs. Uncertainty is a beginning stage in any search, and this is often accompanied by feelings of anxiety — which is a powerful motivator to either get on with the work, or to give up entirely. Kuhlthau’s work was pioneering in several ways, particularly in its attention to the role of affect in information behavior. Like Taylor, she was concerned with stages (or phases) of a search process; her research on students demonstrates how they reach a state of closure as regards their information needs — the final level of Taylor’s typology. Uncertainty reduction is a key component in all of Kuhlthau’s research. A recent application of her model is found in Hyldegard (2009).

4.2.3. Making Sense

While Artandi (1973) saw uncertainty reduction as one approach to defining “information,” she also saw a second framework: semiotics. Semiotics deals with the study of cultural products (such as language) as a system of signs that convey meaning by way of established conventions. So Artandi’s second approach to understanding information stressed the production of meaning, or “sense.”

The most ambitious attempt to explain the origins of information needs lies in the work of Brenda Dervin and her colleagues (see Dervin, Foreman-Wernet, & Lauterbach, 2003; Tidline, 2005) on “sense-making.” This line of inquiry has been applied widely in the context of what Savolainen (1995) calls “everyday life information seeking,” rather than traditional research on the use of factual information from libraries, television, newspapers, or other sources. Perhaps for that reason, sense-making tends to emphasize the *feelings* rather than cognitions “in situations where humans reached out for something they called information” (Dervin, 1992, p. 68).

Brenda Dervin believes that we have a need to “make sense” of the world. Dervin believes that it is safe to assume that need

implies a state that arises within a person, suggesting some kind of gap that requires filling. When applied to the word information, as in

information need, what is suggested is a gap that can be filled by something that the needing person calls “information.” (1983b, p. 156)

Dervin prefers to define information need as a compulsion to *make sense* of a current situation, when faced with problems or worries, or the need to understand or make a choice:

The individual, in her time and place, needs to make sense.... She needs to inform herself constantly. Her head is filled with questions. These questions can be seen as her “information needs.” (p. 170)

In the sense-making characterization, a search for information starts with questions directed at making sense of the situation; communication is central to the process of “bridging the gap” to reach some kind of information or help desired. The strategies employed are shaped by the searcher’s conceptualization of both the gap and the bridge, and by the answers, ideas, and resources obtained along the way. They are engaged in a search for meaning (Cornelius, 1996; Wilson, 1984). As others have noted (Kuhlthau, 1991, 2004a; Mellon, 1986), emotions are at least as important as cognitions in “gappy” situations: searchers may be intent upon reducing their anxiety as much as their uncertainty. And one need for information may easily lead to another, in a kind of chain, as Shenton and Dixon (2004b) have demonstrated in their research.

4.2.4. The Spectrum of Motivations

That people have information needs is a fundamental assumption regarding information seeking. Taylor, Belkin, Kuhlthau, and Dervin all provide frameworks to discuss a phenomenon that remains beyond our observation: the activity in human minds that leads to an individual recognizing an information need.

The work of all four scholars has been widely applied and continues to be cited. Robert Taylor’s writings have been used in the training of public service workers such as reference librarians. The work of Nicholas Belkin has been widely applied in the development of information systems, in which the emphasis is on the way that queries evolve as the answers they generate are evaluated. Carol Kuhlthau’s research is used a great deal in educational settings, at both K-12 and university levels.

Dervin’s work has been, like Taylor’s, used to understand what takes place in question-answering arenas like the library reference desk. Dervin’s “neutral questioning approach to the reference interview”



Figure 4.2: The spectrum of views regarding motivations.

(Dervin & Dewdney, 1986) advocates asking open-ended questions (e.g., “What are you trying to understand?” or “How are you planning to use this information?”) to better understand the questioner’s situation and needs. Because of sense-making’s emphasis on emotions as well as cognitions, it has also inspired a broader audience of professionals involved in human services, such as specialists in health care and in social welfare.

