This source provides the most current access to titles in the Sessional Papers, since the official Sessional indexes are not published until the end of each session.

NON-PARLIAMENTARY PAPERS

These papers are produced outside the House as reports of Royal Commissions or government departments and are not formally presented to Parliament. The number and variety of non-parliamentary papers have increased with implementation of the Treasury Circular of 1921, a money-saving effort following the paper shortage and monetary crisis of World War I. The trend has continued; papers that used to be considered Parliamentary are now published as non-parliamentary papers and are, therefore, no longer included in the Sessional Sets. Unfortunately, non-parliamentary papers are difficult to identify. Some are indexed by the Ford lists and breviate, which include some non-parliamentary papers, but they are listed only in the Catalogue of Government Publications. Once identified, they can be very difficult to obtain.

CONCLUSION

The Parliamentary Papers have a long and eventful history and one that is particularly meaningful to librarians. It is a story of public demand for information about government, the rise of the free press, and the innovations of a few men, like Abbot and Hansard, deeply committed to organizing and publishing information.

The Papers can be a challenging resource to utilize. Awareness and identification of a library's holdings of the Papers is just the beginning. Learning to access the collections of specific documents precisely and to locate them is task practice and perseverance. Fortunately, indexing of the Papers, like indexing everywhere, is becoming computerized. An index to current years of the House of Commons Sessional Papers, House of Lords Sessional Papers, Hansard's Debates, and Journal of the House of Commons is now available on CD-ROM. Additionally, Chadwyck-Healy provides subscribers to its microfiche edition of the House of Commons Parliamentary Papers with a CD-ROM index, which is derived from POLIS, the Parliamentary On-Line Information System, generated by the House of Commons Library. The computerized indexes, however, only cover the more current British Parliamentary Papers and are expensive. Most researchers, therefore, will continue to rely on the traditional tools listed throughout this guide.

Integrating the British Parliamentary Papers into the reference repertoire is challenging, but the benefit to students and researchers is enormous. Each access point to the Papers is a window, with its own angle and view, into a treasurehouse of documents, and each document is a truly exciting window into the past.

REFERENCES AND NOTES

1. Lloyd George, Speech to the House of Commons, 9 May 1918, Parliamentary Debates (Commons), 5th ser., vol. 105 (1910), col. 2358.
3. The General Indexes were compiled by a series of officers of the House and are commonly known by the name of the compiler. They are Cunningham's Index (1545-1659), Flaxman's Index (1660-1697), Forster's Index (1697-1714), Moore's Index (1714-1774), Dunn's Index (1774-1800), Rickman's Index (1801-20), and Vardon's Index (1820-1852). From 1852 to 1886, the indexes were produced by the House of Commons Librarians, and from 1886 to 1879, by the Journal Office Clerks. The indexes vary in arrangement and quality, and may not be available in American libraries. Standard decennial indexes began to be published in 1880.
4. It is difficult to assign a beginning date for the Sessional Papers of either House, since many early papers have not been collected.
5. See note 4 above.

Models of User Satisfaction:
Understanding False Positives

What does it mean to say that users are “satisfied” with online searches or other products of information retrieval systems? A review of research in library and information science, computer systems, marketing, and psychology reveals three models of user satisfaction. In the Material Satisfaction Model, product performance (e.g., recall and precision) determines whether a user’s stated question is answered (material satisfaction). In the Emotional Satisfaction Model—Simple Path, users are “happy” or emotionally satisfied when their questions have been answered. In the Emotional Satisfaction Model—Multiple Path, users’ happiness depends on the questions answered (material satisfaction) but also on factors such as setting and expectations. This last model allows an understanding of the phenomenon of “false positive” emotional satisfaction: users who are happy with bad searches. The choice of model has implications for user services and for research into user-system interaction.

Rachel Applegate

Rachel Applegate is Reference/Online Librarian at the College of St. Scholastica in Duluth, Minnesota. Submitted for review May 14, 1992; revised and accepted for publication February 16, 1993.

A major goal of consumer education is the elimination of “false positives” of consumer satisfaction. A false positive occurs when a consumer is satisfied with an inferior product. Through efforts such as articles in Consumer Reports, consumers are taught what to expect from various products and, therefore, when they should be satisfied (true positives) and when they should not be satisfied (true negatives). This review examines the phenomenon of satisfaction with the products of information retrieval systems in order to outline a theoretical framework within which the false positive of unwarranted user satisfaction can be addressed.

