USERS' RELEVANCE CRITERIA IN MUSIC RETRIEVAL IN EVERYDAY LIFE: AN EXPLORATORY STUDY

Audrey Laplante

École de bibliothéconomie et des sciences de l'information
Université de Montréal, Montréal, QC, Canada
`audrey.laplante@umontreal.ca`

ABSTRACT

The paper presents the findings of a qualitative study on the way young adults make relevance inferences about music items when searching for music for recreational purposes. Data were collected through in-depth interviews and analyzed following the constant comparative method. Content analysis revealed that participants used four types of clues to make relevance inferences: bibliographic metadata (e.g., names of contributors, labels), relational metadata (e.g., genres, similar artists), associative metadata (e.g., cover arts), and recommendations/reviews. Relevance judgments were also found to be influenced by the external context (i.e., the functions music plays in one's life) and the internal context (i.e., individual tastes and beliefs, state of mind).

1. INTRODUCTION

In recognition of the need to provide researchers with an infrastructure for the evaluation of MIR systems and algorithms, Music Information Retrieval Evaluation eXchange (MIREX) was established in 2005. Contests have been held annually since then. Like Text Retrieval Conference (TREC) experiments from which it is inspired, MIREX uses precision-recall measures to evaluate system performance. These are used to measure "the probability of agreement between what the system retrieved or failed to retrieve as relevant (systems relevance) and what the user assessed as relevant (user relevance) where user relevance is the gold standard on the basis of which evaluations are made" [1]. Hence, to establish 'ground truth' for the evaluation of MIR tasks that called for human judgment (e.g., audio music similarity), user surrogates (as opposed to the real users who originated the search queries) were asked to judge *a posteriori* whether the results retrieved were relevant [2].

Although this approach has the advantage of taking into account the human dimension of the process, it also presents limitations. The validity of relevance judgments made in experimental setting is questionable since the criteria used by participants might not correspond to those used by people in real situations. Studies on user-defined relevance conducted in naturalistic settings show that, apart from content-based criteria, criteria pertaining to the user and the user's situation play a significant role in the evaluation of relevance [3, 4]. Therefore, failing to take

into consideration the situation within which relevance judgments occur raises concerns and stresses the importance of studying how relevance judgments are made in real life. While research on relevance criteria used in textual information retrieval can have some utility for the MIR community, studies on video and image information retrieval suggest that differences in information type can result in differences in the criteria used to make relevance judgments [5, 6], hence the need to conduct research on user-based relevance in the context of MIR. Unfortunately, this area of research has hitherto remained essentially unexplored.

The present study was designed to bridge this gap by investigating how young adults make relevance inferences about music items when searching for music for recreational purposes. More specifically, it aims to address the following research questions: (1) What clues do young adults use to make relevance inferences about music items? (2) How do individual characteristics (e.g., knowledge, experience) influence their relevance judgments? (3) How does the context influence their relevance judgments? By providing a rich understanding of relevance judgments in context, this study will be beneficial in many ways. It will provide the MIR community with a better understanding of the behavior of current and potential MIR systems users, which may translate into improvements in MIR system design and evaluation measures.

2. RELATED RESEARCH

Since the 1990s, information scientists have conducted numerous empirical studies on user-based relevance. This has led to a redefinition of the concept of relevance and to an increased knowledge of the criteria used by people when making relevance judgments.

2.1 Concept of Relevance

Researchers distinguish system-oriented (or objective) relevance from user-oriented (or subjective) relevance [7]. According to the former, a document is considered relevant if it is topically related to the search query, a measure that has the useful property of being objective. From the user's point of view, however, topicality was found to be the most important but not necessarily the only relevance criterion. Therefore, a user-oriented definition of relevance was proposed where a document is considered relevant if the user who originated the query judges that it meets his/her information need. This conception of relevance implies that relevance judgments are interpreta-

tional: they can only be made by end-users and are closely tied to the context within which they occur.

