Motion in Casual InfoVis and the Interrelation Between Personality, Performance, and Preference

David W. Sprague and Melanie Tory

Abstract—Current real-world casual information visualization systems use motion to presumably increase system appeal and repeated use. Despite empirical evidence suggesting that motion is distracting and not particularly good for data encoding, its continued use may indicate benefits independent of traditional utility measures. We present two experiments designed to highlight differences and disparities between task performance, perceived system utility, and subject preference for static and moving glyphs. A music library visualization provided a casual, practical, nuanced, and pertinent domain application. Song glyphs represented beats per minute and beat strength using a line encircling the glyph. Beat encodings used one of four methods: a static sinusoidal pattern (static), an oscillating movement pattern (motion), an oscillating sinusoidal line (redundant), and a sinusoidal line that moves to the currently playing song (extraneous). Subjects first performed a pair-wise glyph comparison task to identify conditional differences in speed and accuracy. Then, a more realistic playlist generation task was performed using these beat encodings to explore opinions and behavioral patterns using a qualitative approach. Our results show clear performance decreases associated with motion use in the comparison task. Despite this, more than half the subjects preferred a moving glyph to the static encoding. We identified three groups of subjects (utility, fun, and intuitive) where subjects were consistent within groups but distinctly different between groups. This suggests that casual visualizations could be customizable based on high level groupings, where pragmatism can be emphasized for some users and aesthetics and style can be enhanced for others.

Index Terms—Motion, casual InfoVis, qualitative analysis, music library

1 Introduction

What is the relationship between task performance, system appeal and motivation in casual information visualization (*casual InfoVis*) and how do these factors differ between people? For many casual InfoVis systems such as the digg lab visualizations by Stamen Design [7], the 'fun' factor and aesthetics are critical for adoption by some users [14]. This is in contrast to professional data analysts and scientists who may put more emphasis on productivity. After all, improving a visualization's utility makes it a better tool. We are primarily concerned with the more pragmatic subset of casual InfoVis systems where users are personally motivated to perform data analysis and exploration tasks rather than motivated by a formal job description or financial incentives. In these scenarios, users are not forced to use any particular system and thus aesthetics, fun, and even the perception of productivity could greatly influence initial program use, user motivation and system adoption. Perhaps for this reason, many 'real world' visualizations like the digg visualizations [7] use motion to increase appeal, possibly at the expense of performance. Motion use may have a cost, but the ultimate goal is to create a system people want to use. The complex trade-offs between user motivation and performance merit further research.

We define motion as the perceived sequence of changes in a visualization's appearance over time. This may include icon movements, transitions between system states, and colour changes [18]. We define *extraneous motions* in InfoVis as any motion that does not encode data directly used in a task. This includes icons moving randomly [7], unnecessary movement during an animated transition, and icons that change colour or shape for aesthetic purposes. *Motion based data encoding* refers to the use of temporal changes to encode relevant information [22]. Our research primarily focuses on extraneous motion and motion based data encoding because there are no clear performance benefits, yet motion should affect aesthetics and it is commonly found in real world systems.

In the current two experiments, we examine the use of motion to both encode data and for aesthetic purposes. In particular, we examined three factors. First, we studied whether using a pulsing motion can successfully be used to encode beat rate (tempo) and beat strength information. Second, we measured whether redundantly encoding information using motion offers performance or preference benefits. Third, we explored whether extraneous motion used to improve aesthetic appeal, system novelty, or fun decreased user performance metrics. We chose to study motion due to its appeal, its frequent use in real world casual InfoVis applications, and its documented drawbacks such as distraction (for utility/appeal trade-offs)[22].

Most of the previous motion research has been performed in highly controlled environments, independent of the final proposed application domain. We continue this tradition in our first experiment by conducting a pair-wise glyph discrimination task to see how motion affects data readability. This experiment's primary research question is: *what are the quantifiable costs associated with different styles of motion encoding (motion, redundant, and extraneous)?*

The second experiment used a qualitative approach to observe subjects performing an appropriate 'casual' data analysis task. This study was also designed to examine complex user-system interactions with cost/benefit trade-offs. A controlled quantitative experiment would not adequately address the complexity of our research concerns. Ultimately, a music library visualization was used since tasks were practical and motion should semantically map to a song's beat. Motion encoding was designed to be intuitive and practical so that its use would provide potential advantages over static beat encodings. We examined how participants conducted a playlist generation task to see how beat encodings and tasks affected participant opinions, preferences, and task performances. Trade-offs between task performance and system appeal were examined using interviews. This experiment's two primary research questions are:

1. *Is motion more enticing to users than static encodings?*
2. *What factors influence user preferences and how?*

2 Previous Literature

2.1 Benefits and Drawbacks of Motion

Motion's attention attracting and transitional nature have made readability and preattentive perception key factors influencing its use in

- *David W. Sprague is with the University of Victoria, E-mail: dsprague@cs.uvic.ca.*
- *Melanie Tory is with the University of Victoria, E-mail: mtory@cs.uvic.ca.*

visualizations. Despite earlier claims that motion facilitates learning, Tversky et al. [18] warn that static and animating conditions were not equivalent in most of these experiments and many proposed motion benefits to learning may not be justified. They further warn that for animation to be useful, the motion must map well to the data (congruence principle) and it must be readily and accurately perceived and understood (apprehension principle). Motion in our experiments is designed to address both these principles. Ware and Bobrow [22] reported the perceptual thresholds for differentiating circular oscillating points, thus quantifying motion readability. Oscillation frequency, amplitude, and phases could all be used to differentiate points, with subjects being most sensitive to relative phase patterns and least sensitive to oscillation frequency differences. Motion is perceived preattentively, meaning that identification of a moving object within a person's field of view does not rely on conscious recognition [3]. Motion in the periphery also triggers an 'orienting response,' attracting the user's focus unless consciously ignored [4]. Focusing on a task and ignoring movement thus requires an increase in cognitive load.