What scholars say about information needs can be illustrated on a continuum that reflects their assumptions about the nature of information, *why* people seek it, and what they *use* it for. We might call one end of the spectrum the *Objective* pole and the other the *Subjective* pole (Figure 4.2).

At the Objective end of the continuum are those who view information as reflecting an objective reality, and information seeking as driven primarily by a rational judgment that some uncertainty exists that would be resolved by specific information; emotional motivations of the search process, such as anxiety, tend to be set aside. The prototypical search from the Objective point of view is one in which there is a well-defined need to retrieve a specific fact to make a decision or solve a problem. From this perspective, information needs are thought to be relatively fixed. The early writings of communication scholar Charles Atkin (1972, 1973) best illustrate this view; he acknowledges “noninstrumental” information seeking motives, but he defines them as out of his scope (1973, p. 205).

In contrast, the Subjective pole represents the idealized view that many (and perhaps even the majority of) searches for information are prompted by a vague feeling of unease, a sense of having a gap in knowledge, or simply by *anxiety* about a current situation. This view does not deny that purposeful thought leads to information seeking, but rather emphasizes that humans are often driven to “make sense” of an entire situation, not merely its component “data,” and that rational goals are often overstated. Under such a view, information needs are highly dynamic. Brenda Dervin’s work is an exemplar of the Subjective view, although most IB researchers tend toward that end of the spectrum.

Of course, saying that there are two views of information needs makes an artificial distinction. The differences between the Objective and Subjective camps are not all that great; there are many examples of overlap and agreement. But proposing these stereotypes helps to explain why different approaches to studying searches and sense-making have evolved. The Objective school tends to focus on the psychological aspects of processing

information; it tends to view some types of information seeking as trivial or irrational — particularly in cases in which people do not use the most authoritative sources of information or in which they ignore seemingly relevant information. The Subjective camp holds that an understanding of the receiver as “making sense” of the world leads to more accurate picture of when and how messages are “received” — and when they are not.

In ending this section I must note two facts. First, that in labeling the work of these four authors as being about “answers,” “uncertainty,” and “gaps” I am grossly oversimplifying their respective works. My stereotypes are simply meant to make their ideas both more digestible and more comparable. The reader is urged to follow up with the “Recommended for Further Reading” suggestions for this chapter.

Second, neither do I mean to suggest that these characterizations of information needs — as answers, as uncertainty reduction, or as gaps — belong exclusively to the four authors mentioned. Rather, it is merely that these four authors are frequently cited in discussions of how information needs arise and are attended to in certain kinds of information seeking.

Connections between information acquisition and the concepts of uncertainty, ambiguity, and curiosity, in particular, were widely explored by psychologists from the 1940s through the 1960s such as Allport and Postman (1947), Berlyne (1960), Driscoll and Lanzetta (1965), and Miller, Galanter, and Pribram (1960). Many other authors in the last half-century have invoked the idea of uncertainty reduction as a prime motivator for information seeking. The notion of “making sense” also has roots in the work of sociologists like Cicourel (1964), Garfinkel (1967), and Schutz (1962, 1964, 1967).

4.3. The Trouble with Information Needs

Wilson (1981, 1997) says that, while researchers fret over a definition of information *need*, much of the time they are really studying information seeking *behaviors*. Belkin and Vickery (1985) point out that observing an information need is problematic, because it exists inside someone’s head and must be inferred by any interested observer while a search is in process, or after it has taken place:

Less tractable is the issue of why people look for information at all; that is, what is the status of the concept or category of *information need*?... [I]s there such a thing as a need for information, which can be considered on its own ... or is information-seeking behaviour contingent upon the desire to satisfy other types of needs, or to resolve situations which are not in themselves information-dependent? (p. 6)

Indeed, other scholars (e.g., Wilson, 1981; Poole, 1985) believe that the notion of an information need is an unrealistic concept, as most information needs could be said to be accounted for by more general needs, and in any event they cannot be observed. An example of their first type of objection would be that our need to know the prices of items (e.g., bread) may be driven by our need to eat (surely a more basic human need), or our need to conserve our resources (less basic, but compelling to most humans). That needs and motivations may be hierarchical in this way is also a central assumption of Terror Management Theory (Pyszczynski, Greenberg, & Solomon, 1999), discussed in Chapter 7. Others suggest that information may indeed be a basic human need, but deny that needs are hierarchical (Max-Neef, 1992).