In most previous research, the term “satisfaction” has not been clearly defined, nor have varieties of it been considered separately. Therefore, this examination begins by making a distinction between material and emotional satisfaction. Next, this article suggests that satisfaction research in library science, computer systems, marketing, and psychology can best be understood as following three models: the Material Satisfaction Model (MSM), the Emotional Satisfaction Model—Simple Path (ESM-SP), and the Emotional Satisfaction Model—Multiple Path (ESM-MP). These models are not named as such in the literature; rather, researchers use and define “satisfaction” in a way that reveals their reliance on one of these underlying ideas. The results of research carried out under the assumptions of the first two models
suggest that they are inadequate. The third model explains much more of the variance in satisfaction—what it is and what causes it—allowing a better understanding of false positives.

INTRODUCTION: MATERIAL AND EMOTIONAL SATISFACTION

When librarians began to conduct research on and with information storage and retrieval systems, the first and most important factor on which they concentrated was performance, measured in terms of how close a system came to producing a product that matched a user’s needs: relevance and pertinence, precision and recall.1

In this initial examination of system performance, when the word “satisfaction” was used it usually meant what is termed here “material satisfaction.” In an information retrieval system, material satisfaction refers to a match between what the user requested and what the user received. This material satisfaction was determined by system performance. User satisfaction was not measured. Rather, users were described as “satisfied” when they responded positively to some material measure, such as, “were the search results useful?”

In the late 1970s and into the 1980s, library researchers began using “satisfaction” with another and more direct meaning: users were explicitly asked whether they considered themselves satisfied. This internal, personal, and affective response is in contrast to material satisfaction: comparing giving “emotional support” to a political cause (“We’re with you all the way!”) and giving “material support” (“Here is a donation.”). How do users feel about the products received, or with the fulfillment of their information needs?

There are two major reasons why this emotional satisfaction is worth attention for its own sake: evaluation and behavior. An evaluative model employs emotional satisfaction as a criterion to judge information retrieval systems.2 There are problems with this. A user’s emotional satisfaction may be determined by something other than material satisfaction. Therefore, an emotional satisfaction question (such as, “Are you pleased with the results?”) may not answer the evaluator’s material satisfaction interest in whether the system has delivered a product that matches the stated need. However, emotional information can be valuable. Emotional satisfaction may be at least a partial indicator of material satisfaction. Also, a system that provides material but not emotional satisfaction may be capable of improvement in order to deliver both.

Behavior is the second reason to study emotional satisfaction, based on the premise that people who are satisfied behave differently than those who are not satisfied. This is an explicit assumption of marketing research3 that has been debated in library research (with, for example, Tagliacozzo,4 Hitchinham,5 and Tissier, Crouch, and Atherton6 arguing in support, but D’Elia and Walsh7 and Hill and Johnson8 raising questions).

Satisfaction can affect behavior positively or negatively. Emotionally satisfied patrons, for example, may use the system more in the future. Users who are not satisfied may either complain to the librarian or try to improve the system on their own. Both of these are positive outcomes, in the sense that users are benefiting the most from the system and from the services of the librarian. Consider possible negative situations: users who are not emotionally satisfied avoiding the system in the future, or users who are satisfied not realizing that they have not actually achieved material satisfaction. In this context, dissatisfaction leading to complaints is a positive outcome, as librarians do not always know when users have not had their needs met unless users are able and motivated to alert them.

Most library research on satisfaction states that the majority of people who use information retrieval systems are “satisfied.” Because of this marked positive skew, the problem of negative satisfaction is minor. There are relatively few self-identified unsatisfied users. Moreover, when such an unsatisfied patron complains, the librarian has a chance either to correct the situation (poor material satisfaction, a “true negative”) or to explain (good material satisfaction, a “false negative”). The problem lies far more with the reported abundance of satisfaction: should all these users really be satisfied? “True positives”—users with material and emotional satisfaction—are the goal. There is significant evidence that the “false positive” phenomenon (users with emotional but not material satisfaction) is a real problem.

