Audio Preview Cues: Support for Exploration of Music Information Spaces

m.c. schraefel
Department of Electronics and Computer Science
University of Southampton
Highfield, Southampton, UK
E-mail: mc at ecs.soton.ac.uk

Maria Karam, Shengdong Zhao
Department of Computer Science,
University of Toronto,
Toronto, Canada
E-mail:{mkaram | sszhao} at dgp.toronto.edu

ABSTRACT
This paper presents a novel mechanism that seeks to allow people to explore large collections of loosely structured audio. The approach provides a lightweight preview mechanism that allows people to explore the audio collection by providing supporting information (analogous to the use of tooltips in visual interfaces) We present an evaluation of these "preview cues" towards developing a design heuristics for their deployment.

KEYWORDS: Multimodal interfaces, information access, data exploration, audio interfaces

1 INTRODUCTION

There has been rapid growth in online music delivery, from sites that sell CDs to peer-to-peer networks that let users swap music files. People browsing for music, however, are poorly supported by these interfaces, which rely largely on keyword searches. Such searches bias discovery towards content that is already known. If users do not know a name of a piece or an artist they wish to acquire, they are likely out of luck. Yet browsing/searching for music is a standard method for content discovery, with many record shops such as Amazon.com or HMV, and peer-to-peer clients like Limewire and Kaza oriented toward this mode of delivery.

This paper presents a novel mechanism that seeks to allow people to *explore* large collections of loosely structured audio without relying on previous expertise about that domain (which a keyword search requires). The approach, called "preview cues" provides a lightweight preview mechanism that allows people to explore an audio collection by providing supporting information (analogous to the use of tooltips in visual interfaces or Earcons in audio interfaces) at the point of interest.

In the following sections, we consider the current state of the art of online music delivery, and how it fails to support resource discovery. We then consider Preview Cues as a possible solution.

2 BACKGROUND

2.1 Music Discovery

It is difficult to find another domain outside music where we generally have less key information to make a determination about an acquisition, whether we are online or off. Imagine buying a picture for one's home the way we are usually constrained to buy music: the images would be in racks, each image wrapped in brown paper, with only a title of the image on the wrapper, and perhaps a picture of the artist on the label. The fundamental information about which we make decisions about the picture – the picture itself – is withheld. To imagine looking for visual art this way is laughable, and yet this is exactly how, for the most part, we shop for music: we have to know what we want, whether through recommendation or direct knowledge, and ask for it by name, otherwise we have little means for accessing the critical content – the music itself – to make a determination about preference.

The Real World. While it is true that many music store chains have "listening bars" where theoretically people can evaluate CDs before purchase, the overhead for checking out music in this way at a record store requires a significant commitment of time. In a design session we held with 6 participants, varying in age from 16 to 45, we learned that a trip to the Listening Bar had to be planned: discs that one *had already heard about* needed to be located and a request to the store to cue up the disk needed to be made. Headphones then had to be donned, tracks found and evaluation initiated. One of our participants, a single mother, expressed frustration at the "presumption" that she would have this kind of time.

The Virtual World. With high speed internet access and multimedia computers, it would seem obvious that a networked computer would be able to enhance the online music discovery experience. This does not seem to be the case. The current premier music-purchasing site is Amazon.com. At the Amazon site, one can pick areas of musical interest (country, rock, classical, etc). After one has chosen an area, one is presented with a simple list of the titles in that category, with pictures of the associated album covers per each entry. If the user clicks on an entry in this list, a page for the sale of the album is presented. If the user scrolls to the bottom of this page (sometimes more than two page scrolls down) the user will find a list of album tracks, and be able to stream 30 second samples from selected tracks. In other words, it is only after multiple selections have been made, from type of music, to genre, to individual album that one is presented with any audio samples for making a decision about an album.

2.1.1 *Classical Music: a Test Case for Discovery.*

Even sites that are presumably written to inform users about a particular music genre, and that have no ties to music vendor sites, use the same model for supporting information discovery as the CD sales sites like Amazon or the AllMediaGuide.com. We considered classical music sites in particular, since this is an area with a well-defined nomenclature and well-established categories for organization, which might give users more dimensions through which they could explore a site. Classical music is also an area that, from a preliminary design survey, we found most participants had the least amount of musical knowledge but the highest degree of interest in discovering new music (just edging out jazz). So, we considered what participants who wanted to discover new (to them) classical music would be able to do using online resources to help them.

