Relevance for Browsing, Relevance for Searching

David Bodoff
Department of Information and Systems Management, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong. E-mail: dbodoff@ust.hk

The concept of relevance has received a great deal of theoretical attention. Separately, the relationship between focused search and browsing has also received extensive theoretical attention. This article aims to integrate these two literatures with a model and an empirical study that relate relevance in focused searching to relevance in browsing. Some factors affect both kinds of relevance in the same direction; others affect them in different ways. In our empirical study, we find that the latter factors dominate, so that there is actually a negative correlation between the probability of a document's relevance to a browsing user and its probability of relevance to a focused searcher.

Introduction

The concept of relevance has received a great deal of theoretical attention, as has the relationship between focused search and browsing. Nevertheless, it is our observation that the literature has not adequately compared the concept of *relevance in the browsing task* to the concept of *relevance in focused search*. This article aims to fill that gap. In the first four sections we review the literature and put forward a combined model that highlights the structural relationship between the two kinds of relevance. In subsequent sections we use the model to frame a number of research questions. We investigate the relationship between the two kinds of relevance by tracking which documents are judged relevant by the two kinds of user. A number of empirical results indicate fundamental differences in relevance criteria for browsers[1] versus searchers. In this manner, we aim to show the usefulness of the integrated model for framing research questions and guiding empirical research into the two kinds of relevance.

[1]*Browsers* means "users engaged in browsing"; it does not refer to the World Wide Web navigation software.

Received January 27, 2004; revised October 19, 2004; accepted November 8, 2004

 • Published online 2 November 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/asi.20254

Browsing Versus Searching

The browsing literature includes numerous typologies, i.e., types of browsing. However, there does not appear to be a clear consensus regarding the *definition* of browsing. This section reviews some of the literature on browsing, with the modest goal of understanding the various possible definitions. We will then be in a better position to adopt a definition that suits our current purpose.

Definitions of Browsing

We distinguish between definitions of browsing as a kind of behavior and definitions of browsing as a kind of task or information need. Bates (2002a) clearly defines browsing as a behavior that can be defined on its own mechanical terms: "[Browsing] involves successive acts of glimpsing, fixing on a target to examine visually or manually more closely, examining, then moving on to start the cycle over again." Chang and Rice (1993, p. 237) similarly identify strictly behavioral—even biological—definitions in the works of Morse and O'Connor. A study such as that by Qiu (1993), which measures whether users choose to employ browsing or analytical searching *for a given task*, clearly defines browsing as a behavior, independently of the task. In that study, browsing includes the behaviors of paging through nodes and looking through tables of contents, and analytical search relates to the behavior of submitting string searches.

In contrast, numerous definitions of browsing refer to characteristics of the user's goals or task or his or her information need. We will refer to this as the *cognitive* definition of browsing. For an example of this approach, we refer to another article by Bates (Bates, 2002b), in which *browsing* is defined as whatever one does in the course of an undirected, active search. *Active* is behavioral; *undirected* is a characteristic of the need, not of the method or behavior: "Here we have no special information need or interest,[2] but actively

[2]It should be noted that when these definitions refer to "undirected" or "unspecified," the intention is not (only) that these are not known-*item* searches, but that the searcher does not even know what specific information he or she needs.

expose ourselves to possibly novel information." Apted's similar definition is "examination of sources . . . in the hope of discovering unspecified new, but useful, information" (Apted, 1971, p. 228). This definition refers again to the nature of the need rather than to specific behaviors, i.e. browsing is whatever one does when one is looking for something helpful, but not something specific. Apted puts forward this cognitive definition, even though his own review of the word's etymology reveals a more behavioral focus. Chen and associates (1998) also define browsing in terms of the task. They compare two kinds of behavior—navigating links and keyword search—for a browsing task that is defined as "find something of interest," then study both behaviors again for a focused searching task. They treat the behavioral issue as orthogonal to the question of focus, and they opt to define browsing in terms of the latter.

We have thus far contrasted the cognitive with the behavioral definition of browsing. Of course, the leisurely scanning behavior may be related to the lack of a specific target or need, so it is not surprising that the two senses should blur. For example, Cove and Walsh (1988) give a behavioral definition—"actions of moving about"—but then write that "it is related to searching where the initial search criteria are only partly defined" (p. 31). In her berry-picking work (Bates, 1989), Bates also seems to include both aspects. First, when arguing that a typical user does both browsing and directed searching, Bates defines *browsing* as not having a focused need. But the remainder of Bates's article discusses design features, and in that context, *browsing* is used to indicate a kind of scanning behavior.

In this article, our concern is with how relevance differs in browsing versus directed search. For our purposes, therefore, *browsing* is defined with respect to the unspecificity of the information need. A behavioral definition would not be useful in this study because a user's mechanical behavior does not affect whether a document is relevant. Rather, it is the unspecificity of the information need that is hypothesized to affect the notion of relevance; that is the operative element.

Combining elements from three studies (Bates, 1986; Huber, 1991; Vandenbosch & Huff, 1997), our cognitive-based definition is the following: *Browsing* is actively looking through information (active) or keeping one's eyes open for information (passive), without a particular problem to solve or question to answer (unfocused need).

The particular type of browsing we will encounter in our empirical data is current awareness browsing. The term *current awareness* describes the state of keeping up with new developments. When a user approaches a database with the hopes of keeping current in a broadly defined area, he or she does not have in mind any particular document or topic. Moreover, the user is not even particularly interested that there should *be* any new developments. Rather, the user's goal is to be aware of any new developments that may exist.

But how can we identify which users are engaged in non-focused current awareness browsing? Because this cognitive difference is not directly visible, we will use a surrogate measure. The most popular technologies that support browsing are hypertext and directories (Chen, Houston, Sewell, & Schatz, 1998). In the special case of current awareness browsing, there exists a special case of a directory, called a *clipping service*. A clipping service gathers into one place all documents related to one area. To provide this kind of clipping service, the service provider surveys numerous document sources and identifies—either manually or automatically—all the documents that pertain to the given topic. Once these documents have been identified and colocated, the user can easily browse them. In this way, the user's current awareness is supported. The users in the browsing part of our empirical study scan a folder containing all of the newspaper and magazine articles from many sources that fall into a particular subject category, for example, tourism or foreign exchange. Marchionini (1995, p. 106) has identified[3] such a behavior with Apted's "generally purposeful" type of browsing. Because these technologies support users who have unfocused needs we will consider use of these tools (a visible behavior) as a surrogate way to identify users who have unfocused needs.

Earlier we distinguished between cognitive and behavioral definitions of browsing and specified that it is the cognitive element that motivates our theory that relevance may mean something different for browsing. However, in spite of the cognitive theoretical definition, when it comes to operationalizing our measurement (identification) of browsers, we take advantage of the fact that the cognitive and behavioral aspects are correlated in practice and that people who have less focused needs are supported by particular technologies. The definition is cognitive, but we identify browsers by using the surrogate measure of the tools they choose.

Relevance

The second main concept in this study is relevance. The literature has established two definitions of relevance, known as *systems relevance* and *user-centered relevance*. A main idea of the user view is that a user may have additional requirements, beyond the document's matching the intended topic. For example, a user may need a particularly recent document, or a particularly authoritative one, or one suitable for beginners, and so on, and these needs determine additional requirements the document must satisfy in order to be useful for—"relevant to"—that user at that time. It is interesting to note that this distinction in the two streams of relevance can also be viewed as the distinction between a behavioral and a cognitive definition (Schamber, Eisenberg, & Nilan, 1990), which would exactly parallel our distinction between two the definitions of browsing. One additional distinction can be made, between the user view of relevance and empirical research on relevance assessments. The user view of relevance assumes a rational judge who considers more than just topicality match when determining which documents satisfy his or her needs. Research on relevance

[3]Marchionini envisions a single-word ad hoc query as opposed to a one- or two-word category folder.

DOI: 10.1002/asi

assessments (e.g., Cuadra & Katter, 1967), on the other hand, does not limit itself by the assumption of a rational judge and also investigates psychological biases and other nonrational factors that affect relevance judgments. In this article, we are concerned only with rational criteria.

The systems view and the user-centered view were identified in Saracevic's review as early as 1970 (Saracevic, 1970). Boyce presented the two as working together in a two-stage process (Boyce, 1982). What began as a debate has become a consensus: Both kinds of relevance are important to users (see, for example, Froehlich, 1994). The new focus is on whether even this combination captures the whole story, or whether we should include a greater focus on whether the user was ultimately able to apply the information to achieve a better outcome (e.g., a better diagnosis or a better travel plan) (Hersh, 1994).

Categorizing Relevance Criteria to Help Contrast Browsing With Searching

How does the literature on relevance relate to the literature that distinguishes browsing from focused search? Both the systems view and the user view of relevance have mostly assumed a traditional focused search. For focused search, the literature has developed a detailed inventory of possible relevance factors that go beyond mere topicality. The apparent gap in the literature is the question, What relevance criteria still pertain to browsing users who have an unfocused need and which do not? And are there *additional* criteria that apply *only* to such users and that have not been included in the inventory of relevance criteria for focused searchers? The literature on relevance distinguishes between a narrow topical view and a more contingent user-centered view of relevance. What can we infer from this distinction regarding the meaning of relevance for browsers?

The broad outline of our answer is that topicality—the main item in the systems view—is of less concern for browsers, because they have no particular topic in mind. Rather, only user-focused factors remain.

On closer inspection, this argument is imperfect. The accumulated list of "user" factors includes many other criteria whose importance for browsers is similarly doubtful. It seems a different distinction is required.

In the following section we offer a categorization of relevance criteria that separates those that apply to searching from those that apply to browsing. As for any categorization, what makes it useful is that it sheds light on an issue. In our case, the categorization guides our thinking on the question, which (categories of) relevance criteria apply only to focused searchers, which to browsing users, and which to both?

