

AN EMPIRICAL INVESTIGATION INTO THE EFFECT OF PRESENTATION MODE IN THE CINEMATIC AND MUSIC LISTENING EXPERIENCE

Scott D. Lipscomb & Mark Kerins

Northwestern University

ABSTRACT

Two experiments were carried out to determine the extent to which presentation mode (stereo vs. surround sound) impacts the music listening and cinematic experiences. Results reveal that, in both contexts (music and cinema), the genre of the excerpts and responses to the various verbal scales were different to an extremely high level of statistical significance. In the music listening study (Experiment 1), however, presentation mode proved to be a less significant factor than in the cinema study (Experiment 2), emerging as a significant factor in only two contexts: when paired with the level of visual training and in a more complex three-way interaction. Also, in both the musical and cinematic experimental contexts, visual training emerged as a more significant factor than musical training, even though the tasks focused on the auditory component of the experience.

A set of prior experiments using a similar empirical method revealed no significant difference between subject responses in either the musical or cinematic context as a result of manipulating presentation mode. In the current set of experiments, even after revising the empirical method in a manner intended to emphasize differences between presentation modes, the results suggest that the affect of surround sound presentation appears to be lost on the music listening audience. Results are more promising in the cinematic context, since there does appear to be a significant difference in subject responses to the excerpts presented as a result of presentation mode. Specific differences in mean responses to the various VAME scales and the manner in which they vary as a result of musical and/or cinematic genre and presentation mode will be presented at the conference.

1. BACKGROUND

In recent years, there has been a significant increase in the number of empirical investigations into the role of music within the context of the motion picture experience (Boltz, 2001; Bullerjahn & Gldenring, 1994; Lipscomb, 1995; Lipscomb, in press; Lipscomb & Kendall, 1994; Marshall & Cohen, 1988; Thayer & Levenson, 1983; Thompson, Russo, & Sinclair, 1994; to name just a few). Little attention has been paid, however, to the manner in which this sonic component is delivered to the viewing – and listening – audience. Despite the facts that a growing number of theaters worldwide are now outfitted with surround sound (almost 44,000 with Dolby Digital, nearly 10,000 with Dolby Digital EX, approximately 22,000 Digital Theater Systems (*dts*), and over 8,000 Sony Dynamic Digital

Sound (SDDS); Dolby Laboratories, 2004) and that this capability is available in millions of home theater systems (approximately 40 million 5.1-channel receivers incorporating Dolby Digital decoders have been purchased by consumers; Dolby Laboratories, 2004), there has been virtually no empirical work carried out to determine the affect of presentation mode on the movie viewing or music listening experience. It is this void in our present state of knowledge that the current study wishes to begin to remedy.

2. SIGNIFICANCE OF THE STUDY

Since the very beginning of cinema, musical sound has always played an important role. Even the earliest silent films were accompanied by piano if, for no other reason, than to cover the sound of a noisy projector (Cavalcanti, 1985). However, the arrival of “talkies” (i.e., movies with synchronized sound) initiated a period of sonic experimentation that continues to the present day, searching for higher fidelity and, in some cases, more accurate spatial representation of the sound sources represented on-screen. Early three-channel experiments with musical sound – front left, front center, and front right – were carried out at Bell Labs in the early 1930s (Holman, 2000), followed shortly thereafter by the introduction of “Fantasound” for the release of Disney’s animated feature film, *Fantasia* (1941) ... not surprisingly, a film in which the musical score played a prominent role.

Though it is beyond the scope of the present paper to review the entire evolution of multi-channel systems (for a detailed accounting of these developments, see Kerins, 2003), it is worth noting some of the more recent developments that have proven significant:

- introduction of “Dolby Stereo,” a 4-channel system (*Star Wars*, 1977)
- addition of a subwoofer to the 4-channel system to enhance the presence and fidelity of low frequency sounds (*Close Encounters of the Third Kind*, 1977)
- separation of surround (rear) speakers into discrete channels, creating a true 5-channel system, plus subwoofer (*Superman*, 1978)
- the introduction of “Dolby Digital” *Batman Returns*, 1992)
- *dts* (*Jurassic Park*, 1993)
- the 7.1-channel SDDS system (*The Last Action Hero*, 1993)

While technological developments have continued at a dizzying pace, little empirical work can be found in the literature of the period to

confirm that there is, in fact, a positive enhancement of the cinematic experience as a result of these “new and improved” sound systems.