In surveying online sites about classical music, we found few that linked to the actual music. Typical sites such as "The Classical Music Pages" [2] allow users

to search classical music by various categories (epoch, form), but the site is more clicking through categories about classical music composers than an exploration of classical music itself. The Indiana University Variations digital library project [7] seeks to provide digitized musical content to its faculty and staff at point of access, rather than requiring users to lookup stack locations for recordings, fetch the recording and find a player for the data. The motivating scenarios for the project, however, are to facilitate users who have a specific piece for retrieval in mind, rather than exploration of the catalogue by non-expert users. The Philip Glass Engine [9] uses sliders to let users navigate through the Glass catalogue, playing whatever composition the user lands on. The sliders effect a one-to-one correlation between a title and a piece, so the user can hear the music of a selected track. The interface is frustratingly slow, however, making it difficult to match up in real time what has been cued up to play with the current state of the interface.

Based on this preview, we then did a fast test in our lab with several volunteers asking them to find 4 pieces of classical music within 20 minutes that they could say they had not known previously and would enjoy listening to. One participant found 2 pieces in 20 minutes; another gave up in expressed frustration after ten minutes. The other two expressed getting bored and frustrated after finding one piece in about 7 minutes each. Both observation and participant comments afterwards indicated that Web page search and navigation did not support either discovery or assessment.

2.1.2 Audio Preview Cues

Based on our survey results, participant design sessions, and pilot study above, we hypothesized that if we combined audio streams with terms in an organized visualization of the domain, we would have an alternative tool to keyword search and target acquisition to support exploration and discovery of music in online sources.

We also hypothesized that these *audio preview cues* would be a fundamental UI improvement to provide effective access to the musical domain for domain-naïve users in particular: by associating audio cues at the location of a domain-specific term (such as sonata or Baroque) we would give users a means to make a determination about whether or not this was an area they wished to explore further.

To test our hypothesis, we designed several prototypes to assess when, where and how audio preview cues may be optimally deployed to support access and exploration of a music domain. In the following sections, we describe the interface designs, the evaluation of these designs, our analysis of the results and our plans for future work.

3 AUDIO PREVIEW CUES CONTEXT

Audio preview cues have two components, the audio cue, and the information representation behind the cue. We describe each in turn, and then situate the cues in terms of related work.

3.1 The Audio Component

The design goal for an audio preview cue is to help associate what a user already knows or can assess (represented in the preview cue) with domain specifics (domain lexicon/organization). We associate audio preview cues with domain-

specific labels, such as Romantic, Agricola or Serenade, to give the user a sense of the audio traits inherent in the terms.

3.2 Information Representation

Audio preview cues presuppose an organized domain: in an information space using preview cues, users might encounter the term Baroque as part of a category like Period or Style. They can trigger an audio cue at the site of that term in order to help them to decide first whether or not this sound/style is of interest to them and second whether or not they wish to explore further the attributes like Composers or Arrangements associated with that term. Exploring further means that the user does in fact have access to *more* information specifically associated with that term: the term forms part of a possible path the user can explore. That means such a path needs to be available behind that term. Such path support is the more implicit part of a preview cue.

There are many ways to support users' access to these paths, and many kinds of representations of these hierarchies, from lists to hyperbolic trees. While we have been experimenting with various visualization models. For this study, we chose variants of columnar views of hierarchies – single, and multicolumn visualizations – which have been evaluated in document design [5]. We describe the rational e for these particular choices in section 4.1 on the experimental design, below.

4 RELATED WORK

Auditory preview cues are related to but distinct from both Tool Tips and Earcons. Tool tips are generally employed as text, which appears when a user brushes (mouses over) an unlabelled icon for a specific command in the tool bar of an application. The text describes what the icon command invokes when clicked. Similarly, earcons are highly structured non-speech auditory cues in which the associated auditory cue represents one specifically defined UI event, such as the selection of a particular tool [3]. More recently, Terry and Mynatt have proposed Sideviews [12] previews for graphics applications in which an artist can preview multiple versions of a filter on an image, rather than a seeing a preview of only one filter setting, as is common now.

Preview cues are similar to tool tips and Earcons in that they provide additional information about a UI marker, but they are also broader – they do not need to be so semantically specific. They are more "intensional," in the Montague semantics sense of the term [13]. That is, rather than defining a specific command or reflecting a specific state, or as with Sideviews a set of explicit states, preview cues, suggest a potential *range* of values associated with a given area of a domain.