Categorizing Relevance Criteria: Monotonic Versus Indeterminate

We want to determine which relevance criteria apply to browsers who lack a specific focus. One way to approach this question is to ask whether a given criterion affects relevance regardless of the user need. For some relevance criteria, this is the case. Some criteria are "the more the better" for all users, regardless of their information need. We call such criteria *monotonic*. An example of "the more the better" is clarity: A clearer document is always more relevant.[4] Equally, if a criterion is "the less the better" consistently for all users regardless of their need, then too we call it a monotonic criterion; this is just a trivial labeling reversal that arises because some "relevance criteria" are in fact labeled in the negative way. "Consensus within the Field" is one such criterion we will encounter; it is basically a bad thing, but it is bad for all users regardless of their particular need, so it is a monotonic criterion. Such monotonic criteria apply equally to focused searchers and unfocused browsers, because these criteria affect relevance irrespective of the user need.

But such is not the case for all criteria. For some criteria, which we will call *indeterminate*, it is impossible to say, without reference to the particular user and his or her need, whether having that criterion makes the document more relevant. A trivial example is a document's topical content. The same information content is good for a user who is looking for it and not good for a user who is not. A nontopical example are the document's depth and focus. A document may be deep and focused, or it may be a general overview, but we cannot say which depth level is better for relevance in general, without reference to the user need; it depends on what the person needs. These criteria are characterized by the fact that their value, i.e. their effect on relevance, depends on the user need. These criteria are less applicable to browsers who have no particular need. We also include in this category criteria that are *defined* in reference to the user or the user need (e.g., the property of being novel to this user).[5]

The main point can be summarized as follows: Document characteristics whose presence or absence always enhances relevance no matter what the user need will apply even to browsers. Document characteristics whose definition or whose effect on relevance can only be ascertained by appealing to the details of the user need will not straightforwardly apply to browsers who lack a particular need.

In our model, relevance depends on the following categories of criteria, which are organized to reflect our distinction between monotonic and indeterminate criteria:

1. Indeterminate document criteria whose effect on relevance depends on the *relationship between the document and the user need*. These include (a) topicality match, (b) topicality relationship (but not match), and (c) matching relationship (but not topicality).

[4]We ignore the pathological case in which for some (historical or academic?) reason, a user specifically seeks unclear documents or other faulty documents.

[5]Otherwise, one could arbitrarily redefine an indeterminate criterion as a monotonic one (e.g., the property of having *the right* topic [always a good thing] or the property of having *the right* depth for this users). And in fact Barry does define some criteria in this way. But these are all indeterminate, whether their value depends on the user or whether their definition depends on the user.

DOI: 10.1002/asi

2. Monotonic document characteristics that affect relevance the same way for all users, regardless of their exact information need. These are relevance criteria *of the document itself*.
3. Characteristics *of the user need itself*.

Category 3 will not appear in our diagrams and will rate only brief mention. This category includes only those user criteria that affect a user's *propensity* to judge *all* documents as relevant (e.g., generous people, people in a hurry). It is possible, for example, that a typical searcher is in more of a hurry than a typical browser, in which case searchers may have a lower probability of relevance overall. But we are not primarily interested in this overall propensity. Rather, we are interested in whether browsers and searchers consider as relevant the same documents or different documents. All criteria that affect this question are included in the other categories. Note that our category 3 is very different from the usual distinction between the systems view and the user view of relevance. The user view of relevance includes everything except topicality match, including document characteristics such as specificity. But our user category is really limited to user characteristics that affect all his or her relevance assessments without regard to the document, the user need, or anything else.

Figure 1a shows our main categories of relevance criteria, and Figure 1b expands this diagram backward by showing that criteria 1a–c depend on the match between user need and document.

We will refine the model later, but it is helpful to give a quick example of categories 1b and 1c. An example of topicality but not match (1b) occurs when a document and query stand in a nonmatching relationship, such as precursor–successor or cause–effect (Green, 1995b). The opposite case, match but not topicality (1c), occurs when the document matches a needed characteristic other than topicality, for instance, appropriate level of expertise.

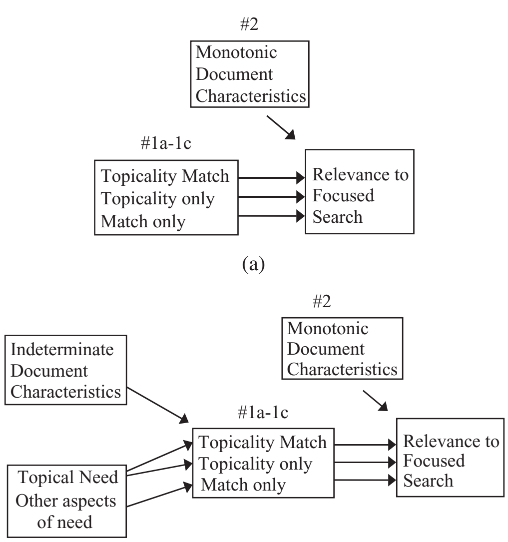

FIG. 1. (a) Relevance criteria. (b) Relevance criteria and the antecedents of match.

Populating the Categories

We "populate" these categories with specific relevance criteria from the literature, in order to be more specific about the meaning of each category. Using an inductive content analysis, Barry (Barry, 1993, 1994, 1998) coded and reported all the criteria mentioned by users as a reason for pursuing (or not pursuing) a document. The criteria are mostly defined as "the extent to which the information [is accurate, is already in the user's possession, etc.]." A number of these criteria were also identified in previous work cited by Barry, lending further validity to her findings. We build on her work in identifying factors that influence relevance, but we abandon her categorization of these criteria in favor of our own three categories 1–3. We go through the factors identified by Barry, distinguishing those characteristics whose presence—or in some cases, whose *absence*—always enhances relevance no matter what the information need is from those whose impact on relevance depends on the user's need. The first group are our monotonic criteria, and the latter group are our indeterminate criteria.

Barry introduces a similar distinction. For each criterion, she records whether users considered it as "desired present" or "desired absent." If all users considered it the same way, this went without comment. But if a single criterion was considered as both "desired present" by some users and "desired absent" by others, she concludes, "It might be argued that these quality categories were not inherently desirable or undesirable, but rather that the desired presence or absence of these qualities varied among users and situations." This parallels our distinction between monotonic and indeterminate criteria.[6,7]

[6] For many criteria in Barry's results, including "Effectiveness" and "Consensus Within the Field," *almost* all respondents agreed that the criterion is either desired present or desired absent, and a tiny number of respondents took the opposite view. Technically, all these are indeterminate because of the one contrary opinion. But our categorization does not automatically consider these as indeterminate. For example, one of Barry's users listed "[methodological] effectiveness" as "desired absent." Barry attributes some such surprising responses to the user's academic/historical interest. For our purposes it is not helpful to consider these cases because an academic/ historian might consider literally anything, including nonrelevance itself as relevant. Thus, in deciding whether to consider each criterion as monotonic or indeterminate, we apply some common sense, consider the meaning of the criterion, and consider the overall consistency in Barry's respondents' judgments.

[7] Difficulties emerge when directly adopting the definitions of Barry's criteria as the basis for a distinction between monotonic and indeterminate criteria. One problem is that some criteria are defined in positive terms, some in negative terms, and some in neutral terms. The first two results in Barry's needing to consider "desired present" and also "desired absent," and in our needing to define monotonic as including both "the more the better" and "the less the better." The neutral labels such as *Affectiveness* make even these meaningless, because it is not defined as positive affect or negative affect, but just affect. A user who expressed like or dislike would in either case be counted as considering Affectiveness. But of course it is impossible to make any sense of Affectiveness's being desired present or desired absent, because it depends whether the affect expressed was "like" or "dislike." For such reasons, the important consideration in determining whether a criterion is monotonic is not Barry's exact definition, but whether that criterion could be defined without regard to the user need, and in such a way that all users would agree that the criterion was "the more the better" or "the less the better."

DOI: 10.1002/asi

We first populate our category 2, monotonic document characteristics whose presence or absence always enhances relevance, regardless of the user need. In this list we include the following of Barry's criteria: Objective Accuracy/Validity, Clarity, Source Quality, Source Reputation/Visibility, Obtainability, Cost (these are monotonic also in Barry's results), Effectiveness, Recency, Consensus, External Verification, Availability Within the Environment.

We turn next to criteria that depend on the user–document relationship. Our category 1a is the most straightforward relationship—"The document had what I was looking for." Surprisingly, this statement was not identified as a factor of relevance in Barry's 1994 article. However, her 1998 article does introduce such a category, Information Content Only, which represents a full 62% of all responses. We also include Content Novelty and Personal Availability. These factors are closely related to the criterion of topical match (1a), except that these criteria, rather than indicating whether the topic is the needed one in principle, indicate whether the content is *still* needed, in light of other documents or other knowledge the user already has. As elaborated by Boyce (1982), this kind of novelty is logically related (secondary) to topicality, because being novel is a benefit only if the topic is right. This idea also corresponds to Bean and Green's view that many such user-centered factors help to limit the result set that is originally based on topicality (Bean & Green, 2001, p. 117). Because these factors work in conjunction with topical match, we consider them as part of our category 1a. Note that Barry and others who are concerned with contrasting the systems and user views include these factors as part of the user view.

Turning to category 1c—match but not topicality—we include the following criteria from among Barry's list: Depth/Scope, Background/Experience, Subjective Accuracy/Validity; Affectiveness, and perhaps also Tangibility. As previously discussed, this category includes factors whose effect on relevance depends on the particular user need or whose definition is stated in reference to the user need. Depth/Scope is the best example. A document may be at the right level of depth for a given need, but we would not ordinarily say that the document was written at *the correct* level of depth for all conceivable purposes. Background/Experience indicates whether the user has background in the document's topic and so is *defined* in terms of user–document relationship. Similarly, Barry's Source Novelty, Stimulus Document Novelty, Ability to Understand, and Relationship with Author are all *defined* in terms of the way the document relates to the particular user's knowledge or situation.

We have so far identified factors that fit our categories 1a, 1c, and 2. In a number of papers, Green (Bean & Green, 2001; Green, 1995b) focuses on nonmatching topicality. This completes our picture by developing our category 1b—topicality but not match. Green explains that there are many kinds of topicality relationship in addition to topicality *match*, such as hierarchical relationships, and cause–effect relationships. For example, in the query "*To whom* is God merciful?" the relevant text stands in a *recipient* relationship to the topic. Numerous other nonmatching topical relationships are reviewed. We adopt Green's analysis of the various kinds of topical nonmatching relationships as the basis for our category 1b.[8]

Finally, category 3 entails criteria that depend only on the user. One example from Barry is the criterion Time Constraints, which we would include in this category. As previously explained, we do not pursue this category further.