Motion may also indirectly affect performance through aesthetics. Motion can be visually appealing to many users resulting in potential increases in system use. The halo effect states that aesthetically pleasing interfaces are described in more positive terms, even if the positive trait has nothing to do with the interface [1]. Norman's [13] three processes of human judgment suggests that behavioral evaluations involving logic and utility (behavioral) are influenced by previous experiences (reflective) and aesthetics (visceral). We therefore expect motion's evocative nature to improve perceived usability despite possibly being distracting. Our experiments are designed to help clarify the fun / functionality trade-offs affecting user opinion.

2.2 Uses of Motion

Motion is traditionally used in visualizations for two main purposes: to alert users [4, 8], and to provide animated data transitions [5, 10, 16]. However, we will limit our discussion to motion data encoding and extraneous motion, which are our primary foci.

Motion data encoding research has been limited due to well documented distraction concerns [3]. Instead, motion based data encoding is normally used temporarily as a means of data highlighting. The preattentive perception of motion makes it effective for filtering and brushing subgraphs in node-link diagrams, thus encoding subgraph relationships [21]. Ware and Bobrow [22] also suggest that static data encoding may be more efficient for displaying information in the user's foci since there is no need to temporally track a glyph's state. More recently, Robertson et al. [15] have shown motion in trend visualizations to be effective and highly appealing for demonstrations but ineffective for analysis tasks. Bartram et al.[4] demonstrate the use of motion encoding for glyphs but suggest that motion should be used sparingly because of distraction concerns. For a brief review of motion based encoding methods see Ware and Bobrow [22]. Our current work extends motion based data encoding by examining its use in casual InfoVis and to what we believe to be a rare, semantically appropriate application domain for its use: music visualizations.

Extraneous motion has been used to improve system appeal. For example, the Vizster Friendster visualization uses randomized node movements to increase system appeal and reinforce the system's informal nature [9]. Similarly, the PartyVote visualization uses animating icons to increase the system's sense of fun [17]. Song icons in the Burst Labs visualization [6] drift constantly within an enclosure seemingly for stylistic reasons alone. Finally, the digg visualizations [7] consistently use motion when representing recent digg articles and their popularity. These movements attract attention, indicate changes in content popularity, and may add a playfulness and style to the interface. However, it is currently unknown whether motion truly improves system appeal, how substantially motion affects data readability, and how motion based distractions affect user opinions. Our experiments attempt to address these unknowns.

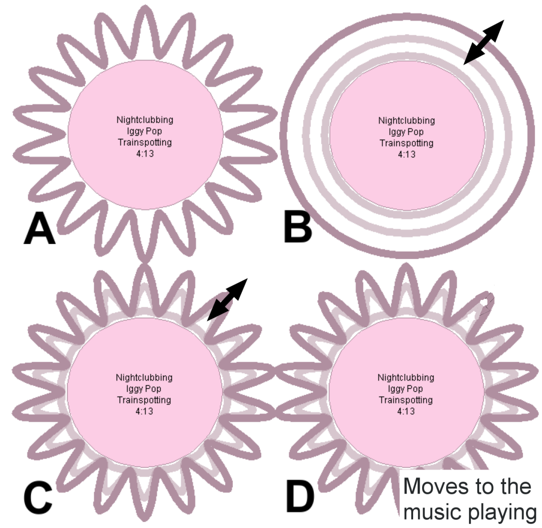

Fig. 1. Song glyph representing Night Clubbing by Iggy Pop. Beats per minute and beat strength encodings differ between conditions: A) static, B) motion, C) redundant, and D) extraneous. Faded lines indicate movement paths.

3 Experiment 1: Glyph Comparison Task

3.1 Methodology

We conducted a traditional forced decision task using a within subjects design with four beat encoding conditions. For consistency with the playlist generation experiment (described later), each glyph represented a song with song tempo (beats per minute or *bpm*) and intensity (beat strength or *BStr*) being mapped to an animating pattern or a sinusoidal line pattern (see Figure 1). Conditions were:

- **Static Encoding:** A sinusoidal line encompassing the glyph represents tempo using the number of line oscillations and beat strength using amplitude
- **Motion Encoding:** an expand/contract animation of a circle encompassing the glyph represents tempo using the motion frequency and beat strength using movement size or amplitude. All movements oscillate at the same frequency as the beat they represent and follow a sinusoidal movement rate. Thus the motion encoding is a temporal equivalent of the sinusoidal line curve.
- **Redundant Encoding:** Tempo information was encoded redundantly using a sinusoidal line that expands and contracts the same as the motion encoding.
- **Extraneous Motion Encoding:** The static encoding method was used to represent the glyph's beat information with the line animating with the beat of the currently playing song. Thus, when a song was playing the glyphs all appeared to 'dance' to the music.

Using a sinusoidal line enables us to keep conditions almost identical except for the trait being manipulated. The movement's repetition should also address the apprehension principle[18]. Sound has traditionally been mapped to sinusoidal wave patterns and to repeated movement. The motion encoding should provide an appropriate cognitive mapping to a beat's regular pattern (particularly when motion and beat are synchronized) [12, 2]. Thus we expected the encoding methods to be more intuitive than changes in colour or luminosity. Bartram et al's [4] work on motion and distraction suggests that growing and shrinking motions used will be less distracting to users than translation, rotation, blinking, or explosive motions.

3.1.1 Subjects

Sixteen subjects aged 19-39 ($\bar{x}$: 25.33, σ = 7.9, 6 female, 10 male) were recruited. Nine undergraduate students, six graduate students, and one professional participated. Subjects were financially reimbursed for their time.