This false positive, false satisfaction, may lead to negative behavior: a lack of complaining, or a lack of demand for material satisfaction once emotional satisfaction has been achieved. This is the phenomenon Plutchak called “the satisfied and indifferent user,” and Quint warned against when she described some users of a system as going away more ignorant of their subjects than when they were when they arrived.10 To eliminate this mismatch between emotional and material satisfaction, a better understanding of the two is needed. What determines emotional satisfaction? What are the factors that influence it, and what are the relative strengths of these factors? For example, if both material satisfaction and expectations affect satisfaction, which is more important? What else affects satisfaction?

The models discussed below, with their variables and problems of measurement, have been derived from research literature from several different fields: library science and computer systems, marketing, and psychology.

Library science is most important in terms of material satisfaction: measuring how users’ information needs have been met. Computer systems research is similar to both marketing and library science research. It focuses on how people interact with computer systems such as electronic mail and management information systems.

Marketing research is business-oriented and has an intense interest in both inducing customers to be repeat purchasers and preventing them from complaining to other customers. Marketing research uses the most elaborate models of the components of satisfaction, although the measures employed are relatively simplistic.

Research in psychology on client satisfaction appeared at about the same time as in library science in the 1970s and the 1980s. Before that time, psychologists seemed to have the same opinion as the people who studied information retrieval systems: we therapists, we researchers know when you the clients should be satisfied; we know when the therapy is right for you, we know when you should be pleased. In the last twenty years, however, psychologists and other mental health and general health professionals have begun to look directly at client satisfaction for three specific reasons: to increase client happiness, to enhance treatment compliance, and to prompt appropriate future behavior. Unlike most marketing or library situations, in psychology repeat business is not a goal in itself, but if clients do have problems, counselors do not want dissatisfaction with previous therapy to prevent them from seeking help. The psychological literature also makes a very important contribution to the study of emotional satisfaction with measurement techniques, which are more sophisticated and better examined than in other fields.

MATERIAL SATISFACTION MODEL

Research conducted according to the assumptions of the Material Satisfaction Model (MSM, see figure 1) postulates that system features determine system performance, which determines material satisfaction. Emotional satisfaction is either ignored, explicitly equated with material satisfaction, or argued to be unimportant. Consequently, user behavior is considered to be determined by material satisfaction (for a discussion of behavioral findings, see Emotional Satisfaction Model—Simple Path).

Researchers identify and measure system features, system performance second, and material satisfaction third. System features include depth of indexing, amount and type of access points (e.g., natural language/controlled vocabulary), and extent of the database.12 Measuring these variables is relatively straightforward and “objective.”

There are four traditional variables in system performance measurement: relevance, pertinence, precision, and recall. Relevance and pertinence are document-level variables: they record how “good” a particular document retrieved by a system is. With relevance, a document is “good” if it matches the question asked. With pertinence, a document is “good” if it matches the asker’s underlying
information need. Precision and recall are search-level variables: they are, respectively, the proportion of “good” documents within a search and the proportion of “good” documents included in the search results. When precision and recall are determined the question remains, of what relative value are they? Cooper argued that unretreived documents (low recall) do not harm the user and are thus unimportant in system evaluation; furthermore, the cost in time and money of attaining 100 percent recall outweighs any proposed benefits. Cleverdon attacked both recall and precision: recall assumes that the user wants all relevant items, and precision assumes the user is intolerant of irrelevant items. Neither of these assumptions, he argued, has been demonstrated. In most research, however, both measures are used; for example, both are present on the Search Evaluation Questionnaire developed by the ALA/RAD Machine-Assisted Reference Section’s Measurement Committee, where they are explicitly stated to suggest user satisfaction or dissatisfaction.

The third step in the MSM is to measure a user’s material satisfaction. Once there is a measurement, it can be compared to system performance (recall and precision) to determine what effect performance has upon it. Most measures employed involve a user describing a search as “useful” or “valuable.” There is, however, no general agreement: table 1 displays definitions and measurements of material satisfaction used in a number of studies.

Definitions commonly used in the MSM follow a circular description of “satisfaction.” If a “satisfactory” search is defined, as, for example, by Burroughs, as one that “retrieves a high number of pertinent items,” then material satisfaction is not only determined by but is identical with system performance (recall and precision). In a similar tautology, Lawton, et al., found that “amount learned” was related to “value of bibliography.”