To summarize, we have the following criteria, organized by category:

1. The relationship between the document and the user needs (a) topicality match: Information Content, Content Novelty, and Personal Availability; (b) topicality but not match: Green's relationship types (cause–effect, etc.); and (c) match but not topicality (indeterminate characteristics): Depth/Scope, Background/Experience, Subjective Accuracy/Validity, Affectiveness, Tangibility, Source Novelty, Stimulus Document Novelty, Ability to Understand, and Relationship With Author.
2. Monotonic document characteristics: Objective Accuracy/Validity, Clarity, Source Quality, Source Reputation/Visibility, Obtainability, Cost, Effectiveness, Recency, Consensus, External Verification, Availability Within the Environment.
3. Characteristics of the user need itself: Time Constraints.

Two Categories Represent Two Competing Hypotheses of Document Popularity

The model set out in the previous section leads to two competing theories as to why some documents are relevant more frequently than others. We use the term *popularity* to refer to the frequency of a document's relevance. If we adopt a frequentist view of probability, this is also related to the document's probability of relevance. Suppose we calculate the number of times each document is identified as relevant by a user and draw the resulting histogram. The literature on Web caching shows that the distribution of document downloads—a rough surrogate for relevance—follows a Zipfian or other power law (Breslau, Cao, Fan, Phillips, & Shenker, 1999). But such statistics do not explain what makes some documents more popular than others, i.e., what is behind the shape of that distribution. Our model of relevance categories 1–3 gives the following two competing hypotheses for what

[8]Green's examples require a clarification. If (say) we had a novel about downtrodden people, and a query about God's mercy, then according to religious belief, the subject of the book (i.e., downtrodden people) would stand in a recipient relationship to the subject of the query (i.e., God's mercy). And Green would be correct that typical search engines cannot readily consider this kind of nonmatching topicality because they cannot know that the subject of downtrodden people may have some relationship with the subject of God's mercy. But if the query asks, "To whom does God show mercy?" and the book says, "God shows mercy to the downtrodden," we consider this as a topical match, and something that a search engine could find, but Green apparently still considers this as a nonmatching relationship.

DOI: 10.1002/asi

makes a document popular, i.e., what makes it relevant more often than other documents:

1. The "repeating need" hypothesis, corresponding to category 1
2. The "good document" hypothesis, corresponding to category 2

The repeating need hypothesis holds that a *document is popular* when it matches a *popular query or information need*. In other words, the document is relevant often, because it matches an information need that arises frequently. On the other hand, the good document hypothesis says that a document is popular because of the document's internal qualities (e.g., it is thorough, accurate, readable, by an authoritative author or publisher); in short, it is a high-quality document. These two "hypotheses" summarize nicely the basic structure of the model, and it will be useful in the section Relevance in Searching, Relevance in Browsing. (There is no parallel hypothesis for category 3, because these factors do not affect the relative popularity of different documents across all users).

Relevance in Searching, Relevance in Browsing

In the previous section we developed a categorization of relevance criteria on the basis of literature that largely assumes focused search. But the categorization was specifically designed with the intention of allowing discussion and comparison of relevance for focused search versus relevance for browsing. This is the focus of this section.

All the monotonic document characteristics of category 2 are expected to play a role in determining relevance for browsing, just as they did for searching. The definition of these characteristics is that their effect on relevance does not depend on the precise user need, so these factors should similarly affect relevance for browsers. On the other hand, the main idea of our approach is that relevance criteria in category 1—criteria pertaining to the *relationship* between document and user need—may affect browsers very differently than they affect searchers. The reason is twofold: First, because the notion of matching user needs does not straightforwardly apply when the user has no focused need, some of these criteria may not even apply. Second, to the extent these criteria do still apply, browsers may simply have different needs than focused searchers. Either case would cause a different set of documents to be popular to searchers and to browsers.

We set the stage by first introducing an extreme position, according to which all the indeterminate criteria simply do not apply to browsing (see Figure 2). We will not ultimately adopt this extreme position, but it serves to highlight our fundamental approach, according to which *the commonality between searching and browsing relevance is due to the monotonic criteria, whereas the differences between them are due to the dubious and different effects of the indeterminate criteria on browsers.*

We have so far argued that monotonic document characteristics will play a role in browsing relevance comparable to

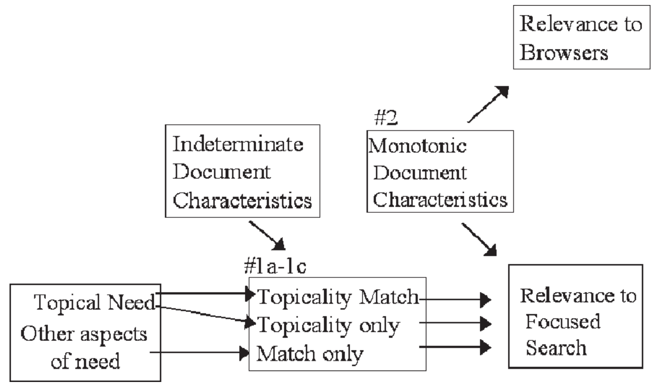

FIG. 2. Initial model of relevance factors for browsing versus searching.

their role in searching relevance but that many indeterminate characteristics related to the document–query relationship will not straightforwardly apply. We will fully explore this idea more fully later. The next question is the opposite, whether there are any *new* relevance criteria, which were *not* found in the context of focused search, that apply uniquely or especially to browsers.

Some such "new" criteria for browsing may actually be modified forms of traditional criteria for searching. For example, Content Novelty and Personal Availability, as originally presented by Barry in the context of focused search, indicate whether the content was *still* needed, a characteristic that is inextricably related to whether it was *ever* needed (i.e., whether it was on the right topic). But in a modified form, we can imagine a criterion of Pure Novelty, which means only that the content was new to the user, with no condition that it was also on the "right" topic. This criterion could certainly apply to many browsers, depending on their goal.

The literature on browsing suggests an additional category of criteria that are part of browsing relevance but not searching relevance: the category of document characteristics related to enjoyment. Browsing in the context of consumer behavior and in our context of information search are sometimes compared. In the case of consumer behavior, browsing "may be pleasurable in itself" (Chang & Rice 1993), and it may also lead to an unplanned purchase. The same applies when browsing for information, except that with information, there is less of a dichotomy between pleasure in itself and browsing that leads to an unplanned purchase, because for information products, the enjoyable perusal is itself the consumption. In any case, the comparison with consumer behavior would support an intuition that relevance for browsers, unlike for focused searchers, may consider document characteristics that make the process more enjoyable "in itself." Barry also includes an enjoyment measure called *Affectiveness* and applies it to focused search. But this measure differs in one sense from "enjoyment in itself." In the context of focused search, emotional affinity is not a rational criterion, unless (1) the document is already relevant in some *other* sense, and (2) the reader's emotional affinity will make the relevant information easier to absorb, and so forth. This is different from enjoyment in itself

DOI: 10.1002/asi

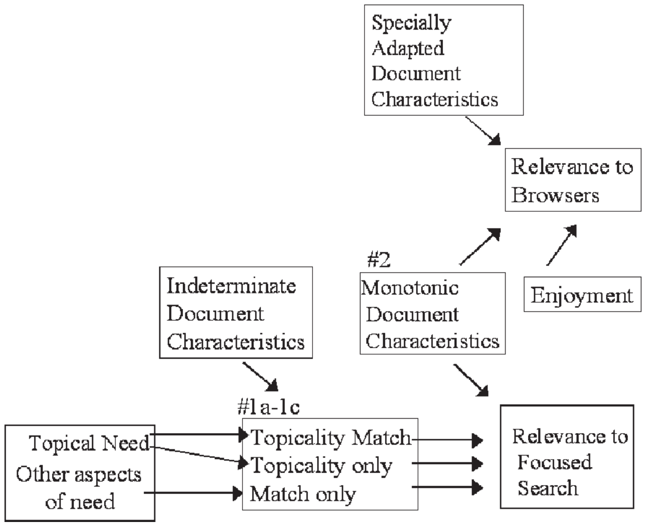

FIG. 3. Partial model of relevance factors for browsing versus searching.

which, in the case of browsing, can make a document relevant even in the absence of any other basis.

In constructing a model of relevance for browsing comparable to Figure 1 for focused search, we have so far included the traditional monotonic document characteristics, characteristics such as Pure Novelty that are adaptations from criteria found for focused search, and a new category of document factors that we call *enjoyability*. Figure 3 presents a summary.

We now return to the question of the document–query relationship, i.e. indeterminate factors from our category 1. For simplicity, up to this point, we have considered that because browsers do not have a particular information need (recall that we *defined* browsers in this way), speaking of a relationship between the document and the information need is difficult, so that category 1a does not straightforwardly apply, and categories 1b–c are questionable. But this is an oversimplification, for a number of reasons. First, even regarding topical match (1a), browsers may *discover* their latent needs through browsing. Browsing users, such as those engaged in scanning behavior, "consult specified sources regularly because they probably contain items of interest" (Chang & Rice, 1993, p. 235). The user did not set out to locate a document on this topic, but he or she recognizes it as being it "of interest" in the sense that is activates a latent topical need.

A second reason regards nontopical needs. When the literature speaks of a browser as lacking a focused need, it ordinarily refers to a lack of focus in the most traditional aspect of relevance, i.e., topicality. As the concept of relevance has expanded to include relevance criteria beyond topicality match, the same intuition of the browsing literature—that not everyone is so narrowly focused—may apply to other relevance criteria as well as to topicality match. So, should we expect that browsers have less focused nontopical needs? Should we simply suppose that for browsing, none of the indeterminate criteria applies?

Ultimately, we do not adopt this extreme position. Reviewing the list of criteria in category 1c—match but not topicality—we find that a number of criteria regard such basic needs, that they should apply even to browsing. For example, the Ability to Understand is so basic that it should certainly apply, even though it is about the "match" with the user. Even regarding less basic criteria such as Depth/Scope and Tangibility, we do not hypothesize that they do not apply to browsers, i.e., that browsers have no preferences about these.