3.1.2 Procedure

The task consisted of subjects performing a series of timed pair-wise glyph comparisons to evaluate performance differences between conditions. The test application was written in Java and had a refresh rate of 40Hz for all conditions. Tasks were conducted on a 1.66 GHz Intel Centrino Duo Core laptop with Windows XP and 1GB of RAM.

Trials followed a 4 (encoding condition) $\times$ 2 (question type) $\times$ 3 (value) $\times$ 2 (distraction) design. Each configuration was repeated three times. Conditions orders were counter-balanced for all tasks using a Latin squares design, which repeated every four subjects. For each trial, a question was provided, users pressed the space bar to start the trial, and after five seconds without a decision, the beat encoding line was removed to force a decision. The keyboard arrow keys were used to choose between the left and right glyphs. Subjects were asked to answer as quickly as possible while maintaining a greater than 70% response correctness (displayed at the end of each condition based set of trials). Glyphs had no text, were colour coded by question type, and glyph bodies were approximately 5 cm in diameter. Question data was synthetically generated a priori based on beat data ranges from the library used in experiment two. Short breaks were provided between sets. Experimental factors were:

- **Encoding**: static, motion, redundant, or extraneous beat encodings. Extraneous motion rates were constant.
- **Question Type**: Two questions were asked: "Which song has the fastest tempo (beats per minute)?" and "Which song has the stronger beat?"
- **Value**: The difference in beat data between glyph pairs was held constant at 31 for bpm and 3.37 for BStr questions. Beat values between trials varied between high ($BPM : 184/153, BStr : 16.06/12.69$), medium ($BPM : 113/144, BStr : 11.85/8.48$), and low ($BPM : 73/104, BStr : 7.63/4.26$).
- **Distraction**: For each question type, the other beat information dimension was either identical between glyphs or differed by the pair-wise beat difference (31 or 3.37).

Trials were grouped by condition and question type with bpm and BStr questions randomly chosen to start. Value and distraction factors were randomized within a subset. Beat differences were made sufficiently large (31 bpm) to avoid minor frequency differentiations (which are difficult to distinguish [22]). Glyph readability was measured using task correctness and task time. A training period was provided at the start of the task, where subjects were required to correctly answer four beat strength questions and four bpm questions with each encoding method.

3.1.3 Hypothesis

Beat data will be read most effectively using static encodings, followed by redundant and extraneous encodings.
Static curves can be read immediately and do not require monitoring over a period of time like motion encoding. Movement phase issues are also irrelevant with the curve encoding.

3.2 Results

Experimental results confirm our hypothesis. Significant differences in trial time ($F_{TimeBPMCond}(3,39) = 13.282$, $p < .001$, $F_{TimeBStrCond}(1.603, 6.451) = 20.844$, $p < .001$) and task correctness ($F_{IsCorrectBPMCond}(3,39) = 15.713$, $p < .001$) were identified between encoding conditions, with post-hoc analyses showing static data encoding being significantly faster than the motion ($p < .001$), redundant ($p < .001$) and extraneous ($p < .018$) conditions (see Figure 2). Static encoding was significantly more accurate than the moving conditions ($p < .001$, $p < .023$, $p < .012$ for motion, redundant and extraneous respectively). Motion encoding was also significantly slower than the extraneous condition ($p < .031$) and less accurate than both the extraneous ($p < .002$) and redundant conditions ($p < .001$). The presence of a distraction factor was also shown to increase trial time ($F_{TimeBStrDistr}(1,13) = 13.676$, $p < .002$ and $F_{TimeBPMDistr}(1,13) = 15.344$, $p < .001$) and decrease task correctness ($F_{IsCorrectBStrDistr}(1,13) = 5.194$, $p < .038$) (see Figure 3).

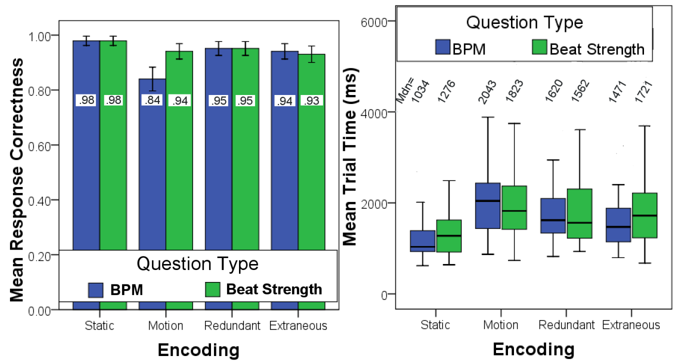

Fig. 2. Glyph comparison task results showing the interaction between question type and conditions on mean trial time (milliseconds) and correctness (%/100). Error bars indicate 2 standard errors.

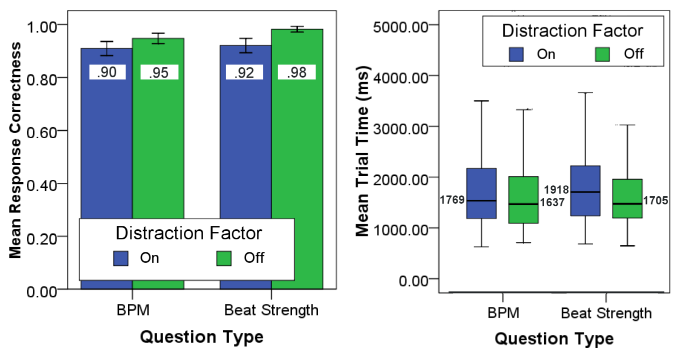

Fig. 3. Glyph comparison task results showing the interaction between the presence of a distraction factor and the question type on mean trial time (milliseconds) and correctness (%/100). Error bars indicate 2 standard errors.