Separate measures such as “usefulness” or “value” have yielded weak or mixed results when compared with the variables of system performance. Fenichel found no relationship between recall, precision, or unit cost and the searcher’s opinion of the “goodness” of a search. Tagliazuczo, in two studies, did find a relationship between number of citations retrieved (a recall surrogate) and satisfaction. Finally, Hickey and Hurrych reported a relationship between greater precision of a search and greater value placed on that search, but their conclusions are based on a 10 percent return rate from questionnaires sent with search results.

Thus, the MSM is unable to explain the relationship between system performance and satisfaction.

EMOTIONAL SATISFACTION MODEL—SIMPLE PATH

There are three major features of the Emotional Satisfaction Model—Simple Path (ESM-SP), see figure 2, which is centered around the identified quality of “emotional satisfaction.” First, emotional satisfaction is considered a distinct entity that can be independently measured. Second, emotional satisfaction is considered worth measuring because it leads to certain behaviors. Third, it is considered to be caused largely or solely by material satisfaction, either measured directly or considered as system performance.

The first problem to be considered, then, is “What is emotional satisfaction?” Often when referring to “emotional satisfaction” the word “emotional” is not used. What distinguishes material satisfaction research from emotional satisfaction research is that in the former users are described as “satisfied” based upon their responses to material measures; thus, a patron asked whether a search is “useful” is described by the researcher as being “satisfied” when he or she answers “yes.” Emotional satisfaction research, on the other hand, involves an attempt to measure the user’s actual feeling of “satisfaction,” rather than simply trying to infer its presence.

In emotional satisfaction, definition and measurement are tightly intertwined around two issues: direct versus indirect, and multivariate versus univariate. A direct measure of satisfaction employs one or more questions using the word satisfaction: “Are you satisfied with the online search?” An indirect measure asks another question that is presumed to reflect on satisfaction: for example, “What grade would you give the search experience?” Multivariate measures ask questions about a variety of components (e.g., “Are

---

**TABLE 1**

**Material Satisfaction Measures**

<table>
<thead>
<tr>
<th>Definition and Measure of Satisfaction</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>items/bibliography called “useful”</td>
<td>Lipsett &amp; Schultz, 1984</td>
</tr>
<tr>
<td>search called “helpful”</td>
<td>Tagliazuczo, 1973</td>
</tr>
<tr>
<td>search called “helpful” and “useful”</td>
<td>Tagliazuczo, 1977</td>
</tr>
<tr>
<td>items/search called “valuable”</td>
<td>Hickey &amp; Hurrych, 1985</td>
</tr>
<tr>
<td>items/search called “valuable”</td>
<td>Weller, 1985</td>
</tr>
<tr>
<td>items/search called “valuable/usable”</td>
<td>Lawton, et al., 1978</td>
</tr>
<tr>
<td>amount learned</td>
<td>Kiewit, 1975</td>
</tr>
<tr>
<td>number of pertinent items retrieved</td>
<td>Lawton, et al., 1978</td>
</tr>
<tr>
<td>search called “appropriate” and “no”</td>
<td>Burroughs, 1989</td>
</tr>
<tr>
<td>“I found information on my topic” (agree/not)</td>
<td>Stallvict &amp; Harmerly, 1983</td>
</tr>
<tr>
<td>“I successfully answered my question” (agree/not)</td>
<td>Allen, 1989</td>
</tr>
<tr>
<td>outcome: use of a retrieved citation in a subsequent paper</td>
<td>Beghiol, 1989</td>
</tr>
<tr>
<td>intermediary searchers’ opinion: “results were very good, good, neutral, bad, very bad;” also a semantic differential scale</td>
<td>Fenichel, 1980</td>
</tr>
</tbody>
</table>

Sources:


---

**FIGURE 1**

Material Satisfaction Model

- **System Features**
  - see Saracevic and Rees, 1976
- **Product Performance**
  - relevance
  - pertinence
  - precision
  - recall
- **Material Satisfaction**
  - "use"
  - "value"
  - see Table 1
you satisfied with the speed of the search service?"?); the responses may or may not be added to yield a composite score. Univariate measures ask single "global" questions (e.g., "All in all, how satisfied are you?"). Confusion between definition and measurement arises when trying to decide whether multivariate measures are measuring the same phenomenon as univariate measures and whether an indirect question can measure the mental phenomenon of "satisfaction" better than a direct question.