Rather, we propose that regarding indeterminate criteria, browsers may have *different* (latent) needs than focused searchers or may typically be in a different situation. In other words, matching criteria may apply in principle to unfocused browsers, but because browsers have different needs or a different situation, these criteria may result in browsers' choosing different documents than focused searchers. For example, a typical browser may have different (latent) needs regarding depth, or may have different background than a typical searcher, or may have different ability to understand compared with a typical focused searcher, and so on.

In summary, there are two ways in which the indeterminate criteria cause differences between relevance for browsing and relevance for searching. First, there are some differences in whether the relevance criteria apply. For example, enjoyment in itself pertains more to browsing than to focused searching; topical match applies to browsers only if we introduce an extra stage in which latent needs are recognized; and some nontopical needs may be less focused for browsers. Second, even when the same matching criterion applies in principle, browsers' *actual needs* may differ from searchers' actual needs; as a result different documents may be relevant to the two groups.

Figure 4 shows the structure of our model of browsing relevance and incorporates it into the previously presented model of search relevance. The model includes monotonic criteria (box B4) that affect both kinds of relevance similarly; it includes criteria such as Pure Novelty (box B6) that apply differently to browsing; criteria such as Enjoyment (box B5) that are new to browsing; and, in light of the preceding discussion,

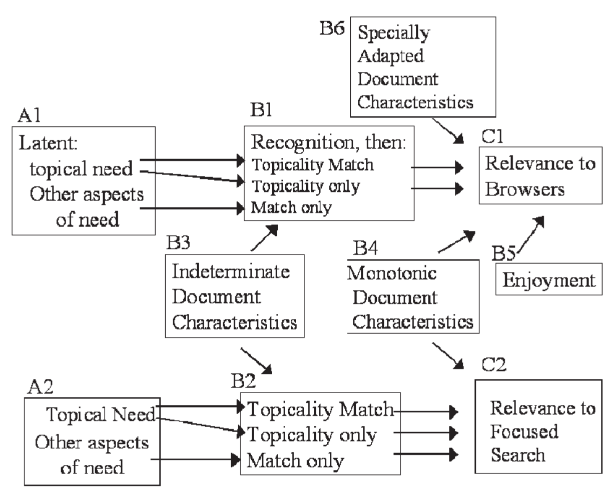

FIG. 4. Integrated model of browsing and search relevance.

DOI: 10.1002/asi

the final model for browsing relevance includes the indeterminate criteria (box B1), except that the details are different for browsing and for searching. The first difference is that because the user does not set out looking for anything in particular, any notion of "matching" his or her needs depends first on a preliminary "latent need recognition" stage (box B1). The second difference is that the user needs (box A1) may differ from those of focused searchers (A2). Integrating the two kinds of relevance into one diagram helps to guide our follow-up research questions in subsequent sections.

With browsing as with focused search, two competing hypotheses can explain why some documents are more popular than others. According to the good document hypothesis, some documents are higher in quality or more intrinsically enjoyable. This hypothesis corresponds to category 2 ("document characteristics") of relevance criteria for browsing, which appear as boxes B4 and B5 in Figure 4. Alternatively, according to the repeating need hypothesis, some documents are popular because it happens that they frequently match users' latent needs (e.g., they are about a topic that piques people's curiosity or matches their other latent needs when they see it). This hypothesis corresponds to category 1a–c (document–query relationship) of relevance criteria as they apply to browsing, and it appears as box A1 in Figure 4.

Research Questions Raised by the Model

The model not only is an integrated conceptual review, but also helps to frame research questions. In the following sections, we present two research questions that are raised by the model, and the empirical data analysis that is also guided by the model.

Research Question 1: Are the Same Documents Relevant to Both?

The first question that follows from the model's structure asks whether the same documents will be popular—i.e., frequently relevant—in both cases. The answer hinges on the following:

1. The good document hypothesis, when applied to both searching and browsing, states that both kinds of relevance depend on monotonic document characteristics. In Figure 4, this approach would model that C1 and C2 depend on B4–B5. Factors B4, ought to affect both kinds of relevance in the same direction. Thus, to the extent that intrinsic document characteristics determine relevance, then the two kinds of relevance will tend to be positively correlated. Also, to the extent that this effect dominates, there will be added significance to the fact, recognized and bemoaned in Green (1995a) and elsewhere, that search algorithms tend to focus exclusively on topical match and to ignore these factors completely.
2. The repeating need hypothesis, when applied to both browsing and searching, models that both kinds of relevance depend on the ways the document's topic and other features relate to the realized or latent information need. In terms of Figure 4, this hypothesis maintains that C2, C2 depend on A1, A2, B3 through B1, B2. According to this view, the only structural difference between the two modes is that for focused search, a popular document is one that relates to a need that often arises consciously and instigates frequent focused searches, whereas for browsing, a document is frequently relevant if it relates to a latent user need that is frequently recognized after the fact. Recall now the research question, Are the same documents relevant to both browsing and searching? According to this "hypothesis" the answer hinges on the empirical question, To what extent are recognized needs identical to latent needs? Take, for example, the topicality aspect of the need. Imagine that we record the relative frequencies of topics that are consciously sought in focused searches and do the same for the topics that are recognized (after the fact) as needed or "of interest" to browsers. Would these relative frequencies be the same? Are the conscious and expressed focused needs representative of users' latent needs? Or is it the case that some topics are easier to recognize consciously, and other topics tend not to be recognized consciously until the information is observed? The same question applies also to nontopical aspects of the need. This is an empirical question that is raised by the model. To the extent the user need determines relevance—the repeating need hypothesis—then the relationship between search relevance and browse relevance depends on the connection between searchers' needs and browsers' needs. If both have similar needs except that one is latent, then we will expect the same documents to be relevant to both. Otherwise, we might not find a positive correlation; in an extreme case, we might even see a negative correlation, at least in this respect. The repeating need hypothesis thus suggests the need for empirical investigation into whether browsers' and searchers' needs tend to be similar or different.

Note one important difference between the two hypotheses: The good document hypothesis could only lead to positive correlation between relevance in the two uses. On the other hand, the repeating need hypothesis may lead to either positive or negative correlations because browsers and searchers may have different needs.

Research Question 2: Can Browsing Relevance Predict Focused Search Relevance?

A second research question regards information retrieval effectiveness. The question is, Is it possible to predict a document's relevance to a focused search from its relevance to browsers? If the first research question asks about the overall correlation—whether the same documents are popular in both modes—this question more specifically asks whether we can predict one from the other. The question of prediction differs from the question of correlation for a number of reasons, and these reasons also explain why we framed the question as predicting search relevance from browse relevance rather than vice versa. First, for focused

DOI: 10.1002/asi

search, we usually have the user's query statement, so the more practical and more interesting question is whether the document's relevance to browsers can help predict its relevance to a searcher, *above and beyond* the predictions we usually make on the basis of the query statement alone. The unconditional correlation cannot tell us that. The second reason is related to the question of timing. In the newspaper situation we will study, we learn about a document's popularity to browsers soon after the document is published. The document's popularity to focused searchers trickles in more gradually over time, as the document is relevant to archival searches. Practically speaking, then, searcher relevance data will not available in time to help predict relevance to browsers, but browsing relevance data may be available in time to help predict relevance to searchers. Grand total historical correlations can tell us whether the same documents are ultimately correlated, but not whether the browsing data arrive in time to help in the practical prediction of relevance for the archival searchers.

Empirical Study

In exploring these questions, we adopt a structural approach to the data. Rather than measuring the individual constructs of the model, we address the research questions by taking advantage of the structural relationship between the two kinds of relevance. This empirical work may be fairly characterized as exploratory, in that we had no theoretical basis to predict whether in fact the same or different documents would be relevant to the two uses (first research question) or whether browsing relevance statistics could help to predict search relevance even given the query statement (second research question).

Our point of view in the data analysis is based on documents. That is, in most cases a record of data will be related to a document, as opposed to a user or a query. For example, in the section Answer to Research Question 1, we will measure and compare two outcomes for each document: its relevance to browsers and its relevance to searchers.

Data

The data are taken from Wisers Information Ltd., a Hong Kong company that provides http-based subscriptions to most major Chinese- and English-language daily newspapers in Hong Kong, Peoples Republic of China, and Taiwan (Republic of China), as well as other information services. Every document is full-text indexed and is accessible to subscribers via English- or Chinese-language keyword search. In addition, every document is categorized into zero, one, or more standard or customized "folders" (e.g., Finance, Environment). This structure allows Wisers to provide a clipping service for current awareness browsing. Wisers Information records in a database every time a user views, saves to the local disk, or e-mails (to himself or herself or to a friend) an article. For each of these events, the Wisers database distinguishes whether the document was originally accessed via a keyword query (and if so, which keywords) or via a browsing folder. As an example, suppose a user submits a keyword search "x y z", gets a list of retrieved documents, and chooses to view one of these documents, D99. Then the database records that this user viewed document D99 as a result of query "x y z". On the other hand, if the user was browsing the Finance folder and chose to view this document, then the database records that document D99 was viewed by this user as a result of his or her browsing the Finance folder. Wisers provided the researchers with 3 months of such log data. The data recorded every document access event, for all of August–October 1999. The data also recorded which "user" did the action, but these users were actually corporate logins rather than individuals.

The restriction of data to August–October 1999 requires clarification. There are two kinds of dates in this application: (1) the dates on which documents were created and (2) the dates on which a user judged the documents as relevant. We limited our study to documents that were both created and judged as relevant during the 3-month period. The generalizability of the analysis is therefore limited to relevance patterns for relatively new documents.

Separating Browsers From Focused Searchers

The degree of focus is not an observable construct. Instead, we use the technical tool used for document access, whether via a folder or via a keyword search, as a surrogate to distinguish those users who had a focused search from those who did not. As discussed in Definitions of Browsing, the literature associates use of clipping services with current awareness browsing, and the use of keyword search with the more focused searches. This correlation allows us to use an observable behavior as a surrogate measure to identify which users probably lack a focused need.