Similarly, in the realm of music listening, the music industry has developed – and is now in the process of heavily promoting – two competing formats of multi-channel audio designed especially for musical recordings: Sony’s super audio compact disc (SACD) and Philips’ DVD-Audio (DVD-A). The primary purpose of the present investigation is to determine whether or not there is a significant positive affect of these new technologies upon either the cinematic or music listening experience.

3. RESEARCH QUESTIONS

Presentation mode (i.e., spatial location and number of sound sources), in the context of the motion picture and music listening experiences, is the variable of particular interest in the present study. Many past investigations, referenced above, have shown the dramatic impact that a musical score can have on viewer perception of a given visual scene within the cinematic context. Though a great deal of technological development and a significant amount of associated marketing have been carried out by various corporate entities within the music industry in recent years, little empirical research exists investigating the affect of presentation mode upon the viewer-listener’s experience. The fundamental questions for the present paper are the following: 1) Does presentation mode significantly influence subject verbal reports about the music listening experience? 2) Does presentation mode significantly influence subject verbal reports about the cinematic experience? 3) Does gender exert a significant influence on a subject’s verbal responses to the listening experience? 4) Does an individual’s level of musical training significantly influence their verbal responses to the listening experience? 5) Does an individual’s level of visual training significantly influence their verbal responses to the listening experience?

Two experiments were designed to answer these questions.

4. EXPERIMENT 1 – MUSIC

4.1 Participants

Eighty-five students enrolled in courses at Northwestern University volunteered to serve as participants in the present investigation (mean age = 21.25; SD = 1.825). Between-subjects attribute variables of interest included gender, musical training, and visual training. Two additional between-subjects variables were added to account for the presentation mode factor, as described in the “Method” section for Experiment 1. The subject pool consisted of 47 males and 38 females. The amount of formal musical training was operationally divided into three categories: low (3 years or less; n=30), moderate (4 to 8 years; n=29), and high (more than 8 years; n=26). The amount of formal visual training was also operationally determined, though divided into only two categories: untrained (no formal training; n=50) or trained (i.e., at least one college-level course in the visual arts, including – but not limited to – painting, sculpture, and film; n=35). A description of the within-subjects variables

incorporated into the repeated measures design will be provided at the conclusion of the “Method” section for Experiment 1.

All students participated in both of the experiments described below. In all cases, the musical stimuli were presented first (Experiment 1), followed by the movie excerpts (Experiment 2).

4.2 Equipment

The experimental procedure was carried out in a free-field classroom in Northwestern University’s Music Administration Building, part of the School of Music. This room is outfitted with a 5.1 surround sound system for movie and audio playback. Heavy insular blankets were hung over the windows to limit the amount of light entering the room and to minimize the reflections off of these surfaces. The same heavy blankets were hung across the rear and side walls of the room to further minimize ambient reflections. A Sony DVP-NS755V DVD/SACD player with six discrete RCA outputs for each channel of audio (front left, center, front right, surround left, surround right, and subwoofer) was used to play SACDs for the music listening experiment. The audio outputs for both the music and movie excerpts were fed into a 310-watt Yamaha RX-V795a receiver, which routed the subwoofer channel to a 250-watt Event Electronics 20/20 sub and the other five audio channels to one of five appropriately located Alesis Point Seven shielded reference monitors, mounted approximately 8.5 ft. above the floor and angled downward at approximately 20°.¹

4.3 Stimuli

Musical excerpts were chosen to include two representative examples of five musical styles: dance, classical, jazz, alternative rock, and adult contemporary. Excerpts were selected only if an SACD recording was available, allowing both surround sound and stereo playback (see Appendix I for a list of the specific musical excerpts used). Prior to each experimental session, the same recording (“Man in a Box” by Alice in Chains) was used to set the maximum loudness level to 85 dB_{SPL} (C-weighted).