With preview cues, people who previously knew nothing about "Baroque" and would have therefore passed over this name in a list of links about equally obscure classical music terms, would now quickly be able to decide whether or not they liked baroque music and wanted to hear more. Preview cues provide the user with useful information in context and at a low cost. In the time it takes the user to trigger a cue in the interface, they have information about an area of a domain.
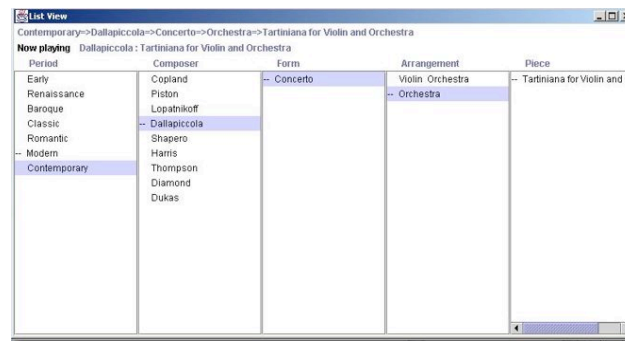
5 EXPERIMENTAL DESIGN

Figure 1: Spatial Multi-Column Layout view
(with Labels).

On the experimental level, we wish to consider whether audio preview cues improve music domain exploration and discovery; on the practical level, we hope our research will lead to principled design heuristics for audio exploration interfaces. Therefore, our experimental factors compare attributes of extant Web/Peer-to-Peer like interface designs in order to see how cheaply and effectively existing designs may be adapted to take advantage of the technique.

5.1 Protocol

We used a one-between/two-within participants repeated measures ANOVA design. We evaluated two treatments of two interfaces types, and tested the same two conditions in each treatment, counterbalancing 24 participants in each treatment for a total of 48 participants. The study was gendered balanced in each treatment, and ranged in ages from 18-54.   The interface types compared a single column (temporal) view with a multiple column (spatial) view of the domain hierarchy; the audio condition compared when in the hierarchy a cue is available (at each point in the hierarchy; only at the final level of the hierarchy). One treatment used persistent text labels for each attribute in a hierarchy; the other used only mouse over text to reveal label name. This yielded a total of 8 interface conditions. To reduce the possible fatigue of participants, we used the Label/No Label treatment for the between subjects portion, so that participants saw only 4 interfaces. A pilot test with 20 participants revealed that the Label/No Label split preserved task focus better than trails that mixed Label/no Label treatments, which proved particularly disorienting to some users. There was no such disorientation when only label or no label interfaces were used. We describe the rational for each treatment/condition below.

5.1.1 *Persistent Attributes Across Conditions.*

In each interface we represented a classical music data hierarchy, organized as Period, Composer, Form, Arrangement and Piece. Beneath each category header, users see elements of the hierarchy. For instance, Period shows Early, Renaissance, Baroque, Classic, Romantic, Modern and Contemporary elements. Also, each interface displayed the path to the current element in the upper left of the window. It also showed the composer and name of the work currently playing in an area labeled "Now Playing." The path to the current location in the hierarchy was also always highlighted: if the user selected Romantic in Period, and Mozart under Composer, both Romantic and Mozart would be highlighted with a pale blue band, as per Figure 1, showing the spatial version of the layout.

5.1.2 Column Condition.

The single column view simulated Web-based exploration of hierarchies, where clicking on one level of a hierarchy generally takes a user to a new page representing the next level of the hierarchy (the Internet Directory, Yahoo.com). Context of where one has been is largely maintained temporally, but for textual information that describes the current path. We refer to this view throughout as the Temporal interface. We use a multicolumn view to maintain spatial context: the previous part of the path remains persistently available.

Extant research would suggest that the multicolumn view would be preferred to the single column view. Such work has not considered the inclusion of path information in the single column view. By reevaluating single column views with path information (the Web model) against multicolumn representations (the simple shift of variable from temporal to spatial views) we can assess two qualities: first, if preview cues enhance the exploration experience for the Web-like, single column model, then this will be a cheap mechanism for improving existing page designs. If the column or spatial views with preview cues afford significantly greater improvements in experience and efficiency, then this would provide a compelling heuristic for interaction designers who wish to best support audio information delivery.

5.1.3 Audio Condition.

In each interface condition, we also tested when audio cues would be available. In the Early condition, preview cues were available for each label in each level of the hierarchy. In the Late condition, preview cues were only available when the user reached the final "Piece" level of the hierarchy. The late case simulates the manner in which Web sites which support any kind of audio audition present audio cues: only at the selection of the final level of the hierarchy, such as an album in a commercial site after download in a peer-to-peer client. The early/late audio conditions would help us confirm when audio cues are best suited to be available. This condition proved to return the most surprising result.

5.1.4 Label Treatment.

While our core evaluation interest lay in comparing standard Temporal/Late interfaces against against modified Early/Spatial ones, the between subjects Label treatment allowed us, at little extra cost, to investigate a visualization issue from a multi-modal perspective: potential scalability of this technique over a large data set. We wanted to see if, in a space where the labels are themselves largely meaningless to the majority of the domain-naïve participants, the absence of labels had any effect on exploring a space, where exploration would be largely informed by the audio rather than persistent text-based feedback. Thus, in the label condition, each element of the hierarchy was presented as a text label. In the No Label treatment, each element in the hierarchy was presented as a vertical bar. When the user brushes over the bar, a tool tip of the label associated with the bar is revealed (Figure 2, left pane).