Though this surrogate has been used in other studies of browsing, we believe it is an especially reliable surrogate here. Wisers's management explained to us that all subscribers during the covered period were corporate customers, and that the majority were media and financial companies. It was the estimation of management on the basis of knowledge of their customers that professionals in these fields perused pertinent folders each morning, as part of current awareness. The media professionals who covered a certain "beat" (e.g., technology, Hong Kong politics) wanted to know what stories were being covered in other printed media than their own. The finance people routinely skim the finance-related folders as a way of covering many daily newspapers for important stories before the trading day begins. On the basis of this information from Wisers's management, we were satisfied that the primary use of the folders, as access points to the newspaper stories, was to meet unfocused information needs. Apart from access via these folders, all documents were also available via full-text search. We treat the use of keyword search as a surrogate for the users' having a more focused need.

The documents in this study are all periodicals. The searchers are those searching periodicals, and the browsers

DOI: 10.1002/asi

are those primarily engaged in maintaining current awareness by using periodicals and using the technology of clipping folders. Regarding external generalizability, one never knows whether browsers who are in another context or are using a different technology might act differently. This lack of knowledge limits our ability to generalize our results. Nevertheless, the browsers in our study are representative of all that typifies browsing. We *know* they are using browsing technologies and behaviors, and we have reason to believe that they have unfocused information needs. In addition, the theoretical model of relevance applies to others who have unfocused needs. We therefore have no reason to suspect that our empirical results will not apply as well to others who have unfocused needs.

Surrogates for Relevance

In this study we investigate models of relevance, and our experimental data are a record of page views, page saves/downloads, and page e-mailings. A user views a page after seeing it listed in a results set or in a browsing folder. The listing includes the headline/title, as well as the newspaper or other source, date, and author (byline). A user's saving or e-mailing the document after viewing it is arguably a strong (and unobtrusive) surrogate for document relevance. On the other hand, merely viewing the page is not as good a surrogate, and the data logs do not distinguish these three kinds of events. In personal discussions with the researcher, users explained that the headline is normally sufficient for an accurate appraisal of the document's relevance, especially because they are conducting current awareness in an area of their expertise; the deep domain expertise makes it fairly easy for them to ascertain the document's relevance from the title. Also, such data have been used as a surrogate for relevance in field studies (Cooper & Chen, 2001). Moreover, in the present context, we believe the question can be dismissed on principle. For testing our model, a measure of *expected relevance* is as good as a measure of *actual relevance*. If one is testing the effectiveness of a document retrieval algorithm, then one wants to know which documents were actually relevant, not only which appeared relevant. But we are comparing which documents are relevant to browsers and which to users. For this purpose, data that indicate which documents appeared relevant to users suffice to tell us whether users have similar relevance *criteria* in the two cases. All that changes after a user sees a particular document is that the user then knows more about the particular document. Thus, our research questions do not depend on the difference between a document's actual and anticipated relevance.

Answer to Research Question 1: Detailed Data Analysis and Results

Our analysis of this research question is presented in parts: the main result, ruling out a possible confound, searching for explanations in topical categories, and searching for possible explanations in other indeterminate document characteristics.

Main Result for Research Question 1. We took the master list of all 18,551 documents that were categorized by Wisers into at least one folder during the period August 1999 through October 1999, and so were accessible via both full-text search and topical folders. For these 18,551 documents, we recorded the total number of events (views, saves, e-mailing) that occurred under each type of use, browse or search. Each such number represents the frequency of relevance of that document under that use. A total of 18,551 pairs of numbers resulted. Table 1 shows a sample of these event counts for a few documents. The first statistic for our purposes is the correlation between these two.

A Pearson's correlation coefficient gave a result of −.13. But inspection of the data showed an uncanny inverse relationship that seems to surpass this magnitude. The sample in Table 1 is typical and seems to show a much stronger inverse correlation. More thorough data analysis, which was encouraged by reviewers, revealed the following:

A large number of unpopular documents were exclusively relevant only to one or the other group (a strong negative correlation), and a smaller number of highly popular documents were popular to both browsers and searchers (a positive correlation).

Tables 2a–b show what happens when we split the documents into two groups, low values of browsing relevance

TABLE 1. Sample of data for correlation study.

Docid	Browse	Search
199908030040031	0	1
199908030040032	1	0
199908030040040	0	1
199908030040059	3	0

TABLE 2(a). Majority of cases, doc relevant to 0 or 1 browsers.

Correlations			
		Browse	Search
Browse	Pearson correlation	1	−.585**
	Sig. (2-tailed)		.000
	N	15236	15236

**Correlation is significant at the 0.01 level (2-tailed).

TABLE 2(b). Minority of cases, doc relevant to 2–17 browsers.

Correlations			
		Browse	Search
Browse	Pearson correlation	1	.302**
	Sig. (2-tailed)		.000
	N	3315	3315

**Correlation is significant at the 0.01 level (2-tailed).

DOI: 10.1002/asi

(0 or 1 users only) versus higher values. For each group separately, we calculate the Pearson correlation coefficient between browsing relevance and searching relevance.[9] We can also split the data in other ways (e.g., choosing documents with low numbers of search relevance instead of low numbers of browsing relevance or defining *low* as less than 3 or 4). Any way we sliced the data, the following structure emerges: The lower the values of search or browsing relevance, the more we find a strong negative relationship.[10]

In a case such as this in which the relationship is not linear, a linear parametric correlation will always give an averaged and therefore understated measure of the relationship. In our case the averaged result was −.13. The dilution of the negative relationship was unusually strong in this case because the smaller number of positively related documents have higher values, and these exert particular influence on the correlation, much as regression outliers with strong "influence" do.

Because we know the relationship is not linear—in fact it changes direction, not only slope—use of a parametric test of linear correlation is not appropriate. Instead, either nonlinear models or nonparametric tests should be used. We do not pursue nonlinear models because we lack any sound theoretical basis to hypothesize such a relationship between browsing and searching. Nonparametric tests of association include Spearman rank coefficient based on ranks and Kendall's tau coefficient based on concordance. Because our variables' relationship reverses from negative to positive, even these nonparametric tests will report an averaged and therefore understated effect. But because these non-parametric tests work with ranks or concordance instead of frequencies, they are not *unduly* affected by the high values of the popular documents. The overall Spearman's rho using all the data was −.5 and Kendall's tau was −.46, both highly statistically significant. These are strong negative correlations between the two kinds of relevance and more accurately represent the inverse relationship that is obvious when inspecting the data, which largely resemble the sample in Table 1.

An alternative, which allows us to continue using parametric tests, is to split the data, either explicitly or implicitly. Table 2 shows regression results after explicitly splitting the data. Within each of the two partitions, parametric tests are reasonably appropriate; within each subset, regression residual plots look normal. A related approach is a piecewise regression, which is an implicit way of splitting the data. A piecewise regression is a "single" regression that allows the model to change the intercept and slope at a particular point(s) in the predictor's values. We ran such a piecewise regression to predict search relevance from browse relevance, allowing the model to shift the line when browsing popularity is greater than or equal to 2. The results of this regression are presented in Table 3. There are two sets of intercept and predictor variables, one set for the lower values of the predictor and one set for the higher. The first set consists of "Constant" and "Browse," and the second set is "Browse_ge2_0" (an intercept) and "Browse_ge2_1." As expected, the slope is negative (−.649) for the lower values and and positive (.999) for the higher values of the predictor. There are numerous outliers, and these present a threat to the meaningfulness of the model, but the residuals appear normally distributed.

TABLE 3. Piecewise regression results.

Model summary[b]

Model	R	R square	Adjusted R square	Std. error of the estimate
1	.421[a]	.177	.177	.916

Coefficients[a]

	Unstandardized coefficients		Standardized coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	1.168	.011		103.876	.000
browse	−.649	.011	−.646	−58.982	.000
browse_ge2_0	1.582	.039	.389	40.207	.000
browse_ge2_1	.999	.023	.369	42.686	.000

[a]Predictors: (Constant), browse_ge2_1, browse_ge2_0, browse.
[b]Dependent Variable: search.

Because we did not hypothesize it, the idea that the relationship changes direction must be considered as an exploratory result and needs to be confirmed in other data sets before it is generalized. Although it is exploratory, it is nevertheless robust within this data set. A single split in the data results in a strong negative correlation for 15,000 + data points, and any similar split yields similar results. For this reason, our subsequent data analysis will present these split-data results in addition to the more conservative nonparametric results over the whole data set.

Our primary aims are to present an integrated model of browsing and search relevance and to show the model's usefulness in exploring the relationship between the two kinds of relevance. The model helped us to frame the following basic question: Are the same or different documents relevant to the two kinds of user? Our investigation has resulted in an unexpected split structure. Using nonparametric methods on all the data, we find a strong negative correlation. Using parametric methods on the two subsets separately, we find a strong negative correlation for the vast majority of documents. The remainder of the article will pursue this negative correlation.

Excluding a possible confound. To confirm the negative correlation further, we undertook one additional step to rule out a possible confounding factor, which relates to the fact that the two access methods are potential substitutes. If a *given individual* had already viewed/saved a document while in either a searching or a browsing mode, then that same

[9]When splitting the data, we do not apply this criterion to both kinds of relevance, only to one or the other.
[10]For more complicated models and data sets, the methods of threshold regression can select cutoff points.

DOI: 10.1002/asi

individual would not ordinarily need to view/save that same document again when operating in the other mode. So, at the level of a given individual, this substitution effect could cause a negative correlation. At the aggregate level we would still expect to see a positive correlation if good document qualities are fundamentally appealing or if the same topics tend to be sought in both modes. Moreover, the "users" identified in the database are actually corporate identifiers, each of which represents dozens of individual users; the log does not identify individual users. Thus, the substitution effect is far-fetched, as it would require that individuals at a company would no longer need a good document via searching, if anyone in that company had already downloaded it via browsing (or vice versa). Nevertheless, in order to verify the robustness of this negative correlation absolutely, we eliminated this substitution effect. We did this as follows: In the original calculation, as in Figure 3, we had a single pair of numbers for each document corresponding to the total aggregate frequency of relevance under the two modes. In the modified method shown in Table 4, we begin with a separate record for each interuser pair. For example, for document 1, we total the number of times the document was relevant to user 1 when user 1 was browsing, and the total number of times the document was relevant to user 2 when user 2 was focused searching (see Figure 4, as an example of a single document with 3 users). We proceed in this fashion to consider every nonidentical pair of users. We omit identical user pairs, e.g., number of times the given document was relevant to (saved by, etc.) user 1 under both modes, because any negative correlation in such a case could be an artifact of the substitution effect. After tallying these records for each nonidentical user pair, we averaged all the records for that given document. We proceeded in this way for each document, until we once again had 18,551 pairs of numbers, this time representing interuser browsing and focused searching relevance frequencies for each document. The correlation results were almost identical, with a difference only in the fourth decimal place.