In the present study, two types of audio presentation were utilized: stereo and 5.1 surround sound. This choice, attempting to maintain a high level of ecological validity, was based on the fact that these are the two presentation modes most likely to be experienced by a music listening audience. Two-channel stereo remains, at present, the dominant mode of music listening, but 5.1 surround sound appears to be its up-and-coming successor. Admittedly, the audience for surround sound music listening currently constitutes a very small percentage of the population, but as home theater systems become more prevalent – it is, after all, becoming quite difficult to purchase an audio receiver that does *not* have surround sound capability – the audience will undoubtedly grow ... especially considering the marketing push currently being made by the recording industry for

¹ The steeper angle was required, rather than the more common 10° tilt, due to the fact that the room size was significantly smaller than that used for a typical theatrical installation.

DVD-A and SACD recordings and the spate of re-releases being issued in one of these “new and improved” formats.

4.4 Method

A past study using an experimental procedure similar to that described below (Lipscomb, Kendall, Moorefield, & Tolchinsky, 2003) utilized independent groups – i.e., one heard only stereo presentations and the other heard only surround sound presentations – and selected variables representing the evaluative, potency, and activity dimensions proposed by Osgood, Suci, & Tannenbaum (1957). Lipscomb, et al. found no significant difference between subject verbal responses as a result of presentation mode. Therefore, two significant changes were made to the empirical method for the present study, both specifically intended to emphasize the differences between the two types of presentation. First, each group of subjects was exposed both to stereo and surround sound presentation modes in an alternating sequence. Second, in order to highlight the “immersiveness” factor of the sonic environment, a list of verbal terms was selected that specifically addressed such auditory characteristics. In addition to one evaluative, potency, and activity term (good, heavy, and active, respectively), the new terms included: artificial, complex, diffuse, expansive, full, immersive, intimate, and natural. Participants were instructed to carefully listen to each musical excerpt and provide a series of verbal ratings.² The subject response method utilized was Kendall & Carterette’s (1992, 1993) verbal attribute magnitude estimation (VAME). In contrast to semantic differential scales (Osgood, et. al., 1957), by using true opposites at anchor positions, VAME provides a means of assigning a specific amount of a given attribute within a verbal scaling framework (e.g., not good–good, as opposed to bad–good).

In order to account for differences as a result of presentation mode within- and between-genres, it was necessary to manipulate two additional between-subjects variables. The order of genres within an experimental session was randomized, but subjects always heard both examples of a given genre consecutively. The presentation mode variable determined whether stereo or surround sound was used consistently as the initial example within each genre category. The order variable determined which of the two excerpts within a genre was presented first. This configuration resulted in four possible presentation mode-by-order arrangements (2 modes x 2 orders). Within-subjects variables accounted for in the repeated measures design included genre, specific excerpt within each genre (identified as “song” for Experiment 1 and “movie” for Experiment 2), and VAME scale. The genre variable consisted of five levels in the music listening experiment (Experiment 1) and three levels in the cinema experiment (Experiment 2). The excerpt variable consisted of two levels and represented the specific excerpt of the two musical or

cinematic selections within a given genre to which subjects were responding. The VAME scale consisted of 11 levels, each representing one of the verbal response scales listed above. Four random orders of VAME scale presentation order were generated and then randomly assigned to each excerpt within an individual subject’s response form packet, creating a unique response order for every participant. Subjects responded on an 8-point scale (0 to 7) by placing an “X” in the appropriate space on their response form.