In library and information science literature, one of the most explicit examinations of these issues has been the work of D'Elia and Walsh on satisfaction with public libraries. D'Elia and Walsh define two types of satisfaction: "subjective user satisfaction," which corresponds to emotional satisfaction as measured here, and "objective measures," which correspond to material satisfaction. Because they argue, objective measures are problematic, subjective satisfaction is an appropriate library evaluation measure. This subjective measure can be used for three purposes: "descriptive," to determine the performance of the library as a whole (corresponding to univariate measures); "diagnostic," to determine the relative performance of separate aspects of the library (corresponding to multivariate measures); and "behavioral," to predict user behavior (a separate variable). Subjective satisfaction can be measured in two different ways: directly by questions on "satisfaction" and indirectly by questions soliciting an evaluation of the library in the form of a grade (A to F). In using these measures, however, D'Elia and Walsh have found only a "weak" correlation between the indirect measures of satisfaction (library evaluation grades) and the direct measures.

Murfin and Gugelchuk's work in developing a Reference Transaction Assessment Instrument (RTAI) suggests that multivariate tests have the advantage of being able to detect multiple aspects of satisfaction. In testing their instrument in several libraries, they obtained consistent, reliable results, finding that satisfaction with services and satisfaction with obtained information seemed to be two distinct entities.

Day, Lee, and Johnson directly addressed the multi- versus univariate question. They used both multivariate and univariate measures for both multiple and unitary services. For example, a particular service in the library, such as reference, had a multi-item scale ("how do you like the hours?" "the friendliness?", etc.) and a global scale ("overall, how do you like reference?") and the library as a whole had three individual items (hours, holdings, and performance) and one global question. The various forms of satisfaction all correlated with each other to a low but statistically significant extent (from r = .464 to r = .513), suggesting that they measured a singular satisfaction phenomenon. The most widely used version has eight items (CSQ-8), considered to yield more consistent results than the longer CSQ-18.34

If emotional satisfaction is an entity that can be measured, the questions remaining for this model are: Does emotional satisfaction affect behavior? Is emotional satisfaction determined by system performance/material satisfaction?

Clearly, if emotional satisfaction predicts behavior, knowing a user's satisfaction is valuable. To date, however, findings have been somewhat mixed. On the positive side, D'Elia and Walsh found no correlation between use behaviors and satisfaction among patrons of public libraries. On the positive side, however, Beeler found a correlation between anticipated behavior ("Would you recommend this service to others?") and material satisfaction (results judged "useful"). Bearden and Teel found a correlation between satisfaction and complaint reports. Attiksson and Zwol found a correlation between satisfaction and remaining in therapy and keeping appointments; Sabourin, et al. found a correlation with therapeutic outcome. Ware and Davies found that satisfaction predicted both changes in doctors and disenrollments from HMOs. Both Nath and Hitz and Johnson found that frequent use of computer systems (MIS and an e-mail system) corresponded to satisfaction, but the cause-effect relationship is debatable. The emotional satisfaction measured in these studies does seem to influence behavior.

The remaining question is, what determines satisfaction? This has been left to last because it is the key feature of this model—and of the material satisfaction model. Both the MSM and the ESM-SP assume that if material satisfaction is achieved, emotional satisfaction is also achieved, or is unimportant and nothing further need be done. If emotional satisfaction is determined by system performance (material satisfaction), then this simple-path model is sufficient.

This does not appear to be the case. A few studies find weak relationships between satisfaction and system performance. More often, studies find that
most or even all of the observed variance in user satisfaction cannot be explained by system performance.\(^{11}\)

This unexplained variance is the birthplace of the false positive. Several studies describe individuals who have not achieved material satisfaction but who report emotional satisfaction. Ankeny studied reported success versus satisfaction among 6,000 people who rated themselves satisfied with their careers.\(^{12}\) He found that while 81.7 percent of those who rated themselves satisfied also rated themselves satisfied with their work, 28.8 percent of those who rated themselves dissatisfied also rated themselves satisfied with their work. Therefore, the correlation between satisfaction and success is low. However, 90 percent of the respondents had a false positive, achieving emotional satisfaction with no material satisfaction. Dalrymple studied users of a library catalog and a card catalog; she found that users of the electronic catalog located fewer items but were more satisfied.\(^{13}\)