Searching for explanation in topical categories. In the context of our model for the two kinds of relevance, the negative correlation of relevance can only be reasonably

TABLE 4. Sample of data for interuser correlation study.

Document	As search user	Search events on document	As browser	Browse events on document
D1	User 1	0	User 2	2
D1	User 1	0	User 3	1
D1	User 2	3	User 1	4
D1	User 2	3	User 3	1
D1	User 3	2	User 1	4
D1	User 3	2	User 2	2
Avg #events interuser intermode for D1		10/6		14/6

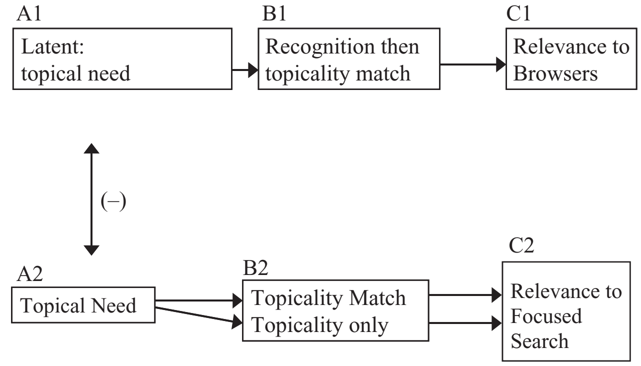

FIG. 5. Portion of complete model highlighting role of different topical needs for browsing versus searching.

explained by the strength of the repeating need hypothesis, combined with the empirically discovered fact that browsers and searchers tend to have different needs. Referring to Figure 4, from the negative correlation of C1 to C2, we infer the negative relationship between A1 and A2 in this data set.

The question is, What aspect of the users' needs is different for the two groups? One possibility is that the *topics* sought by focused searchers are different from those recognized as desirable by browsers when they see them. Figure 5 shows this possibility as the highlighted subset of the complete model. In the figure, the minus sign indicates a negative relationship between the topical needs of browsers and searchers, and this negative relationship eventually flows down to causing a negative relationship between the two kinds of relevance. The idea being proposed is that different documents are relevant to the different modes because the topics sought by focused searchers differ from those recognized as interesting when they see them by browsers. This is an empirical question. How can it be answered?

Every document in our sample was assigned by Wisers to one or more topical categories, so we can use these labels to identify the topics of the documents that were relevant to the two kinds of user. Table 5 presents the number of browse and search relevance events, broken down by topical category. Many documents were assigned to more than one category. In such cases, if the document was accessed via a topical browsing folder (e.g., Air Transportation), then the browsing event was recorded for that category. If the document was accessed via the results of a keyword search, then one search event was recorded for each of the document's categories.

A chi-square test can answer the question, Is there a relationship between topic and browse/search? That is, is the distribution of relevance events across topics different for the two columns, browsers and searchers? The test was highly significant, meaning that the latent topical needs of a typical browser are different from the conscious topical needs of a typical searcher. This initial result means that hypothetically this difference might explain the negative correlation between the two kinds of relevance. Could it be that

DOI: 10.1002/asi

TABLE 5. Two-way frequency counts by categories and search/browse.

Category	Browse	Search
Air_Transportation	57	346
Banking	420	639
Catering	191	167
China_Economics	240	247
Editorial	958	475
Education	5048	2244
Electronics	66	342
Environment	707	465
Foreign_Exchange	471	137
Forum	850	1807
Fund	203	244
Hong_Kong_Economics	441	543
Hong_Kong_Stock	2431	1361
Information_Technology	1237	1122
Insurance	240	102
International_Economics	186	239
Law_and_Judiciary	441	764
Mass_Media	318	648
Medical_Service	425	141
Metal_Industry	5	16
Plastic_Industry	13	30
Politics	980	2179
Property	773	1633
Public_Transportation	332	565
Publishing_and_Printing	58	700
Sea_Transportation	13	22
Telecommunications	1081	592
Textile_Industry	23	99
Tourism	167	193
Toys	47	29
Water_Electricity_LP	50	253

focused searchers tend to choose documents on one set of topics, and browsers tend to pick documents on other topics, and that this explains the negative correlation?

To pursue this possibility, we use partial correlations. We look again at the (negative) correlation between the two kinds of relevance, but this time we control for the category of each document. This method will tell us whether there is still a negative correlation even within each topical category or whether the different topical preferences explain the negative correlation.

Table 6 shows a sample of the data as they were set up for study of partial correlation. This sample should be compared with Table 1. We have the same relevance data, but now we also add each document's topical category, for which we will

TABLE 6. Sample data for correlation analysis controlling for topic.

Docid	Category	Search	Browse
199908030040031	Politics	1	0
199908030040032	Politics	0	1
199908030040040	Law_and_Judiciary	1	0
199908030040059	Medical_Service	0	3
199908030040059	Hong_Kong_Economics	0	3

control.[11] We ran a partial correlation to see whether the two kinds of relevance were different even within each category. We continue our practice of using a nonparametric measure of association based on ranks, because of violations of the linearity and other conditions in the data. We use Kendall's partial tau-b. If we denote the usual Kendall's tau-b between two variables X and Y as τ_{XY}, then the partial of X and Y given Z is:

$$\tau_{XY.Z} = \frac{\tau_{XY} - \tau_{XZ}\tau_{YZ}}{\sqrt{1 - \tau_{XZ}^2}\,\sqrt{1 - \tau_{YZ}^2}}$$

(Gibbons, 1993). In our case, we want to compute $\tau_{\mathrm{BS.C}}$, which denotes the correlation between browsing (B) and searching (S) relevance, given the topical category (C).

The result was a statistically highly significant ($\tau_{BS.C} = -.46$). In fact, this result is identical to the unconditional correlation reported earlier, therefore, the topical category explained nothing about this relationship. *Thus, users are selecting different documents even within a given topical category, depending on whether they were browsing or searching.*[12]

Our analysis so far tells us that (1) browsing and searching relevance are inversely correlated, (2) browsers and searchers are indeed looking for different topics, but (3) even within each topic, searchers and browsers still prefer different documents.

Searching for Explanations in Other Indeterminate Document Characteristics

The analysis of previous sections gives the impression that focused searchers are looking for something different in their documents than browsers, and that this difference is not only topical. Browsers apparently have (latent) nontopical needs that are opposite in some sense from searchers' nontopical needs. This relationship is depicted in Figure 6, which differs from Figure 5 only in that the highlighted operative relationship between A1–B1 and A2–B2 is the nontopical match. That is, we now investigate/hypothesize that the negative correlation between C1 and C2 flows from a difference in the *nontopical* needs of A1 and A2.

[11]A superficial complication arises because many documents are relevant to more than one category. In the example of Table 6, document 199908030040059 was categorized in both the Education and Medical Service folders, so there are now two records for that document. The fact that this document was relevant three times to browsers and no times to searchers is recorded once as a data point for the Education category, and also as a data point for the Medical Service category. This is not a problem of double-counting because we are not counting across categories in this analysis but are instead going to *control* for category. Nevertheless we ran the same analysis choosing only one category when a document was cataloged into more than one and of course also found that controlling for category did not alter the negative correlation.

[12]We will see additional corroborating evidence of this in the section Answer to Research Question 2, in which we control for individual queries, not just for topic category.

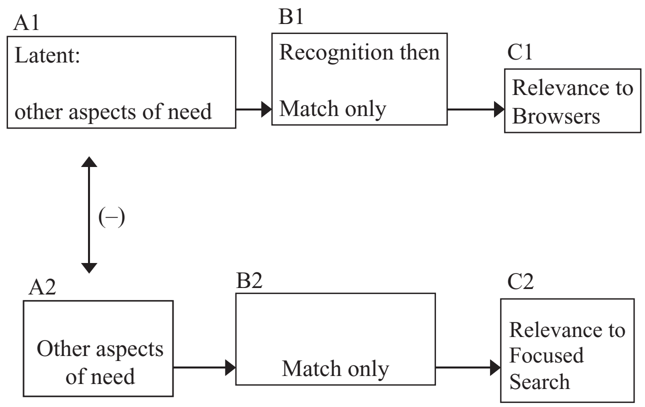

FIG. 6. Portion of complete model highlighting role of different nontopical needs for browsing versus searching.

Our analysis to this point has been based on correlational analysis, taking advantage of the structure of the integrated model. We have not yet sought to operationalize and measure any of the individual criteria that might be desired in browsing but not in searching, or vice versa.

We next probed the possibility of such an analysis, which would measure individual characteristics that distinguish the search documents (i.e., documents relevant to searchers) from the browse documents. As a preliminary test of feasibility, a research assistant was asked to review manually a sample of documents that tended to be popular to browsers and searchers and to report any impressions about characteristics that appear to separate the two groups. The research assistant reported an impression that documents relevant to browsers "are just describing a phenomenon. Let you know what is going on. For example, 20,000 students failed the entrance exam this year. [Whereas] the search documents analyse a phenomenon. . . . For example, some think the entrance exam is unfair to girls."

It was decided to pursue this avenue. On the basis of our own manual analysis and Barry's discussion of the Depth characteristic, we identified the following set of related facets: broad/specific, incomplete/thorough, summary/detailed, report/ analyze, cursory/in depth, diffuse/focused. From among these, the research assistant perceived that the facets that seemed to separate browse documents from search documents were report/analyze and summary/detailed, as in the comment quoted. The research assistant reviewed a random sample of 250 documents and graded each one on two 5-point scales representing the report/analyze continuum and the summary/detailed continuum. At the time of these assessments the research assistant did not know whether each document was relevant to browsers and/or searchers.