3.3 Discussion

The glyph comparison task results clearly show that static encoding was significantly faster and more accurate than the three motion conditions, particularly when distraction factors were involved. We believe that the stability of static encodings simplified data reading, thus improving task scores. Alternately, the motion in the extraneous and redundant conditions may have distracted subjects from reading the sinusoidal line. Subjects during the redundant trials may have chosen to read the data using the motion encoding much of the time rather than use the sinusoidal curve. The ephemeral nature of motion means that glyph evaluations must be timed and likely require working memory [22]. Task performance using the motion encoding was significantly worse than redundant and extraneous encodings, suggesting that reading data encoded in motion is more difficult than reading the curve encodings (even when animating). This suggests that it is better to have motion encode nothing and act as a distraction than to have it solely encode information. Although it could be argued that another motion encoding method might be more effective, this seems unlikely

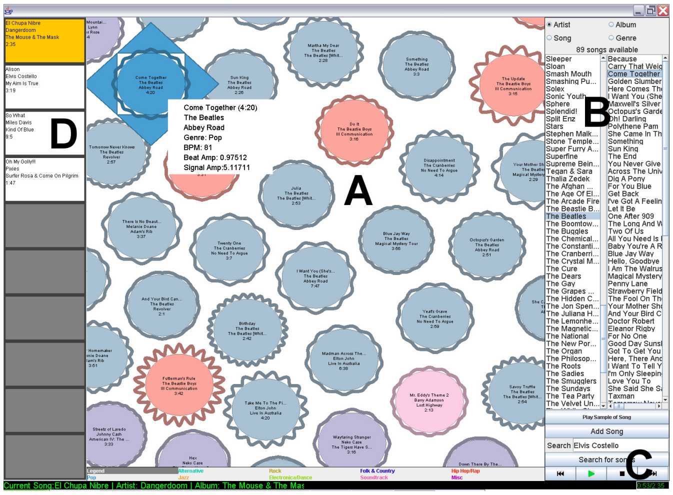

Fig. 4. The visualization used during the playlist creation tasks. Features include A) the zoomable genre coloured music map, B) an 'iTunes like' navigation window, C) music controls, and D) the current music playlist.

based on previous literature [4, 22]. Conversely, the static beat encoding could potentially be improved since it was designed to be intuitive and analogous to the motion, not to maximize readability.

The significant effect of distraction suggests that encoding multiple dimensions of information into motion can lead to problems. The distraction effect on BStr reading is likely caused by the phase discrepancies between glyphs such that the peak movement positions are not simultaneously reached for the two glyphs and visual memory is required to respond. Beat strength distractions may affect bpm questions since pixel movement rates increase with beat strength. These data demonstrate that encoding more than one piece of information using motion can hinder data reading, even when the motions are considered conceptually orthogonal.

4 Experiment 2: Playlist Generation Task

Our second experiment was a qualitative study examining the tradeoffs between aesthetics and performance, and how these affect user preferences. We chose to examine subjects performing a playlist generation task in a custom made music visualization, as this task seemed sufficiently complex, realistic, and open-ended to examine user motivation and thought processes during the follow up interviews. We also wanted to examine casual visualization use under normal circumstances. We focused on exploring behavioral patterns using semi-structured interviews, observations, and recorded user actions, using a method similar to grounded evaluation. See work by Isenberg et al. [11] for a more thorough discussion of grounded evaluation.

4.1 Methodology

The playlist generation experiment was conducted with each subject during the same two hour session as experiment 1 and involved the same four glyphs conditions. Characteristics unique to this experiment are described below.

4.1.1 Software and Equipment

The music library visualization was written in Java and displayed music glyphs on a zoomable two dimensional (2D) music layout which we call a *music map* (see Figure 4). The system provided an editable playlist and traditional 'iTunes like' navigation. Music came from a personal music collection. The text based navigation was only enabled during training to ensure the music map was used during testing. The applications updated at 30Hz to ensure consistency between conditions but update rates occasionally decreased when zooming occurred. Built in speakers were used for music playback and a USB optical wheel mouse was used to pan (click and drag) and zoom (mouse wheel).

Music beat analyses and the music map layout were performed in advance. The final 1432 songs were chosen to be as diverse as possible while emphasizing recognizable music like the Beatles. Human and computer beat identification errors are usually double or half the actual value [20]. We filtered out most songs not in 2/4 or 4/4 time so disparities between the perceived beat and computed beat would remain multiples of each other, limiting confusion. The Euclidean distances between each pair of songs were calculated, with both song metadata and sound similarity metrics, using the same formulation as Sprague et al. [17]. MDSteer [23] then performed multidimensional scaling to generate a two dimensional layout for these vectors. Glyph positions were programmatically scaled and shifted with user guidance to ensure that they did not overlap. Marsyas' bextract application [19] calculated song beat information for each song in the music library. For each song, beat measures were validated and adjusted manually when the computed result was clearly incorrect.

4.1.2 Procedure

The experiment had three stages: a training stage, the *playlist generation tasks*, and a semi-structured interview at the end. At the start of each subject's session, system functionality and the four beat encoding techniques were demonstrated. Subjects were asked to experiment with all system functionality, explore the music library, and locate familiar songs for eight minutes. Subjects were reminded about unused system functions in the last 2-3 minutes of practice. The playlist creation tasks required subjects to find music that they felt was appropriate for a given situation like a romantic evening or a party. The experimenter monitored subjects from another room using a video camera, a feed of the laptop screen and an area microphone; all of which were recorded. Condition orders were the same as experiment 1. The semi-structured interview was conducted for approximately a half hour. Subjects gave background information, their opinions about the playlists they created, and their glyph preferences.