Similarly, consider the fact that Butler and Kortman found that 21 percent of the users of InfoTrac (a general periodical index) selected and used it in preference to a card catalog, and a further 10 percent substituted it for a more specialized index.\(^{14}\) Thus, even though 31 percent of the users by their own reports could not possibly have achieved material satisfaction, only 17.3 percent pronounced themselves neutral or even mildly dissatisfied (with a response of 3–5 on a scale of 1–5, most to least satisfied).

The inability of the “simple path” to fully explain emotional satisfaction leads to consideration of a multiple-path model.

EMOTIONAL SATISFACTION MODEL—MULTIPLE PATH

In the Emotional Satisfaction Model—Multiple Path (EMSP—see figure 3), emotional satisfaction is considered to depend upon not one but three major variables: disconfirmation, product setting, and product performance (material satisfaction). Disconfirmation is a marketing idea that refers to the difference between a person’s expectations of product performance and the actual performance. Product setting embraces a range of variables relating to such things as price, presence of an intermediary, and a “halo” effect from the service or the library in general. Product performance can be measured with traditional material satisfaction variables.

Taking product performance and emotional satisfaction as defined and measured above, there are two categories of remaining variables: user variables and setting variables. User variables are of two types: organic and acquired. Organic variables are characteristics of the user that determines his or her satisfaction-seeking process. Acquired variables are characteristics that the user acquires in the information-seeking process itself.

Demographic characteristics are the largest and most studied group of organic variables. Library research has examined whether demographic variables affect satisfaction; findings have been mixed. Sandore, Hildbrand, and D’Elia found no difference between demographic groups in terms of satisfaction.\(^{15}\) However, D’Elia and Walsh did find significant, though low, demographic differences in the “grades” assigned to libraries and direct satisfaction with public libraries.\(^{16}\) Allen found that males had a more positive opinion of a CD-ROM system, for example.\(^{17}\)

Psychological and marketing research focuses on psychological organic variables as possible determinants of satisfaction. Vallee and Westbrook report finding suggesting that “locus of control,” or perceived “personal competence,” affect satisfaction.\(^{18}\) Sabourin, et al., tested the hypothesis that self-deception and “impression management” might inflate satisfaction scores, finding a significant correlation. On the other hand, Swoeling and Newell reported significantly lower satisfaction scores when respondents were promised anonymity.\(^{19}\) Somewhat similarly, Tagliacozzo found that “learners” rated their scores as less helpful than did “teachers”; however, he also found that longer searches were often rated helpful and teachers more often had longer searches.\(^{20}\)

Experience is another type of organic variable existing prior to a particular information-seeking session. Nath and Hiltz and Johnson found that greater use of a catalog system and an e-mail system (respectively) was associated with higher satisfaction; however, the direction of causality is unsettled.\(^{21}\)

Westbrook studied satisfaction with household appliances to test a hypothesis that extensive past experience would lead to more realistic expectations, thus a higher chance of fulfilling expectations and higher satisfaction.\(^{22}\) He found a curvilinear relationship with prior experience: the highest satisfaction occurred with intermediate experience. In general, measurement of organic variables presents no special problems. Demographic and experiential information is easily gathered. For testing psychological states (e.g., locus of control), generally accepted instruments exist—though they are seldom used in library contexts.

Acquired variables are those that involve each particular information retrieval episode. The most important acquired variable is expectations. Expectations seem to derive from a person’s individual psychological, personal knowledge of the product due to history or experience (organic variables), and information provided in the course of the encounter.\(^{23}\)

Some studies do not individually measure these components. For example, Churchill and Suprenant studied expectations and satisfaction with chrysanthemums.\(^{24}\) Shoppers were given information that indicated the flowers would bloom in ten, fifteen, or twenty days. This manipulation of information was equated with a manipulation of expectations. Similarly, Sullivan, Borgen, and Wipp and Rushinek and Rushinek did not distinguish among factors comprising expectations.\(^{25}\)

Nola and Swan, on the other hand, addressed experience directly. In their study of patrons of a restaurant, they found experience to be a negligible factor; expectations rose slightly after one visit, then gradually declined to the initial level.\(^{26}\) Barbeau also found a weak relationship between past experience and satisfaction.\(^{27}\)

What has received more attention is the question of whether and to what extent expectations affect satisfaction. This derives from the failure of the product performance or material satisfaction model. If product performance does not determine emotional satisfaction, what does?