User-oriented relevance can be studied from different standpoints: some researchers have focused on its cognitive aspects (*cognitive relevance*) [8]; some on its psychological aspects (*psychological relevance*) [9]; some on its dynamic nature (*dynamic relevance*) [10]; and others on its relation with the situation or task at hand (*situational relevance* or *utility*) [11]. Hence, in an attempt to encompass all these dimensions, Barry and Schamber [12] describe relevance in the following terms:

> "[…] relevance is (1) cognitive and subjective, depending on users' knowledge and perceptions; (2) situational, relating to users' information problems; (3) complex and multidimensional, influenced by many factors; (4) dynamic, constantly changing over time; and yet (5) systematic, observable and measurable at a single point in time."

It is with this multidimensional and situational perspective of relevance in mind that the present research project was designed.

2.2 Studies on User-Defined Relevance

In 1998, Barry and Schamber compare the findings of two studies on user-defined relevance and conclude that a high degree of overlap exists in the relevance criteria used by the participants in both studies. Subsequent studies have confirmed it since: there seems to exist a core set of criteria people use to make relevance judgments regardless of the context [13]. However, depending on the nature of the information, the situation or the user, the weight people attribute to each criterion varies and additional criteria may be employed. Of particular importance to MIR is the research on relevance criteria used when searching for non-textual documents. Choi and Rasmussen [6] found that in the context of image information retrieval, authority was less important than in textual information retrieval, whereas subjectivity and affectiveness—the emotional reaction to an image—played a significant role in the selection stage. Yang and Marchionini [5] found that users of video retrieval systems used textual criteria to start their search but mostly employed visual criteria (e.g., style, color, motion) in the final selection stage. In both cases, as in textual information retrieval, topicality was the most common and important criterion.

Also of interest for MIR is the research on relevance in non-problem-solving contexts, which correspond more closely to situations where people search for music for recreational purposes. Xu [14], who studied how users make relevance judgments when searching for information for its epistemic or entertainment value, found that novelty displaces topicality as the most commonly used relevance criterion.

3. RESEARCH DESIGN

The purpose of the present study was to provide a *rich description* of the way young adults make relevance judgments when seeking music for recreational purposes. Considering the complex and subjective nature of the phenomenon, a qualitative approach was considered best suited.

3.1 Data Collection

The subjective and interpretational nature of relevance called for a method that would allow us to gain insights into the internal behavior of participants (e.g., thoughts, feelings, intentions). In-depth interviewing was deemed the most appropriate method to attain this objective. The literature review on user-oriented relevance provided a useful theoretical background for the development of an interview guide. This guide enabled us to ensure consistency in the topics covered in the interviews while facilitating comparison between participants. During the interviews, participants were asked to talk about their preferred music information sources, to discuss the strengths and weaknesses of these sources, and to explain how, why and in which contexts they use them. Participants were also asked to relate in detail a recent music information-seeking experience.

3.2 Participants

Since music behavior is known to vary according to age and culture, we decided to reduce the heterogeneity of the population by limiting our study to the French-speaking young adults (18-29 years) of the Montreal metropolitan community. Participants (*n=15*) were selected following the maximum variation sampling strategy as described in [11]. Recruitment continued until the saturation point was reached, that is when the information obtained through interviews started to be redundant so that no new themes or patterns were emerging from the analysis.

Among the fifteen participants, ten were male. At the time of the interview, five were full-time students, seven were full-time workers, and three were unemployed. All had a high school diploma, 13 had a college diploma (or the equivalent), and ten had a university degree or were currently enrolled in a university program. None of them were professional musicians but six played at least one musical instrument. The group comprised a majority of avid music listeners, although the sample also included a few light or moderate music consumers.

3.3 Data Analysis

The interviews were recorded and transcribed. Each interview lasted between 38 and 62 minutes, for a total of 724 minutes of recording and over 120,000 words of transcriptions and notes. The software package NVivo by QSR International was used to facilitate the encoding and analysis process.