Playlist Generation Task: Subjects were asked to imagine they were looking through a friend's music collection to create a list of songs for four different occasions: a romantic evening, a party with 20 attendees and perhaps some dancing, studying with a personal music player, and exercising with a personal music player. Subjects were allowed to modify these tasks enough for the activity to be realistic for them. For example, exercise music was chosen to be appropriate to the exercises subjects performed. Occasion task orderings were consistent across subjects but condition order varied. Tasks were intended to explore how beat information was used for finding slow music (studying & romantic) versus fast music (party & exercise) and group music (romantic music & party) versus personal music (exercising & studying). Each subject was given seven minutes to create a playlist, however, it was stressed that the number of songs found was unimportant and subjects were to approach the problem the way they normally would at a friend's house. Participants were informed that playlist music would start automatically as an added incentive to choose music they liked. Subjects could control the music volume as long as it remained audible. Beyond this, no formal instructions about how to perform the task were provided, as subject approaches were deemed as important as how conditions affected performance. We also wanted behaviors to be as natural as possible.

Despite using a qualitative approach for the playlist generation task, some metrics were collected to explore the relationship between user action and opinion. Subjects ranked the four glyph encodings and indicated their music collection size. We also recorded the frequency at which a subject sampled songs before adding them to the playlist. We use this as an indirect measure of music novelty as familiar music would not need to be sampled first. Subjects were asked about their song style search criteria for each task to identify when beat was important. They were also asked whether they used the beat encodings to find music. We consider cases where beat was important but the beat encodings were unused to be a *beat contradiction*. This factor shows that the subject did not use the tools provided when they may have helped.

4.2 Guiding Questions

Qualitative evaluations are grounded in the process of pattern discovery rather than hypothesis testing. As such, our study's design was built around two guiding questions:

Q1: How does the presence of extraneous motion and motion based data encoding affect system appeal? Given the use of extraneous motion in commercial visualizations, it is likely aesthetically appealing to many users. However, the extent of this appeal and how it interrelates with task performance is still unknown.

Q2: How do productivity and utility affect system expectations and opinions about casual InfoVis? We hope to discover how user preferences are affected by utility and productivity during a casual task, and whether 'fun' and aesthetic trade-offs with productivity are observable.

4.3 Results

As expected, user reactions to the various visualization conditions were varied. However, users evaluations and comments appeared to fit into three main categories, suggesting that system appreciation is highly dependent on previous experiences and personality. These groups are discussed in the Subject Groups section. We also found that subject strategies and approaches to music selection were complex and varied, despite most subjects pursuing a similar goal (discussed in the Music Selection Approaches section below).

Every condition ranked highest for at least two subjects and lowest for at least two others. The static encoding was generally preferred; ranked first by six subjects. However, six other subjects rated static encoding as one of their least favorite, despite knowing they performed better using it during the comparison task. This result addresses Q2: productivity and utility were not the only factors determining preference for all subjects, however, it was considered very important for half of them. Beat use was associated with the playlist type, however, it was not associated with encoding condition. If motion was substantially more appealing than the static encodings, we expected beat information to be utilized more during the three motion conditions (Q1). We found that the answers to our research questions corresponded greatly with the subject groups, discussed below.

4.3.1 Subject Groups

Consistent patterns of behavior, opinion, and background were found between subject subsets. The recorded interviews were evaluated for common patterns and illustrative quotes. From these records, we found that subjects fell into three major groupings. These groupings corresponded to user comments about utility, fun, and how well conditions mapped to the user's cognitive model about music. Interestingly, these groupings also corresponded with music interests, desire to find new music, gender, and beat contradiction frequency. Each group will be discussed in turn. See Table 1 for a summary.

Utility: "If [motion] was an option, I would turn it off. It's something I don't need." (subject 2)
The eight subjects concerned about task performance, functionality, and utility (subjects 1, 2, 4, 5, 6, 7, 11, and 12) were exclusively male, the vast majority of subjects had a technical background (for example, studying computer science or engineering), all but one made explicit comments about the importance of productivity and system utility, and all but two had the static encoding as their favorite condition (these two thought redundant encoding improved their performance). Comments like "static is fine" (subject 4), "if animating then [I want] the useful animation" (subject 2), and "[motion was] more distracting than anything else" (subject 11) illustrate the general attitude about motion. Only subjects 2, 4, and 12 did not chose the motion condition as their least favorite condition. Utility subjects were significantly less likely to use the beat encodings during playlist creation (only used in 11 of 32 trials). Beat was considered an important criterion in 11 of 21 trials where beat encodings were ignored (beat contradictions). Seven of the eight subjects claimed to search for familiar music rather than discovering new music. For example, subject 11 said it was "easier to find things [he] recognized." Curiously, seven of the nine alternative music fans were utility subjects.

Fun: "Although [motion] was the most distracting, it was certainly the most fun." (subject 16)
The fun and/or aesthetics group (subjects 8, 10, 13, 14, and 16) typically mentioned the importance and appeal of beauty, fun, or aesthetics relating to the beat encodings . Subject 14 thought the static condition was "weird that it just stands still." Subject 13 similarly said "if [the glyphs are] moving, then I think it's more pretty." The static condition was not the favorite condition for any fun group members. Three of the five subjects were women, and the majority had a non-technical background. According to playlist and sampling data, fun group subjects were more likely to play a sample without adding it to the playlist than other subjects (Samples added: Fun group = 18.75% of 672, Others = 25.68% of 693), suggesting they might be finding more unfamiliar music. This is further supported by subject interviews where 4 of the 5 subjects who actively searched for new music were in this group. Beat contradictions were rare with only subject 14 claiming that tempo was important while ignoring beat information. Four of the six pop music fans were members of this group.

Intuitive: "[Motion] gives you the information more directly." (subject 3)
The intuitive group (subjects 3, 9, and 15) found that motion was a more direct, easier to read, or intuitive way to encode bpm information than the curved lines. All subjects were women. Subject 9 stated that motion encoding "kind of clicked" while subject 15 stated "if it's not moving, then I didn't pay attention to it.". The redundant condition was preferred by subjects 3 and 9, potentially suggesting that users thought the motion was intuitive but they still pragmatically wanted the static encoding to "fall back on." The intuitive group used beat encodings the most frequently, utilizing it in ten of their twelve trials. Only one beat contradiction was identified.