Several studies report a direct relationship between expectations and satisfaction.\(^{28}\) However, expectations are more often explored in the context of a model that has expectations come together with product performance to create an intervening variable between expectations and satisfaction. This intervening factor originates in one of three theories: accommodation, equity, and disconfirmation. The accommodation theory postulates that when there is a discrepancy between expectation and outcome, individuals retroactively adjust their expectations to reduce the psychological tension involved.\(^{29}\) Equity theory hypothesizes that an important component of satisfaction is a perceived equality between individual shoppers’ costs and benefits, and the costs and benefits they believe the seller is incurring—if a product meets expectations but the shopper believes the seller has gained disproportionately from the sale, he or she will be dissatisfied.\(^{30}\)
Disconfirmation is the difference between expectations and product performance: confirmation would be an exact match, positive or negative disconfirmation is when the product performs better or worse than expected. Disconfirmation has been found to be a stronger predictor of satisfaction than equity; accommodation is not often tested in research. While Swan and Oliver found a significant effect for equity in their studies, disconfirmation was much stronger. Sullivan, Borgman, and Wippern’s study of satisfaction with end-user versus intermediary online searching operationally equated “general user satisfaction” with disconfirmation. Their instrument to measure “satisfaction with the final free search” was the question: “In comparison with your expectations, to what extent did the results satisfy the information need with which you came?” (four-point scale: results exceeded expectations, expectations were met, results were somewhat disappointing, results were very disappointing). While some studies of users of current versus electronic catalogs included a factor analysis of responses on an attitude survey that suggested expectation as a strong determinant of satisfaction.

The measurement of expectations is in a rudimentary state—similar to many measures of satisfaction, with single-question measurements most common. Expectations, in fact, are often not measured directly. Instead, most research measures disconfirmation. These post-test-only designs operate under the assumed or explicit assumption that pre-test expectations are the same as expectations “remembered” post-test. Disconfirmation is determined not by what a consumer “really” thought “back then” but by what the consumer perceives, post-test, to be the discrepancy between expectations and performance.

An exception to both of these features (one-item measure, post-test-only design), was a study done by Oliver with recipients of flu vaccines. He measured pre-shot expectations with semantic differential scales and post-shot satisfaction and disconfirmation in the form of Likert-scale reaction statements, e.g., “I am satisfied with my decision to receive/not receive the vaccine,” or “If I had to do it all over again I would feel differently about...” Beauden and Teel also measured expectations separately from disconfirmation; in their study, expectations had a stronger effect on satisfaction than did consequent disconfirmation.

Product setting variables are the final category of hypothesized determinants of satisfaction. Setting considered generally has been of concern to D’Elia and Walsh and others, in terms of the separation of satisfaction with a library in general and satisfaction with a particular service. More specific factors identified have been price and interface, or intermediary; elapsed time was considered in early research.

Monetary cost to a user is an obvious factor when considering online search services in general. Measurement is not a problem, although values such as average cost of a search become problematic when some patrons pay nothing. The effect of cost on satisfaction has been found to be statistically significant but trivial.

Work done by Bostrom found that satisfaction with a nondurable good (a potted chrysanthemum) was dependent on disconfirmation as well as product performance; satisfaction with a durable good (a videodisc player) varied only with product performance.

What type of product is an online search? A mediated search? An end-user search of an online system, a CD-ROM, an OPAC? This question has not been addressed directly, but there is some evidence (beyond the fact that a search is usually not a “durable good”) that disconfirmation may play a role. First, Swan and Oliver found that disconfirmation affected satisfaction with a “personal selling” product—the services of a car salesman.

A mediated online search is somewhat similar to a personal-sales product. The impact of the failure of the ESM-SP model. If online services (or other information retrieval products) were like “durable goods,” material satisfaction would directly affect emotional satisfaction. Since it apparently does not, other factors must be involved. Considering the idea of an experimental group and the generally positive satisfaction ratings of online services, a hypothesis might be that persons approach an information retrieval system with low expectations for material satisfaction, so that any degree of material satisfaction achieved, no matter how slight, is higher than expected and results in positive disconfirmation and high satisfaction.