The data were analyzed inductively using the constant comparative method (CCM) as defined in [12]. CCM consists in a step-by-step method according to which the researchers (1) prepare the data for analysis by subdividing the transcripts into units of meaning (in this case into paragraphs); (2) read through the data to identify emerging themes and patterns in order to create a provisional set of categories; (3) categorize each unit of meaning into a category, forming new categories as needed; (4) refine the categories by comparing all units comprised into each category in order to identify the common properties or characteristics, merging, subdividing, or restating categories as needed; and (5) explore relationships and patterns across categories.

4. RESULTS

The interviews elicited a wealth of information about participants' music information-seeking behavior, including their likes and dislikes in terms of music information sources and the way they interact with these sources. The analysis revealed that the participants' relevance judgments were the results of a combination of different criteria and factors: criteria pertaining to the music itself (e.g., quality), the physical document (e.g., disk), the external context (e.g., intended use), the internal context (e.g., disposition), and their personal knowledge and experience. To determine the likelihood that a music item meets these criteria, participants used a variety of clues.

The results presented here cover the clues used to make relevance inferences and the criteria associated with the internal and external contexts of the search. Quotes were translated from French to English, while maintaining as much as possible the level of language used by the participants.

4.1 Relevance Clues

For all participants, listening to the music retrieved played a crucial role in determining its relevance. This, however, requires time and effort that, in some conditions, participants did not consider worthwhile. Moreover, the content of the music was not the only criterion. Therefore, at least in the first stage of their search, participants reported employing a variety of clues or extra-musical information to make inferences about the type of experience a music item offers, and assess the probability that it meets their desired criteria. As these clues are not always self-explanatory, previous knowledge and experience was often called upon to interpret them.

4.1.1 Bibliographic Metadata

Bibliographic metadata refer to the information used to describe an item, which, for music recordings, includes performers, composers, authors of lyrics, titles of songs/pieces and albums, labels, etc. This type of information appeared to be commonly used by participants in the selection process. The names of the main contributors (e.g., singers, bands) seemed to have the greater impact on their selection. A good experience with an artist could even transform some of them into committed fans having an almost unwavering faith in any new project to which this artist contributes. This explains why a few participants admitted buying CDs of their favorite artists without even listening to them beforehand. Conversely, disappointing experiences with an artist also increase the likelihood of discarding an item without listening to it. This speaks of the notion of authority, a relevance criterion also used in other contexts of information retrieval.

In the same way, for the most avid music listeners, trusted labels represented a guarantee of quality. When asked whether he would borrow an album by an unknown artist from the library, one participant explains: "if it's on a label, maybe. […] if it's a bluegrass band, for instance, because it's on Smithsoninan, it must be good." Another affirms that he "pretty much trust[s] what [Matador] release." Of course, metadata regarding contributors or labels are only useful if those are familiar. The fact that labels were only used by heavy music listeners suggests that extensive music knowledge is required to interpret this type of information.

Other bibliographic metadata proved to be of lesser importance for relevance judgments. Composers and authors of lyrics were explicitly mentioned by none of the participants, which might be due to the fact that many of the participants' favorite bands and singers composed their own songs. Album titles did not seem to affect selection either and only one participant affirmed relying on song titles to determine if an album was worth borrowing from the library.