The use of bars with mouse over text is inspired by Microsoft Research's Social Computing Group's Stacks photo browser. The browser presents stacks to represent stacks of photos; the photo image associated with that element of the stack is revealed on brush over [8]. In our case, the user hears an audio preview cue with the label of the cue currently playing, rather than sees a photo with the file name of the selected image for as long as the mouse is over the target.

Based on information visualization research [1], [11], [6] which suggests that providing at least two representations of the same data provides users with a performance boost in navigation and recall tasks in the hierarchies, we wanted to get a preliminary sense of whether or not the audio preview cue behaved similarly to these heretofore visual presentations of the data.

Success with a visually low volume marker like a bar rather than text could prove significant for representing large volumes of data where space for full text labels is at a premium. If successful in the no label trial, the technique could also be adopted to situations where either there is limited or no visual cue feedback, or where the visual channel is taken up, such as driving or walking. The no label condition test would give us insight into whether a combination of Earcon, audio cue and voice label may may be effective.

Interface Operation. In the multicolumn interface (Figure 1), when a user clicks on an element in a category, that selection is used to populate the next level of the hierarchy (Composer) in a pane beside the original elements (Period). So, a user who clicks Romantic, opens a composer listing that could include Beethoven, Haydn, Mozart. In the single column interface, when the user makes an element selection, the expanded area of the hierarchy appears in the current window, replacing the content of the original window.

In both the spatial and temporal interfaces, the upper left corner of the window provided path information so that users could see where in the data hierarchy they were. Where in the spatial condition, the user could click inside an alternate pane to move backwards or forwards in the hierarchy, users of the temporal interface had browser-like back and forward buttons for navigating the hierarchy.

In each interface there were 3 consistent actions:
- *single click* on an element in the hierarchy selected and revealed that area of the hierarchy
- *right mouse click* added the element's associated musical piece to the participant's purchase list.
- *brush over* an element initiated the playback of an auditory preview cue. A piece continued to play until another label was brushed over.

In the Late condition interfaces, brush over started an audio preview cue only at the Piece level of the hierarchy. In the Early condition, all elements brushed over, anywhere in the hierarchy, initiated an audio preview cue.

5.1.5 Selection Window Note:

Brush over also operated in all conditions in the Selection window, allowing participants at any time to review quickly the audio cue associated with the name of the work they had selected, and thereby quickly compare selections to see which of the ones they had they wanted to keep.
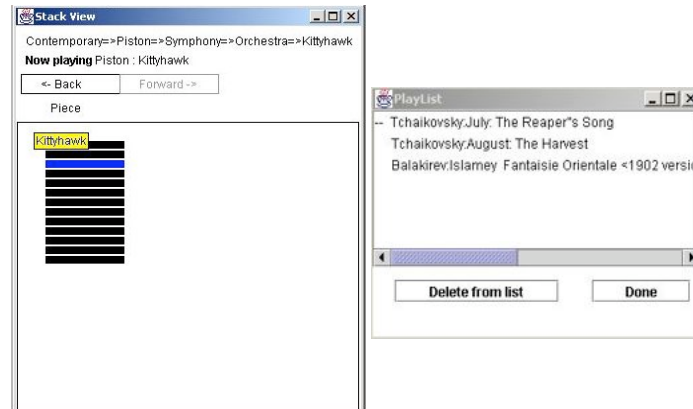

Figure 2. Temporal Single Column Interface, No Label Version (left pane) with Selection Window (right pane).

Task. Participants were given the following scenario: they would be given 100 virtual dollars per interface to purchase four classical music CDs at 25 dollars (CND) per CD. They would have ten minutes per interface to make their selections. They could take the full ten minutes or less. The evaluator would tell them when they had reached five minutes and two minutes left, and then when time was up. Each interface also used a separate window to represent participants' selections (Figure 2, right pane). The window allowed the user to see at any time how many selections they had and what the selections were. The participant could also at any time delete a selection from the window, thus the user could add many selections throughout a trial and then delete the ones they did not want to keep. All UI actions in both the test interface and the selection window were logged automatically so that we could assess where participants spent their time in the interfaces, how much of the data was explored and so on. As well as being logged, the trials were videotaped for reference.