This analysis is to test whether searchers and browsers desire different levels of analysis and detail and whether this difference can help explain the negative relationship between browse and search relevance. After scoring each of the 250 documents on its levels of analysis and detail, we compared the unconditional (negative) correlation between browsing and searching relevance against the partial correlation, *after controlling for level of analysis*. We denote browsing relevance and searching relevance as *BRW* and *SRC*, and levels of analysis and detail as *ALY* and *DET*. For each of the 250 documents we have a value for *BRW*, *SRC*, *ALY*, and *DET*.

Using nonparametric Kendall's tau on this sample, the unconditional correlation $corr(BRW, SRC) = -.50$, and the conditional $corr(BRW, SRC|ALY) = -.47$. This means that level of analysis (*ALY*) explained a small part of the variance.

To test the significance of the mediating effect of *ALY*, we ran a regression for this sample, first regressing *SRC* only on *BRW*, then on *BRW* and *ALY* together. To be conservative and appropriate to the nonlinear relationship between *BRW* and *SRC*, the regression is piecewise for the BRW predictor.[13] As shown in Tables 7a–b, results show that *ALY*

TABLE 7(a). Regression results of SRC on BRW only.

Model	*R*	*R* square	Adjusted *R* square	Std. error of the estimate
1	.422(a)	.195	.185	.712

Coefficients

	Unstandardized coefficients		Standardized coefficients	*t*	Sig.
	B	Std. error	Beta		
(Constant)	.982	.087		11.354	.000
browse	−.606	.080	−.654	−7.572	.000
browse_ge2_0	1.205	.282	.354	4.266	.000
browse_ge2_1	.869	.244	.252	3.556	.000

Dependent variable: search.

TABLE 7(b). Regression results of BRW on BRW and ALY.

Model	*R*	*R* square	Adjusted *R* square	Std. error of the estimate
1	.501(a)	.251	.239	.688

Coefficients

	Unstandardized coefficients		Standardized coefficients	*t*	Sig.
	B	Std. error	Beta		
(Constant)	.559	.130		4.288	.000
browse	−.521	.080	−.562	−6.512	.000
analyze	.184	.044	.244	4.232	.000
browse_ge2_0	1.092	.274	.321	3.980	.000
browse_ge2_1	.765	.238	.222	3.217	.001

[13]Because the linearity assumption is violated, the predictive effect of *BRW* would be understated in a regular (nonpiecewise) regression, and the variance remaining to be explained by *ALY* would be artificially high. This would give an unfair "advantage" to the effect of *ALY*. In this case, where we are testing effect of *ALY*, using a piecewise predictor for *BRW* is a much more conservative approach, as well as the more appropriate choice given the nonlinearity between *BRW* and *SRC*.

DOI: 10.1002/asi

had a significant and positive coefficient, and a real effect on the *R*-square. These results indicate that *ALY* is responsible for some of the negative correlation between browsing and search relevance in this sample.

Discussion. After presenting an integrated conceptual model in Figure 4, we have aimed to demonstrate its usefulness in guiding empirical inquiry. The model shows the relationship between relevance criteria for searching and those for browsing. It proposes that some factors called monotonic document characteristics (category 2) will tend to affect both kinds of relevance in the same direction, and other indeterminate factors such as "match but not topicality" (category 1c) may affect the two kinds of relevance in different directions. Overall, we found a negative correlation between browsing relevance and searching relevance. Returning to the model to interpret the possible reasons for this, we see that this could only be caused by indeterminate factors—the repeating needs hypothesis—and then only if browsers and searchers empirically desire different characteristics in their documents. In terms of Figure 4, the negative correlation is caused by differing patterns of need between items A1, A2 in the model.

Items A1, A2 suggest the difference might be one of differing topical needs, or differing nontopical needs. We investigated both possibilities. Topical differences could not explain the negative correlation. But the one *nontopical* need we explored, Depth level, was a source of explanation. Any number of other, specific hypotheses may become the focus of further empirical testing and theoretical development guided by this integrated model.

We are left without a full explanation for all the negative correlation. One possible explanation returns to the question of topicality. It may be that within the topical categories, which we were able to control for, are more subtle topical differences. A much deeper study would be required for this; such a result would have important practical implications for search engines, clipping services, and other information providers.

A research assistant has induced a second and intriguing explanation from the data. Many keyword searches used proper nouns, especially names of famous people, places, or companies. It may be that typical archival searches are less topical in nature and typical browsing is more topical. In other words, it is not that focused searchers are looking for different topics than browsers, but that searchers tend to be more interested than browsers in people, places, and companies. This possibility is the subject of ongoing work.

An alternative regards other nontopical needs such as Background/Experience, Source Novelty, Stimulus Document Novelty, and Affectiveness. A typical browser may have different background knowledge than a typical focused searcher, and so the documents that match a typical browser's background, or that are novel to him or her, or that tickle his or her fancy may be different from the documents that match the focused searcher's background or that are novel to him or her or that tickle his or her fancy. If these are the causes, there is less immediate practical application, because it would be difficult to quantify these characteristics for each document automatically without human assessments. In any case, we hope we have demonstrated the usefulness of the conceptual model in first framing the research question and then differentiating alternative explanations that could be disambiguated with further research based on the model.

Answer to Research Question 2: Detailed Data Analysis and Results

For decades, researchers have bemoaned the fact that our search engines only consider topicality, specifically topicality match. The practical barrier has always been, How can an automatic process do otherwise? How can we measure the user's nontopical needs and the document's nontopical features? The structural nature of our approach circumvents these problems of direct measurement. In this section we apply our structural approach to consider the possibility that a document's browsing relevance may act as a surrogate measure of all those difficult to measure nontopical characteristics. Specifically, we investigate whether browsing popularity can be used as a surrogate measure for unspecified nontopical characteristics, which may be desirable or undesirable to focused searchers. Rather than attempting to measure all those nontopical needs directly, this structural approach allows a system to consider nontopical criteria in an indirect way.

The practical question is whether it is possible to predict a document's relevance to a focused search, from its relevance to browsers. In light of the negative correlation investigated in the previous section, the question is actually whether we can predict a document's relevance to a searcher from its *nonrelevance* to browsers. More specifically, because in the case of focused search we possess the user's query, the question is whether a document's relevance to browsers can predict its nonrelevance to a focused searcher, *beyond the information already found in the user query*.

To test this, the data are arranged in records that relate to document–query pairs. For every document and query, we calculated the following:

Independent variables:

MAT is the top 100 rank score of this document with respect to this query (100 is the top ranked document, 99 is the second highest, etc.).

PRVSRC is the number of times that that document had been relevant to focused searchers on dates before this query date.

PRVBRW is the number of times that that document had been relevant to browsers on dates before this query date.

Dependent variable:

SRC_{binary} is the binary value of whether the document was relevant to the user who submitted that query on that date.

DOI: 10.1002/asi

QUERY	QDATE	DOCID	MAT	PRVSRC	PRVBRW	SRC_{binary}
一條龍	1999-10-12	199908130030061	85	0	3	0
一條龍	1999-10-20	199908130030061	85	0	4	1
一校一社工	1999-10-30	199910200330035	78	3	2	1
一校一社工	1999-10-30	199910290290005	73	0	2	0
一國兩制	1999-10-19	199909140010101	35	0	3	0
一國兩制	1999-10-15	199909140040209	80	1	0	0
一國兩制	1999-10-19	199909140040209	80	1	0	0
一國兩制	1999-10-15	199909230330076	79	0	0	0
一國兩制	1999-10-19	199909230330076	79	0	0	1

FIG. 7. Data input for regression.

Referring to Figure 7, the regression only uses SRC_{binary} (dependent variable), *MAT*, *PRVBRW*, and *PRVSRC* (independent variables), but the first three columns *DOCID*, *QUERY*, and *QDATE* are included to help explain the meaning of each record. Each record regards a particular query submitted on a particular date, and a particular document whose search engine matching score with the query we know, together with information about the document's historical popularity.

In terms of a regression equation, we hypothesize

$$Logit\ (SRC_{binary}) = b_0 + b_1\ PRVBRW + b_2\ MAT$$

In the previous section, *SRC* and *BRW* denoted the total number of times the document was ever relevant to searchers or browsers during an entire 3-month period. In contrast, in this section a record in the regression represents a single document–query pair at a particular moment in time. Therefore, SRC_{binary} is a binary relevance judgment. *PRVBRW*, which stands for *Previous Browsing*, still represents a nonbinary total of the browsing relevance events for that document but is limited to the number of times the document was relevant to a browser *previous to the date* of this particular query, because we are trying to find whether browsing data can be useful at the moment of query processing. As a control variable, we also included *PRVSRC*, which is the total number of times the document was previously relevant to focused searchers before the current search.

Should we expect that *PRVBRW* can (negatively) predict SRC_{binary} after controlling for the document–query topical match? If the reason for the negative correlation between search and browse relevance is only that the two groups like different topics, then after controlling for *MAT*, the topical match, *PRVBRW* should have no effect. But in the section Answer to Research Question 1 we observed that even after controlling for topical category, there was still a negative correlation, which would lead us to expect some predictive power remaining in *PRVBRW*. In this section, we take this one step further by controlling for the search engine match (*MAT*) with the individual query statement, not just for the document's topical category. This allows us to focus on the very practical question of whether *PRVBRW* can help predict SRC_{binary}, above and beyond the search engine scores. In the previous section we further observed that browsers prefer documents with less depth of analysis. In this section, all factors besides topical match are implicitly captured in regression coefficient b_1.

The theoretical questions to be addressed empirically are these: Are the nontopical needs of focused searchers and of browsers important in predicting relevance? If so, then there may be useful information in historical relevance data even after controlling for the topicality match. If browsers and searchers have similar nontopical needs, then historical browsing relevance should positively predict relevance to searching. If not—and we already have a firm belief that the two types have very different nontopical needs—then browsing relevance should negatively predict relevance to searching.

Getting the data in this form involved elaborate data transformations, programming, and data cleaning, and a few data quality issues remained. We began with 5,980 queries that resulted in a view/save/e-mail for at least one of the 18,552 documents within our window. A *query* is a unique pair (query string, date). We wanted to get the *MAT* score for each document for those queries. We did this by rerunning the logged queries through the Wisers search engine in batch mode (and limiting the documents to those that existed on the date the query was originally run). The top ranked document had a score of 100, then 99, and so forth. For technical reasons we were able to do this only for the 3,980 Chinese-language queries and not for the 1,508 English-language queries. Also, recall that all our data are limited to documents that were created and accessed during a 3-month period, so for each query we were left with the intersection between the search engine's top 100 and those within our time window. We ultimately had 1,955 queries for which the search engine gave a positive score for at least one document in our window. On average, for each query approximately 5.2 documents from our window had a positive score, so the final table had about 10,000 records. Table 8 shows results of our first run.