4.5 Results

Results revealed that there were no main effects of any of the five between-subjects variables (mode, order, gender, musical training, or visual training). However, there was a statistically significant interaction between presentation mode and visual training ($F_{(1,42)}=4.790$; $p=.034$) and a significant three-way interaction between order, gender, and visual training ($F_{(1,42)}=5.014$; $p=.030$). Two within-subjects main effects reached the level of statistical significance: genre ($F_{(4,168)}=52.783$; $p<.0005$) and VAME ($F_{(6,507,273,314)}=51.559$; $p<.0005$).³ Interaction effects – only those up to third order are reported for both Experiments 1 and 2 – include: song-mode-order ($F_{(1,42)}=4.939$; $p=.032$), genre-song ($F_{(3,322,139,508)}=8.753$; $p<.0005$), genre-song-order ($F_{(3,322,139,508)}=2.656$; $p=.045$), genre-song-musical training ($F_{(6,643,139,508)}=2.363$; $p=.028$), genre-VAME ($F_{(15,917,668,534)}=67.374$; $p<.0005$), genre-VAME-gender ($F_{(15,917,668,534)}=1.774$; $p=.031$), song-VAME ($F_{(7,345,308,498)}=8.548$; $p<.0005$), and genre-song-VAME ($F_{(15,917,668,515)}=9.045$; $p<.0005$). The interaction between genre-VAME-mode was determined to be significant according to both the sphericity-assumed ($F_{(10,420)}=2.079$; $p=.025$) and Huynh-Feldt ($F_{(10,000420,000)}=2.079$; $p=.025$) adjusted value calculations, but was just barely higher than the .05 value for the Greenhouse-Geisser ($F_{(15,917,668,534)}=10.185$; $p=.054$).

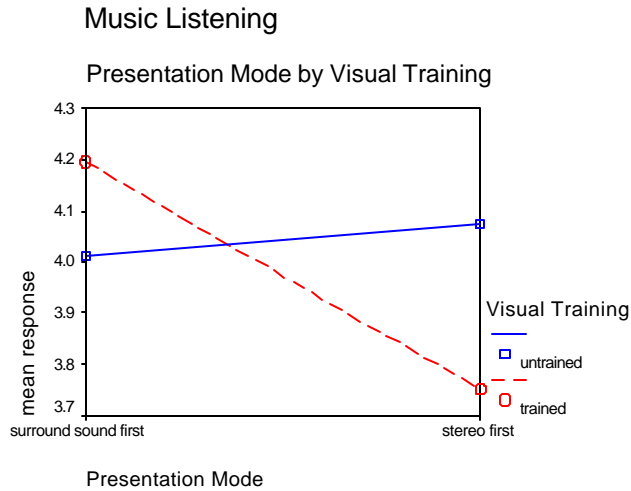
As expected, subjects responded differently to music of varying genres and they responded differently to the numerous verbal response scales. Of particular interest, however, are the apparently strong significance of visual training (e.g., presentation mode-visual training and order-gender-visual training interactions) and the seeming insignificance of musical training in relation to participant verbal responses, with the exception of the genre-song-musical training interaction. This result is particularly interesting due to the fact that the stimuli for Experiment 1 were audio-only musical excerpts, with no visual component. Those participants with some degree of visual training responded significantly higher to the verbal scales when the surround sound presentation came first than those who heard the stereo presentation first, while participants with no visual training responded almost identically regardless of the order in which the two presentation modes were experienced (Figure 1).

² In addition to the eleven VAME scales listed above, these ratings included a familiarity measure for each excerpt and a preference assessment in comparison to the previous excerpt. Only the VAME scales will be analyzed in the present paper.

³ Subject responses across all within-subject variables *except genre* violated the assumption of sphericity as determined by Mauchly’s test. As a result, all F - and p -values reported for Experiment 1 except for the main effect of genre utilize the more conservative Greenhouse-Geisser calculations, with adjusted values for the degrees of freedom.

Other differences and a more more detailed discussion of these results – especially the genre-VAME-gender interaction and the near-significant genre-VAME-mode interaction – will be provided during the conference presentation.

Figure 1. A line graph representing the mean responses for visually trained and untrained participants across presentation mode.