5.1.6 Training.

Participants were given training time on each interface condition before each formal trial with that interface. The training set was in pop rather than classical music, to avoid users becoming familiar with the data set before the formal trial. A maximum time of 10 minutes was set aside for each training period, but users generally indicated that they were ready to proceed to the trial in half that time. To reduce learning effect in the actual trials, the software randomly generated a set of pieces for each interface so that the user would not encounter the same music for each trial. A data set of approximately 400 individual pieces was available per trial.

5.1.7 Preliminary Survey and Post-Questionnaire.

We had one pretest qualification for participants: they had to characterize themselves as daily users of a Web browser and email client. Since we wanted to consider a range of computer-expertise levels, we did not need a higher computer literacy requirement. We tested this qualification in our pilot study,

which assured us that the users who identified themselves this way would be comfortable learning these interfaces.

The preliminary questionnaire collected basic demographic information such as gender and age range, while the remaining questions asked about any formal musical training the participant may have had, the size of their personal music collection, and what amount of those recordings were classical music. The post questionnaire asked about which interface used the participant preferred, which they most disliked and what were the best and worst attributes of both these interface choices. There was a brief section for further comments. The evaluators also asked the participants after each trial if they would like to make any other comments verbally. The evaluator recorded these comments.
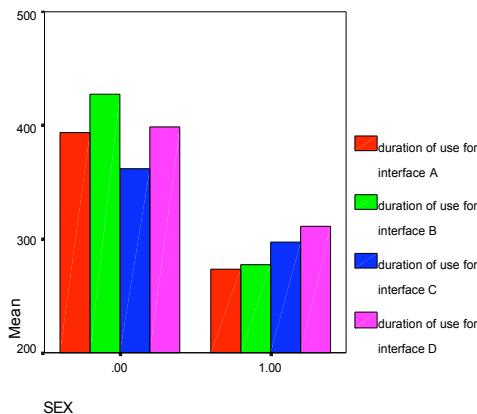
6 RESULTS

6.1 Quantitative

The most preferred interface style was the spatial layout. The Label/No Label condition had no effect on this preference. Neither did the Early/Late audio condition (Table 1).

Interface ∨	Pref ->	Most	Least
Spatial Early (SE)		21	0
Spatial Late (SL)		17	1
Temporal Early (TE)		1	20
Temporal Late (TL)		3	21

Table 1: Most and Least Preferred Interfaces

Graph 1: Duration per interface: Gender effect. Interfaces: A=SE B=SL C=TE D=TL

Duration overall that each user spent per interface was significant (Graph1): women spent less time overall per interface than men (F=6.776, p=.015). While 80% of participants had less than two years of any kind of formal music training, and forty percent of those had no training, there was no correlation among musical training, music ownership, or general level of education and either performance or preference.

Other than this, the general pattern was that users took longer in their first interface to make their selections and sped up throughout each successive trial. Similarly, there was no correlation between interface preference and duration of

use. Participants were as likely to use any of the interfaces longest (and to make the most actions in an interface), regardless of preference. Looking at the number of actions per interface, and time between adding selections, however, did yield significant results. Looking specifically at actions in the interface and times between adding selections, we see that participants who took longer to make their selections clicked/expanded more elements in the interfaces (Table 2).

|	SE	SL	TE	TL
clicks:	.45	.80	.62	.65
brushes	.67	.79	.73	.64

(all values are significant at $p < .01$)