TABLE 8. Regression results, with total number of cases 10096 (unweighted), and dependent variable SRC.

Variable	*B*	S.E.	Wald	*df*	Sig.	*R*	Exp(B)
MAT	.0102	.0010	97.9184	1	.0000	.1077	1.0102
PRVSRC	−.1265	.0288	19.2876	1	.0000	−.0457	.8812
PRVBRW	−.3410	.0326	109.6408	1	.0000	−.1141	.7111

DOI: 10.1002/asi

TABLE 9. Regression results considering also effect of document age, with dependent variable SRC.

Variable	*B*	S.E.	Wald	*df*	Sig.	*R*	Exp(B)
MAT	.0102	.0010	96.3385	1	.0000	.1068	1.0102
PRVSRC	−.0624	.0283	4.8672	1	.0274	−.0186	.9395
PRVBRW	−.3276	.0323	103.0362	1	.0000	−.1105	.7207
DOC_AGE	−.0193	.0015	171.9475	1	.0000	−.1434	.9809

The statistical significance of *PRVBRW* seems to confirm everything we found in the previous section and indicates the possibility that a document's previous browsing popularity may contribute (negative) predictive power even beyond the query statement. But one problem immediately presents itself and calls into the question this interpretation: The coefficient on *PRVSRC* is also negative and significant. This result was unexpected, as it indicates that a document that previously had more numerous relevance events to focused searchers is less likely to be relevant again now, given the new query. It occurred to us that this could be an artifact, with *PRVSRC* (and also *PRVBRW*) serving as a surrogate for age. That is, documents that have accumulated more relevance events are documents that have simply been around longer. In the case of news, older documents are predicted to have a lower probability of relevance. So, we ran an additional regression (Table 9), which also included the document's age, measured as the number of days between the date of the document's creation and the date of the query.

The coefficient of *PRVSRC* is still negative, but the magnitude of its coefficient and *R* statistic declined markedly, with only minute changes in the coefficients of *PRVBRW* and *MAT*. The introduction of *AGE* did not completely eliminate the source of this surprising result related to *PRVSRC*. But considering the combined evidence of the previous section and this section, as well as the effect of introducing DOC_AGE on *PRVSRC* but not on *PRVBRW*, we believe there is strong evidence for the robustness of the main result regarding the (negative) predictive power of *PRVBRW* beyond the search engine's *MAT* score of topicality match. *On the practical question*, the effect suggests that considering this indirect evidence may improve retrieval results that consider only topical match. *On the theoretical question*, it appears that browsing relevance indicates the presence of unspecified nontopical features that tend to be undesirable for focused searchers.

Discussion

In this section, we have addressed our second major research question. The practical question is, Can a document's popularity to browsers help to predict (non)relevance to a focused search, given the user query? The data show that it can. By controlling for the specific user query and for the time of its submission, we extend the result of the section Answer to Research Question 1, in which we accounted for only the document's general topic. The combination of these results indicates that the negative correlation between browsing and searching relevance is caused by nontopical aspects of the user need.

Further Discussion and Summary

The data analyses used in the two preceding sections differ in their treatment of time. In the section Answer to Research Question 1 we analyzed the total relevance counts for the whole 3-month period. In contrast, each record in the regression in the section Answer to Research Question 2 corresponds to a relevance judgment at a certain moment in time, and variable *PRVBRW* (and *PRVSRC*) considers the relevance history of the document up to that point in time, as well as the matching score with the query, in an effort to predict relevance. The latter approach suggests the possibility of a time series analysis of this data. In Bodoff and Zhang (2000) we indeed approached the data with more of a time series view. Instead of considering the historical relationship between browsing and searching, that article focused on the pattern of a document's popularity over time, analyzed separately for browsing and for searching. For browsing, if a document was popular when it was new, then it was more likely also to be relevant to browsers many weeks later. In contrast, regarding focused searches, the best indicator of whether a document will be relevant to searchers long after its publication is not whether it was popular when it was new, but whether it was relevant to other late searchers. The idea seems to be that for browsing, initial response is a good indicator of later interest, but for focused search, *late* use demonstrates the document's longevity as an archival source of information for still later searchers. This result adds another consistent element to the picture we are forming about differences between the two kinds of relevance.

We began this article with a review of two literatures, on browsing and on relevance. We related the two literatures by comparing and integrating the model of relevance for focused search with a model of relevance for browsing. It is hoped that the model itself will contribute to theoretical understanding.

We then reported results of an empirical analysis of the two kinds of relevance data. All our analyses indicate that there is a qualitative difference between the needs of browsers and those of searchers that extends beyond the user's degree of focus. In the section Answer to Research Question 1 we saw that there is an overall negative correlation between the two kinds of relevance in this data set. This negative correlation is not explained by the (true) fact that users in the two modes have different topical foci, but it may be explained by other criteria such as depth level. In the section Answer to Research Question 2 we saw preliminary evidence that relevance to browsers can predict non-relevance to searchers, even when we account for the query statement. Browsing relevance may thus be indicative of unspecified characteristics that match browsers' needs but not searchers'. These particular results are limited to the particular case of periodical data, but it is hoped that the model and the insights gained from the data analysis will further our understanding of relevance to browsers versus searchers.

Acknowledgments

The author would like to thank Tang Ya for help with the Chinese documents and Susanna Ho Shuk Ying for help with SQL programming.

References

Apted, S.M. (1971). General purposive browsing. Library Association Record, 73(12), 228–230.

Barry, C.L. (1993). The identification of user criteria of relevance and document characteristics: Beyond the topical approach to information retrieval. Syracuse, NY: Syracuse University.

Barry, C.L. (1994). User-defined relevance criteria: An exploratory study. Journal of the American Society for Information Science, 45(3), 149–159.

Barry, C.L. (1998). Document representations and clues to document relevance. Journal of the American Society for Information Science, 49(14), 1293–1303.

Bates, M.J. (1986). An exploratory paradigm for online information retrieval. Intelligent Information Systems for the Information Society. B.C. Brookes, Elsevier Science.

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

Bates, M.J. (2002a). Speculations on browsing, directed searching, and linking in relation to the Bradford distribution. In Emerging frameworks and methods: Proceedings of the Fourth International Conference on Conceptions of Library and Information Science (CoLIS 4) (pp. 137–150). Greenwood Village, CO: Libraries Unlimited.

Bates, M.J. (2002b). Toward an integrated model of information seeking and searching. Paper presented at Fourth International Conference on Information Needs, Seeking and Use in Different Contexts. Lisbon, Portugal.

Bean, C.A., & Green, R. (2001). Relevance relationships. Relationships in the organization of knowledge. Boston, Kluwer.

Bodoff, D., & Zhang, L. (2000). Patterns of document access in browsing and searching. In M.H. Chung (Ed.), Proceedings of the Americas Conference on Information Systems (AMCIS 2000) (pp. 58–64). Long Beach, CA: Association for Information Systems.

Boyce, B. (1982). Beyond topicality: A two stage view of relevance and the retrieval process. Information Processing and Management, 18(3), 105–109.

Breslau, L., Cao, P., Fan, L., Phillips, G., & Shenker, S. (1999). Web caching and Zipf-like distributions: Evidence and implications. IEEE INFOCOM '99, The Conference on Computer Communications, 18th Annual Joint Conference of the IEEE Computer and Communications Societies, New York.

Chang, S.-J., & Rice, R.E. (1993). Browsing: A multidimensional framework. In M.E. Williams (Ed.), Annual Review of Information Science and Technology (Vol. 28). Medford, NJ: Learned Information.

Chen, H., Houston, A.L., Sewell, R.R., & Schatz, B.R. (1998). Internet browsing and searching: User evaluations of category map and concept space techniques. Journal of the American Society for Information Science, 49(7), 582–603.

Cooper, M.D., & Chen, H.-M. (2001). Predicting the relevance of a library catalog search. Journal of the American Society for Information Science and Technology, 52(10), 813–827.

Cove, J.F., & Walsh, B.C. (1988). Online text retrieval via browsing. Information Processing and Management, 24(1), 31–37.

Cuadra, C.A., & Katter, R.V. (1968). Opening the black box of relevance. Journal of Documentation, 23(4), 291–303.

Froehlich, T.J. (1994). Relevance reconsidered—towards an agenda for the 21st century: Introduction to Special Topic Issue on Relevance Research. Journal of the American Society for Information Science, 45(3), 124–134.

Gibbons, J.D. (1993). Nonparametric measures of association. Newbury Park, CA: Sage.

Green, R. (1995a). Topical relevance relationships: 1. Why topic matching fails. Journal of the American Society for Information Science, 46(9), 646–653.

Green, R., & Bean, C.A. (1995b). Topical relevance relationships: 2. An exploratory study and preliminary typology. Journal of the American Society for Information Science, 46(9), 654–662.

Hersh, W. (1994). Relevance and retrieval evaluation: Perspectives from medicine. Journal of the American Society for Information Science, 45(3), 201–206.

Huber, G.P. (1991). Organizational learning: The contributing processes and the literatures. Organization Science, Special issue: Papers in honor of (and by) James G. March, 2(1), 88–115.

Marchionini, G. (1995). Information seeking in electronic markets. Cambridge: Cambridge University Press.

Qiu, L. (1993). Analytical searching vs. browsing in hypertext information retrieval systems. Canadian Journal of Information and Library Science, 18(4), 1–13.

Saracevic, T. (1970). The concept of "relevance" in information science: A historical review. In T. Saracevic (Ed.), Introduction to information science (pp. 111–151). New York: R.R. Bowker.

Schamber, L., Eisenberg, M.B., & Nilan, M.S. (1990). A Re-examination of relevance: Toward a dynamic, situational definition. Information Processing and Management, 26(6), 755–776.

Vandenbosch, B., & Huff, S.L. (1997). Searching and scanning: How executives obtain information from executive information systems. MIS Quarterly, 21(1), 81–107.