5. EXPERIMENT 2 – CINEMA

Participants

The same group of subjects described in Experiment 1 participated in the second experiment. Due to the fact that several subject response forms for Experiment 2 exhibited missing values, the number of subjects in the two experiments differed slightly. Experiment 1 (music listening) included all 85 participants, while Experiment 2 (cinematic excerpts) included only 77, grouped into the following categories using the same operational definitions delineated previously: gender (43 males and 34 females), level of musical training (26 low, 26 moderate, and 25 high), and level of visual training (46 untrained and 31 trained).

5.1 Equipment

The equipment used was identical to that described for Experiment 1, except that a SonyDVP-NS400D CD/DVD player was used to play the movie excerpts and feed the audio signal to the same Yamaha receiver and speaker system previously described.

5.2 Stimuli

Cinematic excerpts were chosen to include two representative examples of three film genres: action-adventure, drama, and suspense-horror (see Appendix II for a list of the cinematic stimuli used). Excerpts were selected only if the DVD recording included both a 2-channel and 5.1 surround sound mix, allowing both surround sound and stereo playback. Similar to the music listening experiment, these choices were based on our desire to maintain a high level of ecological validity. As noted previously, a growing number of

theaters currently present theatrical sound using one of the available surround sound systems, while viewers watching movies at home most often hear two channels of sound coming directly from their television set or from a stereo system into which the audio has been routed. The Dolby Surround 2.0 tracks used for the “stereo” presentation mode in this experiment consist of a matrixed 4-channel configuration that is designed to be played back on a 4-channel system, but also to play acceptably on a 2-channel system, a common configuration for many home systems. The maximum loudness level was calibrated to 85 dB_{SPL} (C-weighted).

5.3 Method

Because a previous study (Kerins & Lipscomb, 2003), using the same basic method as Lipscomb, et al. (2003) but with movie excerpts, revealed no significant difference in responses due to presentation mode, the same changes were made to the experimental method for the cinema study as those described for Experiment 1. As a result, the method for Experiment 2 was identical to that of Experiment 1, except that participants watched and listened to cinematic excerpts before responding on the same 13 VAME scales. Subjects were explicitly instructed, however, to respond to the series of verbal scales based solely on the *sound* component of the movie excerpts.

5.4 Results

Analysis of the verbal responses to the movie excerpts revealed no statistically significant main effects of or interactions between the between-subjects variables. There were, however, a number of significant main effects of and interactions between the within-subjects variables:⁴ genre ($F_{(2,72)}=9.259$; $p<.0005$), movie ($F_{(1,000,36,000)}=7.939$; $p=.008$), VAME ($F_{(5,562,200,249)}=29.886$; $p<.0005$) genre-visual training ($F_{(1,976,71,142)}=4.108$; $p=.021$), genre-mode-order ($F_{(1,976,71,142)}=3.440$; $p=.038$), movie-mode-order ($F_{(1,000,36,000)}=5.460$; $p=.025$), genre-movie ($F_{(2,72)}=12.011$; $p<.0005$), genre-VAME ($F_{(10,087,363,122)}=30.708$; $p<.0005$), movie-VAME ($F_{(7,003,252,100)}=38.216$; $p<.0005$), genre-movie-VAME ($F_{(9,484,341,431)}=15.337$; $p<.0005$).

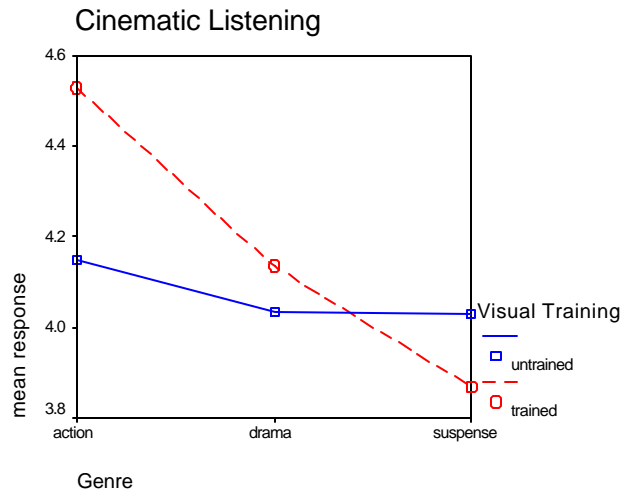
As in the music listening study, participants responded differently to excerpts representing the various cinematic genres and they responded differently on the numerous verbal response scales. Once again, the role of musical training did not emerge as a significant factor either as a main effect or in any of the within-subjects interactions, while visual training had a significant impact upon the manner in which participants responded to the various cinematic genres (Figure 2). Those with no visual training responded similarly to all three genres, while those with some degree of training responded differently to each genre, i.e., much higher to the action-adventure