Beat Contradictions: Numerous utility subjects claimed beat was important for finding playlist songs, but they did not use the beat encoding information. This seeming contradiction was justified in six ways (based on interviews). Subjects 5, 6, 9, 11, and 14 believed that beat encoding was not needed for the current task, although this changed depending on the task. Subject 6 wished to see the actual beat information rather than an approximation so he used the mouse for pop-up beat information. Subjects 4, 11 and 12 believed that music choice could not be simplified to a couple key dimensions like beat. They did not trust beat information to adequately represent a song's style. Subjects 4 and 10 lost their trust in the beat encoding when several sampled songs did not match their expectations (despite being manually verified). Subject 4 expanded on his comments by saying "it's not right or wrong, but what I expect." Thus, trust issues for subjective data like music may remain, even if the data is verifiably accurate. Subjects 1 and 8 claimed to not use the beat encoding because they were zoomed out too far during the trials. Finally, subject 15 claimed that she initially forgot she could use the beat encodings.

4.3.2 Music Selection Approaches

Beat information provided a means of culling the song search space rather than as a selection criterion. In fact, limiting the song search space with genre, beat, and music map location information was one of the most popular tactics by all subjects. Song choices themselves were accomplished almost exclusively by sampling the song first or based on the subject's familiarity with the song. That said, we found beat information was only consciously used in half the trials, as some users though other song information, such as artist and genre, were more critical to their searches.

Subject decisions and approaches were more complex and nuanced than anticipated, despite the tasks being designed to encourage open approaches. Several subjects consciously looked for songs that were the opposite of the primary song style to take a planned break in their activity. For example, subjects 1 and 9 included cool down music for

Table 1. Subject groups and their distinguishing traits. Factors are (in order): gender, favorite condition, least favorite condition, the frequency a subject sampled a song but did not add it to the playlist, the number of times beat information was considered important but not used (beat contradictions), explicit claims a subject looked for familiar or new music, and the size of the subject's personal music collection.

Subject	Group	Gender	Fav. Cond.	Least Fav. Cond.	Sample Only Freq.	Beat Contra. (/4)	Find Familiar (/4)	Find New (/4)	Personal Music (Songs)
1	Util.	**Male**	Redun.	**Motion**	**Med.**	1	**1**	1	**4000**
2	Util.	**Male**	**Static**	**Extra.**	**Med.**	0	**1**	3	400
4	Util.	**Male**	**Static**	**Extra.**	**High**	1	**2**	1	**10000**
5	Util.	**Male**	Extra.	**Motion**	**Med.**	1	**1**	0	**20000**
6	Util.	**Male**	**Static**	**Motion**	**Med.**	3	0	0	**12000**
7	Util.	**Male**	**Static**	**Motion**	**Med.**	0	**4**	0	**2000**
11	Util.	**Male**	**Static**	**Motion**	**Med.**	3	**4**	0	500
12	Util.	**Male**	Motion	**Redun.**	Low	1	**1**	0	**5000**
8	Fun	Male	**Motion**	**Static**	**High**	1	0	0	**1500**
10	Fun	Female	**Redun.**	**Static**	**High**	**0**	1	**4**	**1500**
13	Fun	Female	**Motion**	Extra.	**High**	**0**	0	**4**	**500**
14	Fun	Female	**Redun.**	**Static**	**High**	3	0	**4**	**1400**
16	Fun	Male	**Extra.**	Redun.	Med.	**0**	2	**1**	6000
3	Intuit	**Female**	**Redun.**	Static	High	0	0	**0**	1000
9	Intuit	**Female**	**Redun.**	Motion	Low	1	4	**0**	2000
15	Intuit	**Female**	Static	Motion	Med.	0	1	**0**	2500

their exercise list. Subject 5 chose relaxed mellow music for the party music because this was appropriate for his group of friends.

Actions during the four playlist tasks seemed predominately influenced by whether the music was personal or shared with others. Thus, subjects tended to choose music they wanted (either new or familiar) for the study and exercise tasks, but their choices were consolatory for the romantic evening and party playlists. Romantic evening playlists predominately had people searching for songs about love, jazz, or slow music. For subjects 2, 4, 5, 9, 12, 14, and 15, romantic music was chosen based on what their partner enjoyed. Party music was typically hip-hop or rock and chosen to appease the subject's friends. Subject 10, in particular, chose electronic music her friends enjoyed, but she did not. Study playlists were perhaps the most musically diverse between subjects, however, everyone except subjects 11 and 16 actively sought music that would not distract them. Strategies included searching for music without lyrics, familiar music, unfamiliar music, and music the subject could not easily sing along with. Exercise music choices depended on the exercise chosen (ranging from weight lifting to snowboarding) but subjects predominately chose faster music or music with a stronger beat. As such, several subjects enjoyed using the beat encoding more for this task.

4.4 Discussion

The experimental results show that motion only improved system appeal for a subset of the subjects, while others were acutely concerned about system utility and efficiency. Furthermore, groups of users can frequently have conflicting desires and goals. Even within the context of a constrained experimental task with clearly formulated instructions, the individual freedoms associated with casual InfoVis resulted in extremely divergent results, opinions, and perceptions about system effectiveness. Trust in the effectiveness and validity of the visualized data also varied dramatically which seems to influence behaviors such as zooming and beat encoding use. Furthermore, previous experiences seem to affect perceptions of utility. Despite clear utility / fun trade-off evaluations by subjects, general opinions about the system appear more varied than expected. To address individual differences, visualization customizations may be beneficial but time costly. Provided subject groupings are possible, high level customizations based on group memberships could greatly improve user experiences.