**IMPLICATIONS FOR RESEARCH**

Current understanding of satisfaction suggests several implications for research in library and information science. First, researchers should specify what aspect or definition of satisfaction is being measured in a particular study. Second, researchers should consider using general (non-library) instruments for satisfaction measurement already existing or being developed in other fields, particularly in psychology. Third, researchers should be aware that implicit material satisfaction or simple emotional satisfaction models have generally been unable to explain variation in satisfaction. More complex, multivariate models, the one presented here or others, should be tested to determine if they allow an improved understanding of the phenomena of satisfaction. Finally, researchers must define and separate as clearly as possible the phenomenon (process or product about which they are eliciting satisfaction responses. Marketing research suggests that different factors may be involved in determining satisfaction with different types of products. One model may be best for a “product” such as an intermediary online search; another for the “process” of obtaining information through end-user searches of an online catalog or CD-ROM station.

**IMPLICATIONS FOR PRACTICE**

The most important finding from this research for practicing librarians is a warning against using a simple Yes/No question or questionnaire to determine user satisfaction. It appears that a short-form assessment of user “satisfaction” is not only described and briefly measured, does not provide real information about whether, first, a user’s information need has been met or, second, whether the user is pleased with the exchange. Therefore, it does not answer the question of whether there is anything else the librarian can or should do to improve the situation. Only personal questioning can shed light on this—which is yet another argument for the importance
of the post-search interview, which, like the pre-search interview, should include open-ended, probing questions.

Managerially, the overabundance of false positives is good news. Should an outside agency demand a quantified measurement of “client satisfaction,” research suggests that any manager can cheerfully and confidently comply with such a request: almost every library or service will receive “satisfactory” scores.

The prevalence of false positives calls for a rich, context-sensitive, personal approach to the problem of determining true user satisfaction. Online searchers and other information intermediaries should abandon any reliance on simple questions that provide simple and misleading answers and concentrate instead on establishing a real dialogue in which a user’s needs, expectations, knowledge, and emotions can all be communicated.

In the meantime, research definitions, tools, and models should work together to provide a better understanding of satisfaction. Satisfaction can then be addressed and probed in real reference encounters so that users achieve a level of emotional satisfaction that corresponds to the degree to which their material needs have been met. They will complain more when their needs have not been met, rather than being silenced in the great “false positive.”

REFERENCES AND NOTES


12. Tefko Saracevick and Alan M. Rees, Towards the Identification and Control of Variables in Information Retrieval Experimentation. Comparative Systems Laboratory Technical Report No. 6 (Cleveland: Center for Documentation and Communication Research, School of Library Science, Western Reserve University, Cleveland, 1966).


36. D’Elia and Walsh, “Patrons’ Uses and Evaluations.”
40. D’Elia and Walsh, “Patrons’ Uses and Evaluations.”
43. Sandra, “Online Searching.”
44. Ankeny, “Evaluating End-User Services.”
47. Sandra, “Online Searching.”
48. D’Elia, “User Satisfaction as a Measure.”
49. D’Elia and Walsh, “User Satisfaction with Library Service.”
53. Westbrook, “Correlates of Post Purchase Satisfaction.”
58. Barbeau, “Predictive and Normative.”
60. Howard and Sheeh, The Theory of Buyer Behavior, 221.
63. Swan and Oliver, “The Factor Structure of Equity”; Swan and Oliver, “Automobile Buyer Satisfaction.”
64. Sullivan, Borgman, and Wippern, “End-users.”
65. Dalrymple, “Retrieval by Relational Subsumption.”
70. For example in Beeler, “The Relation of User Fees,” and Lawton and others, A Systems Evaluation.
72. Burroughs, “Clinicians’ Satisfaction.”
73. Lawton and others, A Systems Evaluation.
76. Oliver, “A Cognitive Model.”
77. Churchill and Suprenant, “An Investigation.”
78. Swan and Oliver, “The Factor Structure of Equity”; Swan and Oliver, “Automobile Buyer Satisfaction.”