Bosman and Renckstorf (1996) point out the circular nature of assumptions about information needs as a distinct motivator:

it is in fact an *ad hoc* notion created for practical purposes in order to predict information-seeking behaviour and information consumption. It is rather obvious that people who consume much information on a certain subject will also state that they have a certain need for this information.... However, if one wants to explain why some people do and others do not consume certain information, the information needs concept is as elucidating as, for instance, explaining criminal behaviour on the basis of hypothetical “criminality needs.” (p. 43)

In reviewing issues surrounding human motivations, Hirschman and Holbrook (1986) say that “action theorists” (e.g., Goldman, 1970; Hampshire, 1982) generally argue that wants and desires — when coupled with beliefs about the relationship between means and ends — provide *reasons* for actions. However, these theorists are divided about whether such reasons could be said to *cause* actions.

In the case of information, Bosman and Renckstorf see three overlapping motivations that determine a need for it:

social utility (e.g., in order to have topics of conversation), instrumental utility (e.g., in order to decide whether to buy something) and intrinsic utility (e.g., the entertainment value of the information offered). (p. 46)

The first (social) and last (intrinsic) of these “utilities” are not always counted as “real” information needs. Bryce Allen (1996) points this out when he says

there may be a variety of gratifications that are provided by the information-seeking process that cannot be considered meeting a

specific information need or solving a particular problem. Another way of looking at these information activities is that they meet needs (such as the need for entertainment or companionship) that are not classified as information needs. (p. 56)

Given the multidimensional nature of such needs, how are we to describe them? Harter (1992) argues that to talk about an individual’s “information need” is virtually the same as describing his or her “current psychological state,” because needs shift stochastically as each relevant piece of information is encountered. One bit of knowledge may raise questions, lead to another fact or to a new conclusion, and so forth, which changes one’s knowledge state and hence what one finds relevant and worth seeking. At least Wilson, Pool, Bosman and Renckstorf, and Harter would agree that, however information needs are characterized, they are not something fixed and long-lasting. Kari’s (2010) exploration of possible meanings of “information use” takes a similar view in that “use” may involve the construction of knowledge in response to an evolving need.

Most of what I have said (and quoted) thus far has downplayed the idea that having information is a “basic” human need. Many psychologists would disagree. George Miller (1983b), for example, described information gathering in instinctual terms. Another psychologist, Abraham Maslow (1963), said “I am convinced that man *does* have a need to know,” describing it as “instinct-like,” even though he admitted that he could not prove its existence (p. 111). Wendell Garner, who wrote extensively on the role of information in forming cognitive structures, believed that “the search for structure is inherent in behavior.... People in any situation will search for meaningful relations between the variables existing in the situation” (1962, p. 339). Milton Rokeach’s view was that “we are all motivated by the desire, which is sometimes strong and sometimes weak, to see reality as it actually is” (1960, p. 400). To “see reality,” we *need* information about it.

4.4. Information Seeking and Information Behavior

We come, at last, to information seeking. It may seem counterintuitive, but researchers have had less to say about seeking and use (Savaolainen, 2009b) than they have about needs. Perhaps the meaning of the term is thought to be obvious. Most accounts of empirical investigations do not bother to provide a definition of information seeking, taking it for granted as what people do in response to a need for information. It could be said that information seeking is more closely tied to the concept of “need” than it is to the notion of “information” itself. For instance, Tom Wilson has said that

information seeking is “the purposive seeking for information as a consequence of a need to satisfy some goal” (1999b).