4.4.1 Motion Appeal (Q1)

Our results suggest that motion was not as appealing to participants as expected, although this was group dependent. Although many subjects stated they enjoyed one of the moving conditions, subjects were equally likely to mention negatives like decreased productivity and distractions. Negatives may counter motion's potential aesthetic benefits. Six fun and intuitive group members reported enjoying the motion despite the distractions while utility members tended to concentrate on the negatives. We were also surprised at what features subjects considered 'useless' or distracting. Subject 15 stated "when [glyphs] were all going at once, it was like: OK. That's a bit much." Similarly, subject 2 stated that the chaotic motion of the redundant motion was "like static on TV" and could be tuned out, unlike the extraneous motion. Subject 16, however, stated the asynchronous motion was "super distracting" but the extraneous motion was not. We were also surprised that only two subjects named the extraneous condition their favorite given that the condition was designed to be maximally appealing.

4.4.2 Utility, Appeal, and Sub-Optimal System Use (Q2)

As expected, productivity and functionality are not the exclusively important factors for system appeal. However, the factors affecting appeal were far more complex than a simple fun/functionality trade-off. During the interviews, subjects 1 and 5 believed they performed better using the redundant encoding. This suggests either "the more data the better" preconceptions or motion's appeal resulted in a discrepancy between objective metrics and perceived performance. Alternately, positive aesthetic judgments could influence the perceived performance as suggested by Norman [13]. Motion preconceptions appeared to influence preferences. Subject 2 frequently noted his dislike of blinking html to explain his dislike of extraneous motion. Similarly, higher level cognition and anthropomorphic image interpretations seemed to positively affect motion appeal. Subject 15 stated that the moving glyphs "seem happy that you're using it" while subject 3 liked the redundant condition because it "looks like it's alive." Visualization novelty may play a factor in perceived system appeal, however, only subject 10 made such comments explicit.

For fun group members, the emphasis on fun or style seems to be reflected in their approaches to the task. Fun group members more frequently sampled songs without adding them to their playlists and, as a group, they more actively sought out new music rather than familiar songs. It seems reasonable that users wishing to be efficient and accomplish the task would tend to look for familiar music and only choose to sample music if there was a good possibility it would be added to the playlist. Listening to thirty second samples of new music would decrease the number of songs chosen. The large personal music libraries associated with utility group members might also suggest that large music collections lead to a learned emphasis on utility. Searching for familiar artists, albums, or genres rather than sampling unfamiliar music may be an acquired approach due to their music collection size. Alternately, utility members might be more interested in owning music than other subjects.

4.4.3 Subject Groupings

For the subjects observed, it seems that system preferences could be addressed using group-based visualization customizations. Furthermore, our results suggest that with future research, categories of similar thinking users could be identifiable based on markers such as music collection size, background training, and system preferences. The ideal markers for identifying user groups is an open research question. If systems could account for a enumerable number of groups in a target population, this may be sufficient to address most individual user needs. The set of all possible customizations and permutations would not need to be supported. Similarly, less user information may need to be collected to identify user opinions. This research is a first step into looking at group based visualization customizations so generalizations are limited. Future work should consider a larger sample size, a more representative subset of the population (rather than predominately students), and different application domains.

Numerous explanations for the observed groupings are possible. Personality characteristics may dictate how a user approaches system use. Users concerned about utility may be unwilling to try new features or new music because it may impact their perceived performance, while participants in the fun group may prefer novelty and thus the moving glyph encodings. Caution may explain the reluctance of utility group members to use beat encodings and their emphasis on functionality. Perhaps the lack of beat encoding precision also frustrated utility subjects. Adventurous users may want to find new music and hence sample far more music than they add to their playlists (e.g. the fun group). Subject backgrounds, training/education, and gender would logically affect opinions before the study and these experiences may affect responses and performances more than in-study training. This would explain the group differentiations along music genre, background, opinion, and education lines. We note that for effective group based customizations, however, the underlying reason for the grouping is not required.

Gender and occupational differences between the subject groups seem to be the most obvious and alarming result from our grouping scheme. Utility and intuitive groups broke perfectly along gender lines but this does not necessarily mean gender affects preferences. First, there were significant age and occupational differences between men and women which may explain this result. Men were older and were predominately from technical fields like engineering and computer science. Women mostly had non-technical backgrounds. Second, groupings were based on user comments so this may indicate a gender/culture based language difference.

4.5 Implications for Casual InfoVis

Results from this research have some important design implications. Motion improved aesthetic appeal for some people, suggesting it may have some substantial benefits for casual InfoVis systems. Nonetheless, over half of our subjects stated that they found motion distracting, even if they found it entertaining. Motion based data encoding and extraneous motions appear to be polarizing topics with previous experiences playing a large role in its perceived benefits. The utility users found motion frustrating, the fun group enjoyed the aesthetics, fun or novelty, and the intuitive subject group thought motion was a more natural way of encoding beat data (an opinion we expected more subjects to hold). Motion is appealing to some users while unappealing to others, hence care should be taken before introducing it into a visualization, particularly for entertainment / aesthetic purposes only. Previous recommendations to only use motion temporarily, and only when well designed, continues to be sound advice since it was clearly a negative for some users [3, 18, 22]. Alternative ways to increase fun and system appeal should probably be pursued when possible.

Our identification of subject groupings suggest that effective casual InfoVis system design either needs sufficient customizability or the ability to clearly identify target users to be effective. Results also suggest that although users had vastly different goals, opinions and preferences, knowledge about high level user goals and expectations may simplify group-based visualization customizations. For example, knowing a user is a philosophy major who finds motion appealing may mean she wants to discover new music, she is unconcerned about task efficiency, and a more stylized visualization would be appreciated. Similarly, an engineer with a large personal music collection may want a more pragmatic visualization. More research is needed to establish whether consistent subject groups can be identified and whether groupings exist in other application domains.