Table 2: Brush and Click correlation per interface

There were significant effects of layout and early/late audio on how many actions participants took, both brushes and clicks, which can be seen in the graphs below. There were more actions taken in the spatial interfaces and in the late interfaces. There was a significant effect of condition only for clicking; people clicked more in the labeled condition.
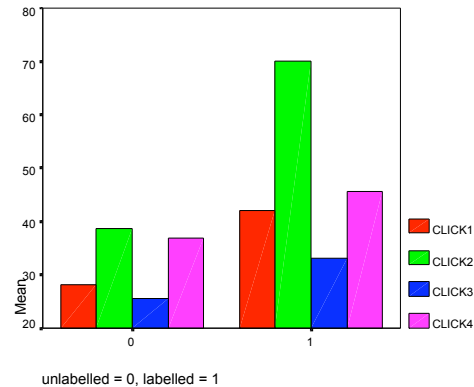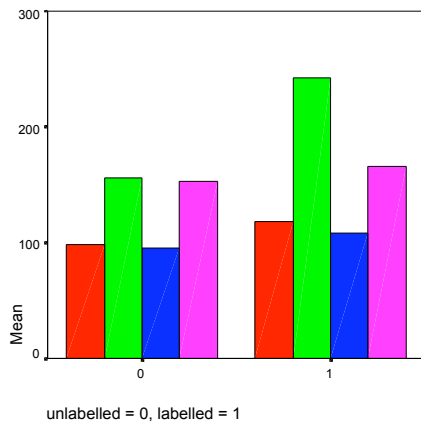
Graph 2 (left). Amount of Brush Actions per Interface   Graph 3 (right). Amount of Click Actions per interface
Color code: red = SE, green = SL, blue = TE, pink = TL
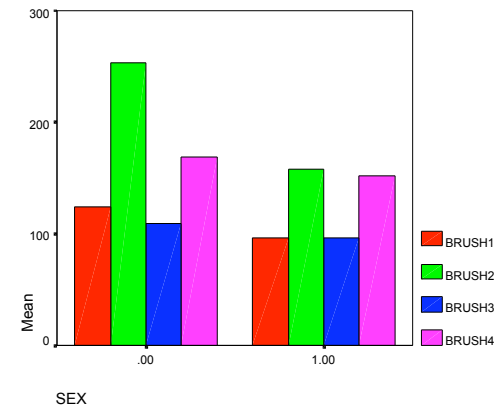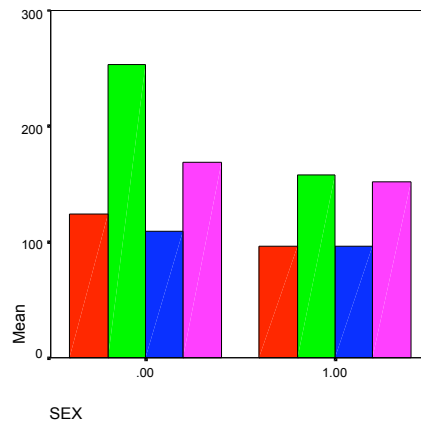
Graph 4 (left). Clicks/interface (0=men, 1= women) layout*gender: F=9.885, p=.004
Graph 5 (right). Brushes/interface (0=men, 1= women), layout*gender: F=6.575, p=.016

There was one gender effect for actions, a layout*gender interaction: men clicked and brushed more than women in the spatial interfaces, and in fact women clicked and brushed about equally in both layouts (Graphs 4 and 5). There was a significant difference between the early and late temporal interfaces for how much people used the "back" action (F=8.276, p=.008). It was used more in the "late" interface. Participants also used back more in the labeled condition, but this effect was not quite strong enough to be significant ($p = .11$). The "forward" command was used too infrequently to be able to test any differences between conditions or interfaces.

In the temporal layouts there was a negative correlation between age and both duration of use and number of actions (clicks, brushes, adds), which was not present in the spatial layouts (Table 3).

	spatial	temporal
actions	-.08	-.45**
duration	-.15	-.34*

$*p < .05, **p < .01$

Table 3. Age correlation for Action/Duration in Temporal Layouts.

6.2 Qualitative

All participants reported that preview cues made the process of discovering music enjoyable. Many participants commented on how the preview cues made finding new music "easy." Comments like they "wished [a certain music store] used this to let shoppers find new tunes," or "I want to take this [software] home and use it," were common. Evaluators noted that participants were frequently reluctant to stop playing with the spatial interfaces in particular, checking for "new" pieces to audition. Participants who had no previous experience of this domain, and reported having had "no way" of accessing it before, reported that they discovered new music to enjoy.

Participant comments made clear that they did not enjoy the temporal interfaces. Participants suggested several problems with this approach, but the main one was perceived lack of context: despite the fact that the path to their current location in the hierarchy was consistently available at the top of the interface window – an attribute brought to participants' attention repeatedly through the training process – many users reported that they felt lost. Comments like "I didn't know where I was…I couldn't see the whole path" were common. However, in all cases, users said they preferred having audio cues available.

The results show a higher percentage of users preferred the Early Spatial interface to the Late Spatial version; user comments indicate what were perceived as the strengths and weaknesses in each of these approaches. Users who preferred the Early Spatial interface recorded enjoying being able to get a sense of each area before exploring it. Similarly, these participants did not like that in the Late version, a brushed over piece kept playing until another brushed-over piece was selected. This meant that the last brushed over piece would continue to play while they explored other areas earlier in the hierarchy. As one participant put it "I went to somewhere else [in the interface] because I wanted it [the brushed over piece] to stop! That's why I was looking for a new tune!!"

Users who preferred the Late Spatial interface responded in almost the opposite fashion, seeing as a plus the continued playing of the last brushed piece. These

comments were similar whether participants were part of the Label or No Label trials. For these users, the preview cues reportedly got in the way of their browsing through the hierarchy: "they were irritating." Though these participants also reported enjoying the immediacy of the preview cues at the piece level. Indeed, they deliberately brushed over a piece label to keep that piece playing while they returned to other parts of the hierarchy to make new selections. "I like to multitask," wrote one participant, "I wanted the music I picked to keep playing while I looked for something else to try." For those who preferred the Late Spatial, however, they also reported that a plus of the Early Spatial condition allowed them to encounter music serendipitously.