4.1.2 Relational Metadata

Lee and Downie [15] define relational metadata as data regarding relationships between music items (e.g., music genres and similarity). Relationship between artists (e.g., collaborations, influences) could be added to this category. For several participants, links between artists allow them to situate an unknown artist in the music sphere, thus helping them assess the probability that the music of this artist corresponds to their taste. Hence, one participant related having discovered a group by reading an interview in a music magazine in which the group members were citing "their inspirations", which turned out to be "things [he] like[s]", which convinced him to listen to their music. This could also explain why *MySpace Music* and *allmusic*, two sources that incorporate a plethora of relational metadata, reached the top of the list as the most popular music information sources on the Web among participants. In *MySpace Music*, the "Friend Space" that connects artists and regular users was considered the source's most useful feature. Browsing the list of friends of a group one loves was a common strategy to discover new (but similar) groups as exemplified by this quote from a participant: "It's also a good way […] to find groups that make similar music and groups with which they do concerts… It's really, really useful!" Likewise, in *allmusic*, participants greatly appreciated the links to influencers, followers and similar artists in the articles, although the 'similar artists' links, which are created by music experts, were not unanimously considered reliable. Indeed, two participants complained about links leading to artists that "were not really similar" or "not similar at all."

While links between artists are primarily used to make inferences about individual objects, music genres, which are meant to bring similar or somewhat similar albums together, were mostly employed at an earlier stage, for instance to discard several items at once in order to get a manageable set of items to browse. Hence, one participant reported going directly to the "Film Music" section at the library, whereas another mentioned regularly searching for "rock" in online music stores. Even though participants widely used genres to search or browse music collections, they also frequently complained about them. Common criticisms included (1) some groups do not clearly fit into one genre ("It's oversimplified. You can't always categorize a group into something."); (2) genres are too broad ("[searching by genres on the Web] generates endless lists", or "[In music stores] they put everything that is rock, alternative, punk, metal together!"); and (3) genres are too narrow ("[On *allmusic*] they have something like 600 styles of music, it's not concise

enough. […] They make up terms I've never seen and there are basically two groups that make this type of music, maybe even just one!"). As illustrated by this last quote, genres that are too specific also often sound unfamiliar and are therefore useless ("I could tell you there are 40-50% [of the genres on Limewire] I have no idea where it comes from or what it is"). The apparent contradiction between participants' complains regarding the level of specificity of genres can be explained by the source used to find music: genres in brick-and-mortar music stores are usually very broad, whereas specialized music sources on the Web tend to use very specific genres. But this does not explain entirely their dissatisfaction. Their knowledge of music or of a particular genre also influenced their perception of genres. While one participant admits knowing "seven genres, eight at most!" another explains that "there are five, six styles of punk."

4.1.3 Associative Metadata

Lee and Downie [15] define associative metadata as data about the relationship of a music item with events or other forms of art, which includes cover arts. While some appreciated cover arts for their intrinsic beauty, participants also appeared to attach importance to them for what they tell about the music. When little or no costs are involved, assessing the potential value of an album on the basis of its cover art could even represent a valid alternative to sampling the music: "At Cheap Thrills [a music store], they have a one-dollar rack. There are groups I don't know. The only thing I can look at, really, is the cover art. If it looks cool… And it often happens to fit. I find music I like in jackets I like!" Most of the time, however, participants employed cover arts only to make a first selection among albums, the following step being to listen to the music. Hence, when asked how she selects CDs in music stores, one participant describes that she looks at CDs "one by one" and selects those who have a nice cover art because "it says a lot" and "it has to represent something." She will then wait to be at home to download the album and see if it is worth buying it. Another participant, whose previous experiences had led him to conclude that cardboard jewel cases often contained music he liked, had come to use that criterion to make a first selection, a method that seemed to pay off ("Often, it happens to be albums I like!"). For two participants, however, cover arts had no influence on their selection. One participant explains that "[shopping for] music is not really visual" and cannot be done "simply by wandering around, looking at album covers." Another mentions that her previous experiences have convinced her that cover arts should not be considered as indicators of quality ("I've bought so many good albums in hideous jackets in my life!").

4.1.4 Recommendations and Reviews

A majority of participants affirmed that they attached importance to recommendations, reviews and ratings. The trustworthiness of the source was determinant in the value they ascribed to the information. Hence, recommendations from friends or colleagues perceived has having discriminating judgment and tastes that are similar to theirs had the greatest influence on their relevance judgments. A few participants also relied occasionally on record store staff.