5 Conclusions & Future Work

Now that we have a better understanding about the characteristics of user groups, quantitative studies and additional qualitative studies need to be designed. Our overall research focus is to establish design guidelines for casual InfoVis and evaluate how various factors affect system appeal and performance. Our results begin to examine the cost/benefit trade-offs between designing for appeal and designing for functionality. Current results suggest that greater research into group classifications and user characteristics may be beneficial. Future research investigating the appeal of different data metaphors and data complexity will be conducted. We also hope to further explore the use of motion in other contexts and its use for data highlighting rather than as a constant interface feature. It would also be interesting to see how aesthetic appeal and perceived functionality changes depending on user goals. For example, emphasizing different performance criteria in a task like playlist generation may result in different judgments and actions.

Our results illustrate the complex and nuanced appreciation for motion in terms of appeal, perceived utility, and functionality in a casual context. For the glyph comparison task, motion based beat encodings were less effective than the static encoding, and motion only data encoding was the least effective. However, performance did not match preferences in our second experiment and some subjects did not realize that the extraneous and redundant encoding conditions were sub-optimal. Three clear user categorizations based on glyph opinions were identified: utility, fun, and intuitive. Differences in subject background, judgment, and performance corresponded with this grouping. Our findings suggest that a group based customization approach to casual InfoVis is a promising new research direction. Furthermore, knowledge about a user group is critical for effective casual InfoVis design. Motion may provide an effective means for attracting a particular subset of users, provided it is designed intelligently. Many of our subjects expressed motion based distraction concerns. Our results support previous recommendations that motion should be used sparingly for pragmatic or utility users, should be transitory in nature (eg. animated transitions and highlighting) and should be subtle to avoid distraction. With more research we can hopefully tease out what costs people are willing to incur for motion's appeal.

References

[1] A. D. Angeli, A. Sutcliffe, and J. Hartmann. Interaction, usability and aesthetics: what influences users' preferences? In *Proc. Conf. Designing Interactive Sys.(DIS '06)*, pages 271–280, 2006.

[2] M. Anshel and D. Marisi. Effects of music and rhythm on physical performance. *Research Quarterly*, 49:109–112, 1979.

[3] L. Bartram. Can motion increase user interface bandwidth in complex systems? In *1997 International Conf. Systems, Man, and Cybernetics*, volume 2, pages 1686–1692, 1997.

[4] L. Bartram, C. Ware, and T. Calvert. Moticons: detection, distraction and task. *Int. J. Hum.-Comput. Stud.*, 58(5):515–545, 2003.

[5] B. B. Bederson and A. Boltman. Does animation help users build mental maps of spatial information? In *Proc. IEEE Symp. Information Visualization (InfoVis '99)*, pages 28–35, 1999.

[6] burst Labs. burst labs home page, 2008. Available at `http://www.burstlabs.com/`.

[7] digg labs. The dig labs swarm visualization, 2008. Available at `http://labs.digg.com/`.

[8] J. Gluck, A. Bunt, and J. McGrenere. Matching attentional draw with utility in interruption. In *CHI '07*, pages 41–50, 2007.

[9] J. Heer and D. Boyd. Vizster: Visualizing online social networks. In *Proc. IEEE Symp. Information Visualization (InfoVis '05)*, pages 33–40, 2005.

[10] J. Heer and G. Robertson. Animated transitions in statistical data graphics. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1240–1247, 2007.

[11] P. Isenberg, T. Zuk, C. Collins, and S. Carpendale. Grounded evaluation of information visualizations. In *Workshop on BEyond time and errors: novel evaluation methods for information visualization*, pages 1–8, 2008.

[12] D. A. Norman. *The Psychology of Everyday Things*. Basic Books, 1988.

[13] D. A. Norman. *Emotional design: why we love (or hate) everyday things*. Basic Books, 2004.

[14] Z. Pousman, J. Stasko, and M. Mateas. Casual information visualization: Depictions of data in everyday life. *IEEE Trans. Visualization and Computer Graphics (TVCG)*, 13(6):1145–1152, 2007.

[15] G. Robertson, R. Fernandez, D. Fisher, B. Lee, and J. Stasko. Effectiveness of animation in trend visualization. *IEEE Trans. Visualization and Computer Graphics (TVCG)*, 14(6):1325–1332, 2008.

[16] M. Shanmugasundaram, P. Irani, and C. Gutwin. Can smooth view transitions facilitate perceptual constancy in node-link diagrams? In *Proc. Graphics Interface 2007*, pages 71–78, 2007.

[17] D. Sprague, F. Wu, and M. Tory. Music selection using the partyvote democratic jukebox. In *Proc. Working Conf. Advanced Visual Interfaces (AVI '08)*, pages 433–436, 2008.

[18] B. Tversky, J. B. Morrison, and M. Betrancourt. Animation: can it facilitate? *Int. J. Hum.-Comput. Stud.*, 57(4):247–262, 2002.

[19] G. Tzanetakis. Musical genre classification of audio signals. In *IEEE Trans. Speech and Audio Processing*, pages 293–302, 2002.

[20] G. Tzanetakis and G. Essl. Human perception and computer extraction of musical beat strength. In *DAFx-02: Proceedings of the Int. Conf. on Digital Audio Effects*, pages 257–261, 2002.

[21] C. Ware and R. Bobrow. Supporting visual queries on medium-sized node-link diagrams. *Proc. IEEE Symp. Information Visualization(InfoVis '05)*, 4(1):49–58, 2005.

[22] C. Ware and R. Bobrow. Motion coding for pattern detection. In *Proc. Symp. Applied Perception in Graphics and Visualization (APGV '06)*, pages 107–110, 2006.

[23] M. Williams and T. Munzner. Steerable, progressive multidimensional scaling. In *Proc. IEEE Symp. Information Visualization (InfoVis '04)*, pages 57–64, 2004.