7 ANALYSIS

7.1 Preview Cues and Spatial Layouts

That users brushed less in the Late condition than the Early is not surprising: brushing in the Late condition only triggered a preview cue in the final level of the hierarchy. That users clicked less in the Early versions, however, and indeed, spent less time between selection additions, suggests that the preview cues were working as designed: they allowed participants to assess *quickly*[14][i] whether or not they *wanted* to click/expand/explore an area of the domain.

While 50% of participants preferred the Early Spatial, 40% preferred the Late spatial. This is a curious finding, since the only information available by clicking through the hierarchy in the Late condition were names or terms that "meant nothing" to the majority of those participants. Such users seemingly preferred to click/hack randomly through the rest of the interface paths just to get to brushing over a set of pieces at the end of the path. This approach seemed counter-intuitive to those of us running the study, but is somewhat clarified by these participants' post-evaluation questionnaires. Many who preferred the Late Spatial condition suggested that they would have preferred the Early Spatial if they could have controlled two effects: (1) *when* or *whether* a preview cue played and (2) that they could make a preview cue continue to play, uninterrupted, as a selection. This suggests that (1) preview cues effectively enable explorations of unfamiliar domains, but that (2) for preview cues to be strongly effective, users need to be able to control when and how preview cues are available.

7.2 Label vs. No Label

The data suggests that persistent labels are not essential for an effective user experience in exploring less familiar data domains when other cues are available. This finding has potential implications for interface design of small or no-screen devices where text labels are not practical. In such cases, small graphic markers with audio cues may be an effective means for facilitating exploration of large data spaces, heretofore inaccessible on these devices. In a no screen device, for example, a car stereo, Earcons may act in place of graphic markers with voice over in lieu of mouse-over text on one channel and Preview Cue on another.

7.3 Gender Effect and Task Focus

It is not clear at this point how the gender effect on overall shorter use and overall fewer clicks in each interface by women relative to men may be interpreted or leveraged at this time. It does seem to be a finding in concert with

---

[i] Humans require only 250ms of audio to determine genre. A similar effect may operate for preference.

studies of women and computer use that show women spend less time "playing" with computers than men, treating them more frequently as tools rather than toys [4], [10]. Thus, even though women reported enjoying using the interface, they gave themselves perhaps less opportunity to go off task in making their selections as expeditiously as possible.

8 CONCLUSIONS AND FUTURE WORK

8.1 Conclusions

This paper presents the findings on the use of audio preview cues for exploration of a structured hierarchical representation of the classical music domain. We tested audio preview cues comparing several factors of common online music information clients like the Web (Table 4, below)

Within Sub	**Common**	**Comparison**
Layout	Temporal	Spatial
Audio	End of Path	Throughout Path
Bet. Sub		
Instance	Label	No Label

Table 4. Variables Tested

The study tested two hypotheses: that audio preview cues would improve effectiveness and efficiency of standard, temporal Web-like presentations of audio, and that a spatial layout with audio preview cues would significantly improve user experience for exploring the domain. With respect to our first hypothesis, from our experiment, we can see that participants brushed/explored more in the temporal view when cues were available throughout, than when cues were only available at the end of the path. Clicks, which switched to the next level of the hierarchy, occurred less frequently in the Early condition. This suggests that users moved forward in the hierarchy with more intent based on a deliberate choice.

With respect to the second hypothesis, the spatial layout was significantly preferred over the temporal layout. We also saw a significant negative effect of age with the temporal layout that was not evident at all with the spatial layout. The use of preview cues with spatial layout of the data hierarchy to be explored does significantly enhance both reported user experience and recorded efficiency of exploration into the hierarchy. Analysis of the results also suggests that giving the user specific control over when and how preview cues are available can further improve the efficacy of preview cues.

Several design heuristics fall out from the work.

- exploration of structured domains representing music can be improved by adding preview cues to the elements of the domain, whether the hierarchy is represented temporally or spatially
- this effect can be significantly enhanced if a spatial layout is use*d*.
- the negative effect between age and temporal representation of hierarchy can be nullified by using a spatial layout.
- while the majority of users prefer having cues persistently available to turn "on" by brush over, it is advisable to offer the user the option to turn cues on or off

8.2 Future Work

While we can see that audio preview cues enhance domain exploration, we also wish to investigate how preview cues may support implicit learning about elements in a domain. For instance, do users begin to build mental models about domain elements based on repeated exposure to preview cues; are these models sufficiently correct to build user expertise, so that users could begin to describe the characteristics of an attribute based on selecting and hearing many cases of these kinds of instances?