The few authors who state an explicit definition of information seeking typically describe a process of either *discovering patterns* or *filling in gaps* in patterns previously recognized. Garner (1962), for example, implies that it is the search for relationships among stimuli. Likewise, Zerbinos (1990) says that

information seeking takes place when a person has knowledge stored in long term memory that precipitates an interest in related information as well as the motivation to acquire it. It can also take place when a person recognizes a gap in their knowledge that may motivate that person to acquire new information. (p. 922)

The basic notions behind what Garner and Zerbinos describe date back to John Dewey’s (1910/1933) characterizations; Dewey saw *inquiry* as motivated by recognition of a problem — of something lacking in a situation. Gary Marchionini’s definition of information seeking is problem oriented: “a process in which humans purposefully engage in order to change their state of knowledge” and which is “closely related to learning and problem solving” (1995, pp. 5–6). Also relevant is the way that Brenda Dervin defines sense-making in terms of confronting problematic situations; indeed, for some investigators information seeking has come to be synonymous with sense-making.

Johnson offers one of a few definitions that are more restrictive than those above: “Information seeking can be defined as the purposive acquisition of information from selected information carriers” (1997, p. 26). In this case there is no reference to the “purpose” itself, or to what motivates a person to select a “carrier” and acquire information from it. Krikelas (1983) describes information seeking in like terms, saying that it starts with a “need-creating event/environment” — a characterization also echoed by Julien and Michels (2004, pp. 547–548) and Westbrook (2008, p. 24).

The reader may have noticed that the definitions of information seeking quoted above emphasize *purposive activity*. There is a broader term that encompasses information seeking yet also includes behaviors that are passive: “information behavior.” Tom Wilson (1999b) defines this term as:

the totality of human behaviour in relation to sources and channels of information, including both active and passive information seeking, and information use. Thus, it includes face-to-face communication with others, as well as the passive reception of information as in, for example, watching television advertisements, without any intention to act on the information given.

In recent years the label of “information behavior” has firmly established itself as a covering term for a broader range of information-related phenomena, many of which topics are receiving fresh attention (see Chapters 10–12 for examples). It is a term whose time has come.

In conclusion, information seeking is a taken-for-granted concept, a catchall phrase that encompasses a variety of behaviors seemingly motivated by the recognition of “missing” information. Although it is the most common term in use, information seeking is typically defined strictly in terms of active and intentional behavior, which limits its applicability to the broad range of research currently being conducted on human use of information.

4.5. Summary

In this chapter I have explored the notion of information need. I have pointed out, as have many others before me, that “need” is an awkward concept, particularly in that it is not easily observable. Rather, needs are more typically inferred *post hoc*, after some action or request has been made manifest.

The notion of an information need is rooted in more basic human needs. The extent to which humans have a “need to know” is disputed, with most scholars identifying it as a secondary need that is much less important than the need for food, shelter, or companionship.

I described four oft-cited conceptions of how information needs arise: models by Taylor, Belkin, Kuhlthau, and Dervin. Each of these is similar in that they all point toward feelings of uncertainty, ambiguity, or uneasiness as the root cause of information needs. You will read more about these authors — particularly the work of Brenda Dervin — in future chapters. For now, here is a brief comparison of what they say about the origins of information needs.

Robert Taylor talks about a series of stages in which “vague sort of dissatisfaction” may (or may not) become “an ambiguous and rambling statement” that in turn may (or may not) become an articulated question. Taylor’s examples suggest that his main focus is the situations in which one person asks another person a question, as occurs at a library reference desk. Taylor’s work helps us to understand the nature of human questioning or, as Taylor calls it, “question negotiation.”