In these cases, as they did not know them personally, the decision to trust a person was based on (1) his/her general look: the person has to look like someone who "pretty much listens to what I listen"); and/or (2) where that person works: people working in small and specialized music stores ("a small, underground CD store") tended to be perceived as especially trustworthy. In contrast, because of past disagreements with critics' reviews, a majority of participants affirmed not being influenced by them.

Recommendations and reviews provided information that could be useful at different stages of a search. Participants sought recommendations to begin a search, to obtain information that has been filtered specifically for them ("It's like a filter [or] a bit like a shortcut"). It saves them time, while increasing their chances of finding something interesting ("You're more likely to come across something good right away"). This information can also be useful to make relevance inferences *en route*. Hence, they were more likely to pick up an album if they had heard of the artist before. Star ratings or popularity sorting, on the contrary, were used at a later stage, to identify albums or songs "that best represent the career of an artist." Reviews or recommendations could even change their initial relevance judgment. Indeed, one participant admitted having changed her mind about some music artists because of friends who "had found arguments" that had allowed her to listen to the music "from a new angle."

To conclude this section on the clues participants used to make relevance inferences about music items, we shall mention the types of information that had little or no influence. Interviews or biographies were mentioned only by a two, who were mostly interested in the professional life of the artists and in the relational metadata (influences, collaborators, etc.) they find in them. Related to that, participants seemed to favor information that could be scanned quickly: ratings or editor's picks, for instance, were far more popular than long, written reviews. Also absent from the picture were the lyrics (or the topic of the lyrics). Although the topic of a song or an album could sometimes be determinant in deciding what one would listen to in a specific situation, the lyrics appeared to be of little importance in determining the relevance of an item during the search process. This might be explained by the fact that although the participants spoke French, a majority mostly listened to Anglophone music.

4.2 Context

Saracevic [16] maintains that relevance "cannot be considered without a context" and defines the context as the result of a "dynamic interaction between a number of external and internal aspects." In line with this definition, our analysis revealed that the context in which a search occurs affected the relevance judgments of the participants in various ways.

4.2.1 Situation

We usually define situational relevance as the relationship between an information object and the user's information problem or task. This definition, however, did not seem entirely appropriate for this study since, according to our participants' accounts, searching for music for recrea-

tional purposes is rarely a task-oriented activity. But this does not mean that the situation had no influence on their relevance judgments. In reality, the roles music plays in their lives affect their information-seeking behavior. Participants used music for a variety of functions. While some reported listening to the same type of music regardless of the context, most affirmed selecting different genres of music in different situations. Thus, when seeking music, they always bore in mind the potential functions the music encountered could potentially fulfill.

It is easy to see how this can affect relevance judgments. Music, for instance, was the soundtrack of mental or artistic work for 11 participants. The reasons for listening to music while working were various: music can (1) contribute to inducing concentration ("when I forget my headphones, I'm kind of not productive"); (2) make the work more pleasant, especially a tedious or repetitive task ("it helps me get through the day"); or (3) provide inspiration for artistic activities ("[it] helps me get into a mood"). When used in this context, the primary selection criterion was that music did not interfere with their thinking, which usually meant instrumental and/or repetitive music such as classical, techno, or electronic; music genres they would not necessarily listen to otherwise. Hence, for one participant who usually listened to old French music, the music that played in the background when she was working was totally different: "If I'm at work and need a lot of concentration, I will put on techno music with no lyrics."

When music is used to maintain or establish interpersonal relationships, the selection is once again affected. As a matter of fact, some participants mentioned listening to different genres of music with different persons so that it would please everyone, maybe even music they wouldn't have listened alone, as illustrated by this quote: "My mother hates Pink Floyd. [...] So when I'm with my father, we listen to Pink Floyd, but when I'm with my mother, I'll listen to something else. I can even listen to some Luce Dufault."