We also wish to consider offering multiple preview cues, in order to determine if one cue may be too coarse a generalization, or similarly, whether more cues are more effective or more distracting in supporting exploration for the domain-naïve user. We have developed a revised classical music prototype which offers three cues rather than one. An instantiation of this is shown in Figure 3, below. Each label has now three explicit cues associated with it for previewing. This approach also poses interesting algorithmic questions about how to best automate the choice of candidate samples for an area of a domain. This question becomes even more critical when considering preview cues for other domains than music.
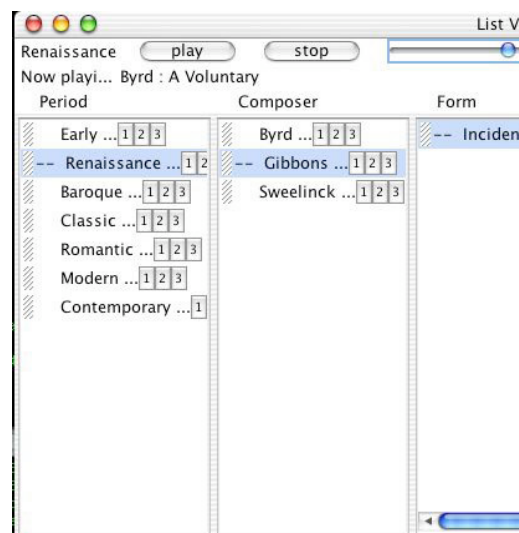

Figure 3. Multiple preview cues offered per label.

We have begun to investigate a generalized model of preview cues for assisting exploration of other kinds of data spaces. How might cues be designed to afford a person exploring JavaScript information about whether a particular area (WindowEvent vs MouseEvent for instance) might be a good place to look to help solve a status bar problem? What might a cue look like to let a health consumer choose to explore one genre of treatments over another for breast cancer *before* digging into those areas of an information space?

Our evaluation of audio preview cues of domain areas have shown that this light weight mechanism assist evaluating possible domain paths is an effective tool, both to improve exploration efficiency and to enhance the users' exploration experience, improving what Dillon refers to as "positive affect." We look forward to investigating the potential for domain preview cues in other contexts.

9 ACKNOWLEDGMENTS

10 REFERENCES

1.   Amento, B., Terveen, L., Hill, W. and Hix, D., Topicshop, Enhanced Support for Evaluating and Organizing Collections of Web Sites. Proc. Uist 2000, 201–209. In *Proc. UIST 2000*, (2000), 201–209.
2.   Boynick, M. Classical Music Pages, 1996.
3.   Brewster, S.A., Wright, P., Edwards,, A.D.N. An Evaluation of Earcons for Use in Auditory Human-Computer Interfaces. In *Proc of CHI '93*, (1993), 222-227.
4.   Cherny, L., Weise, E.R., eds. (ed.), *Wired Women: Gender and New Realities in Cyberspace*. Seal Press, NY, 1996.
5.   Chimera, R. and Shneiderman, B., An Exploratory Evaluation of Three Interfaces for Browsing Large Hierarchical Tables of Contents. In *ACM Transactions on Information Systems (TOIS)*, (1994), 12.
6.   Czerwinski, M., van Dantzich, M., Robertson, G.G. and Hoffman, H., The Contribution of Thumbnail Image, Mouse-over Text and Spatial Location Memory to Web Page Retrieval in 3d. In *Proc. of Interact '99*, (Edinburgh, Scotland, 1999), 163-170.
7.   Dunn, J.W., Mayer, C.A., Variations: A Digital Music Library System at Indiana University. In *Proc of DL '99*, (August 1999), 12-19.
8.   Group, M.R.S.C. Stacks Photo Browser, 2002.
9.   Hochheiser, H., Browsers with Changing Parts: A Catalog Explorer for Philip Glass' Website. In *Proc. on Designing interactive systems.*, (2000), 105-115.
10.   Margolis, J. and Fisher, A. *Unlocking the Clubhouse: Women in Computing*. MIT Press, Mass., USA, 2001.
11.   Risden, K., Czerwinski, M.P., Munzner, T. and Cook, D.B. An Initial Examination of Ease of Use for 2d and 3d Information Visualizations of Web Content. *IJHCS*, *53*.
12.   Terry, M. and Mynatt, E.D., Side Views: Persistent, on-Demand Previews for Open-Ended Tasks. In *UIST*, (2002), 71 - 80
13.   Thomason, R. (ed.), *Formal Philosophy, Selected Papers of R. Montague*. Yale University Press, 1974.
14.   Tzanetakis, G., Automatic Musical Genre Classification of Audio Signals. In *Proc of ISMIR 2001*, (2001).