Nicholas Belkin emphasizes the notion of anomaly and the uncertainty that accompanies it. He also invokes the idea of a “state of knowledge” that is being constantly updated and compared with earlier states to judge whether or not an anomaly has been resolved. Although Belkin’s concepts are very

general, his work lends itself to the modeling of information retrieval systems. In such systems the results retrieved are evaluated and may result in additional query statements if the results are judged to be inadequate, or if the retrieved information results in a “shifting” of the focus and vocabulary of the question. Belkin’s writings are often cited in some combination with those of Taylor and Kuhlthau (e.g., see Attfield & Dowell, 2003; Bruce, 2005; Cole, Leide, Beheshti, Large & Brooks, 2005; Ford, 2004a).

Carol Kuhlthau uses psychological theories of learning to advance the idea of uncertainty as a starting point in library research. She has also emphasized the importance of emotions in information behavior. Like Taylor, she suggests that information seeking is expressed in stages.

Brenda Dervin freely acknowledges the overlap between her ideas and those of Taylor, Belkin, and Kuhlthau. Like Belkin she invokes the idea of a “gap” in life’s experience as a motivating stimulus for seeking information. However, Dervin is concerned with a broader issue than simply getting answers to questions. Her writings emphasize basic issues in human welfare, such as the goals of feeling secure and self-actualized. For Dervin, looking for “information” is only one response to a “gap”; other responses could include seeking reassurance, expressing feelings, connecting with another human being, and so forth.

Many writers on the topic of information needs suggest that it is not a basic human need, comparable to those for food, shelter, security, or companionship. Information needs are said to change constantly with new, relevant sensory inputs. In other words, new questions emerge as old ones are answered or even partially satisfied. Yet some psychologists with existential leanings see information processing as a basic aspect of being human; they are inclined to see information as a basic human need after all.

Finally, I addressed the concept of “information seeking” itself, pointing out that researchers rarely bother to define it explicitly. When it is defined at all, it is described as a reaction to the recognition of an information need — a somewhat circular definition. I make a case for use of the term “information behavior” as better suited to characterizing a broad range of relevant human behaviors dealing with information.

Information needs and information seeking are related to a host of other notions, some of them tightly coupled and others more peripheral. In the next chapter we will explore these related concepts and see where and how they fit in.

Recommended for Further Reading

Allen, B. (1996). *Information tasks: Toward a user-centered approach to information systems*. San Diego, CA: Academic Press.

Although directed at the design of information systems, Allen discusses many aspects of information seeking, including the concept of “need.” His book contains many examples of how social relations and context influence the seeking of information.

Beautyman, W., & Shenton, A. (2009). When does an academic information need stimulate a school-inspired information want? *Journal of Librarianship and Information Science*, 41, 67–80.

Beautyman and Shenton review the concept of information need and apply the ideas of Brenda Dervin to a participant observation study of primary school children. They found that children often seek information on topics related to those they studied in class, based on a variety of motives.

Bradac, J. J. (2001). Theory comparison: Uncertainty reduction, problematic integration, uncertainty management, and other curious constructs. *Journal of Communication*, 51, 456–476.

A comprehensive introduction to the ways in which uncertainty has been dealt within research on interpersonal communication.

Dervin, B. (1983). Information as a user construct: The relevance of perceived information needs to synthesis and interpretation. In S. A. Ward, & L. J. Reed (Eds.), *Knowledge structure and use: Implications for synthesis and interpretation* (pp. 153–184). Philadelphia, PA: Temple University Press.

A well-written explanation of Dervin’s earliest views on the subjective nature of information.

Green, A. (1990). What do we mean by user needs? *British Journal of Academic Librarianship*, 5, 65–78.

Green’s article is broader in scope than either its title or journal of publication would imply. His topic is human “need” and how it has been defined by political philosophers and social theorists. The concept of need is a slippery one, and even more so when we discuss the need for information, which is even more subjective than needs for physical things and states.

Taylor, R. S. (1968). Question-negotiation and information seeking in libraries. *College and Research Libraries*, 29, 178–194.

Taylor’s observations about how a reference question is negotiated between librarian and client still seem fresh after many years. In this piece he suggests how information needs arise and are satisfied.