Music is also used to manage one's mood. Whereas some participants used music to modulate or enhance their mood ("If I get up on the wrong side of the bed, I put on music that will make me happy for the rest of the day"); others sought music that matched their current —usually depressed— mood ("If you're broken-hearted [… and you listen to] *Rose* by Portishead, you clearly know that they feel like shit, just like you.").

4.2.2 Individual Tastes and Beliefs

Not surprisingly, one of the main criteria employed to make relevance judgments was affectiveness or the emotional response to music. In other words, music usually has to meet one's taste to be considered relevant. Indeed, although participants reported occasionally selecting music outside of their regular tastes to fulfill specific functions, they still wanted this music to be as good as it could be. This explains why participants believed listening to the music was an essential step in formulating relevance judgments unless, as mentioned before, they considered that the risk incurred was low (e.g., highly trusted artist, music is free or almost free). For that purpose, partici-

pants frequently visited the *MySpace* profiles of artists or downloaded music illegally to be able to listen to entire songs or albums ("it allows you to really see what the song is like, the melody, see if you like it or not"). Indeed, although most online music stores or other music information sources propose 30-second excerpts, this was considered insufficient to make inferences about the work of an artist.

As seen in Section 4.2.1, music serves different functions, one of which being to help people define their identity, an area of research that has been widely studied by sociologists and psychologists. Through their music tastes, people express who they are—their attitudes, values, and opinions. Of course, people also use music preferences to make inferences about others. This has repercussions on their information-seeking behavior: people want the music they retrieve to correspond to their values and beliefs. Such behavior was common among participants. One participant said that he liked music that "has meaning" and "would feel guilty if [he] liked the music of someone [he] hated." Another admitted attributing a lot of importance to finding underground groups "because I tend to want to be unique." In fact, many participants showed a strong penchant for non-commercial music, which could be explained by their desire to distinguish themselves from others by having unique music taste.

Related to that, the geographical provenance of music artists seemed to influence the perception of a few since five participants showed a marked preference for local artists (one confessed that she was more "opened" to Quebec groups, while another explained that she really liked following "what's happening on the Quebec scene").

4.2.3 State of Mind

One's state of mind also affects music perception. As a matter of fact, nine participants admitted that discovering new music artists or genres required mental effort and an openness of mind they only had in certain contexts. Four said they could only appreciate unfamiliar music if they had "time to waste" so that they could settle down and concentrate on the music. Three affirmed they needed to be receptive to novelty ("I need to be in a different state of mind. [...] You really have to say 'Ok, I need to adapt'"). Moreover, since music is often used for mood management (as seen in Section 4.2.1), one's current mood also influenced music selection.

5. CONCLUSION

This study sheds light on some of the particularities of relevance judgment in the context of MIR, more specifically in situations where people seek music information for recreational purposes. Although we found that a significant set of criteria used in textual information retrieval were also applicable in this context (e.g., quality, authority, familiarity, situation, user's knowledge and experience), some unique characteristics also emerged. Findings suggest that criteria pertaining to the user, especially individual tastes and beliefs, have a greater impact on selection than in other contexts. This could be due to the subjective nature of music perception and to the fact that mu-

sic tastes often act as a 'social badge' that conveys information about people. Moreover, while topicality was found to be the most common and important relevance criterion in textual, image and video information retrieval, our study revealed that it was not used by our participants, possibly because they attached little importance to lyrics. Therefore, genre, which has the property of allowing one to obtain rapidly a manageable set of items, displaced topicality as the most commonly used criterion to start a search. Our analysis also uncovered the importance of recommendations and reviews from trustworthy sources at different stages of the music selection process. On the other hand, this study reinforces findings from previous studies by confirming the importance of affectiveness as a criterion for making relevance judgments about non-textual information [5]; and the importance of novelty in recreational contexts [14].

This study provides indications for the design of MIR systems and interfaces that support users in their relevance judgments. It reiterates the need to provide rich metadata, including links between artists, not only to facilitate the search but also to better assist users in their selection. However, the fact that some metadata were useless to people who did not have the required knowledge to interpret them suggest that systems should also assist the user in this task, for instance by providing descriptions for music genres or music labels. The importance ascribed to recommendations from people one knows and trusts indicates that systems should include social networking tools that facilitate the sharing of information between users. The study also revealed that music preferences could change depending on the context (e.g., one's current mood, functions music plays in one's life). A successful recommender system should therefore be able to handle this complexity and allow people to have multiple 'music personalities,' thus recognizing the dynamic nature of relevance.

6. ACKNOWLEDGEMENTS

7. REFERENCES

[1] T. Saracevic, "Effects of inconsistent relevance judgments on information retrieval test results: A historical perspective," *Library Trends,* vol. 56, pp. 763-783, Spring 2008.

[2] M. C. Jones*, et al.*, "Human similarity judgments: implications for the design of formal evaluations," in *Proceedings of the 8th International Conference on Music Information Retrieval, Vienna, Austria, September 23-27*, S. Dixon*, et al.*, Eds., ed Vienna: Österreichische Computer Gesellschaft, 2007, pp. 539-542.

[3] L. Schamber, "Relevance and information behavior," *Annual Review of Information Science and Technology (ARIST),* vol. 29, pp. 3-48, 1994.

[4] C. L. Barry, "User-defined relevance criteria: An exploratory study," *Journal of the American Society for Information Science,* vol. 45, pp. 149-159, 1994.

[5] M. Yang and G. Marchionini, "Exploring users' video relevance criteria: a pilot study " in *ASIST 2004, Proceedings of the 67th ASIST Annual Meeting*, L. Schamber and C. L. Barry, Eds., ed Medford, NJ: Information Today, 2004, pp. 229-238.

[6] Y. Choi and E. Rasmussen, "Users' relevance criteria in image retrieval in American history," *Information processing and management,* vol. 38, pp. 695-726, 2002.

[7] D. Swanson, "Subjective versus objective relevance in bibliographic retrieval systems," *Library Quarterly,* vol. 56, pp. 389-398, 1986.

[8] P. Wang and D. Soergel, "A cognitive model of document use during a research project. Study I. Document selection," *Journal of the American Society for Information Science,* vol. 49, pp. 115-133, 1998.

[9] S. P. Harter, "Psychological relevance and information science," *Journal of the American Society for Information Science,* vol. 43, pp. 602-615, 1992.

[10] L. Schamber*, et al.*, "A re-examination of relevance: toward a dynamic, situational definition," *Information Processing & Management,* vol. 26, pp. 755-776, 1990.

[11] P. Wilson, "Situational Relevance," *Information storage and retrieval,* vol. 9, pp. 457-471, 1973.

[12] C. L. Barry and L. Schamber, "Users' criteria for relevance evaluation: A cross-situational comparison," *Information Processing & Management,* vol. 34, pp. 219-236, 1998.

[13] T. Saracevic, "Relevance: A review of the literature and a framework for thinking on the notion in information science. Part III: behavior and effects of relevance," *Journal of the American Society for Information Science and Technology,* vol. 58, pp. 2126-2144, 2007.

[14] Y. Xu, "Relevance judgment in epistemic and hedonic information searches," *Journal of the American Society for Information Science and Technology,* vol. 58, pp. 179-189, 2007.

[15] J. H. Lee and J. S. Downie, "Survey of music information needs, uses, and seeking behaviours: Preliminary findings," in *5th International Conference on Music Information Retrieval*, Barcelona, Spain, 2004, pp. 441-446.

[16] T. Saracevic, "Relevance: A review of the literature and a framework for thinking on the notion in information sicence. Part II: nature and manifestations of relevance," *Journal of the American Society for Information Science and Technology,* vol. 58, pp. 1915-1933, 2007.