⁴ The Greenhouse-Geisser estimates for statistical significance will be used for all factors except the main effect of genre and the genre-movie interaction for the same reason cited in the data analysis for Experiment 1.

genre, lowest to the suspense-horror genre, and moderately to the drama genre.

The most important result of this cinema experiment is the discovery of a significant interaction between cinematic genre, presentation mode, and order of presentation. Recall that, in order to include both presentation modes in every experimental session, subjects were randomly assigned to one of four groups, based on presentation order (stereo first or surround sound first) and order of excerpt presentation (within each genre, whether the first or second example was played initially). Therefore, it appears that, unlike the results of the music listening responses in Experiment 1, mean subject responses to the various cinematic excerpts are, in fact, significantly influenced by the presentation mode of the given excerpt. Due to space restrictions associated with the present document, details of the specific responses and the manner in which they vary will be presented at the conference.

Figure 2. A line graph representing the mean responses for visually trained and untrained participants across cinematic genre.



6. CONCLUSIONS

Results of the present study reveal that the answers to the stated Research Questions are multifaceted and complex. In both contexts (music and cinema), the genre of the excerpts and responses to the various verbal scales were different to an extremely high level of statistical significance. In the music listening study (Experiment 1), however, presentation mode proved to be a less significant factor than in the cinema study (Experiment 2), emerging as a significant factor only when paired with the level of visual training and a single 3-way interaction. Also, in both experimental contexts, visual training emerged as a more significant factor than musical training, even though the task was focused on the auditory component of the experience, whether listening to musical excerpts or seeing-hearing movie excerpts.

In conclusion, even after revising the empirical method in a manner intended to emphasize differences between presentation modes, the results suggest that the affect of surround sound presentation

appears to be lost on the music listening audience. Results are more promising in the cinematic context, since there does appear to be a significant difference in subject responses to the excerpts presented as a result of presentation mode. Specific differences in mean responses to the various VAME scales and the manner in which they vary as a result of musical and/or cinematic genre and presentation mode will be presented at the conference.

7. REFERENCES

- Boltz, M. (2001). Musical soundtracks as a schematic influence on the cognitive processing of filmed events. *Music Perception*, 18(4), 427-454.
- Bullerjahn, C. and Gldenring, M. (1994). An empirical investigation of effects of film music using qualitative content analysis. *Psychomusicology*, 13, 99-118.
- Cavalcanti, A. (1985). Sound in films. In *Film sound: Theory and practice* (eds. E. Weis & J. Belton), pp. 98-111. New York: Columbia University Press.
- Dolby Laboratories (2004, May 19). Worldwide statistics. Retrieved June 14, 2004, from <http://www.dolby.com/stats/>.
- Holman, T. (2000). *5.1 surround sound: Up and running*. Boston, MA: Focal Press.
- Kendall, R.A. & Carterette, E.C. (1992, February). Semantic space of wind instrument dyads as a basis for orchestration. Paper presented at the Second International Conference on Music Perception and Cognition, Los Angeles, CA.
- Kendall, R.A. & Carterette, E.C. (1993). Verbal attributes of simultaneous wind instrument timbres: I. von Bismarck adjectives. *Music Perception*, 10(4), 445-467.
- Kerins, M. (2003). Building a history of multi-channel sound for movies. Unpublished paper presented at Northwestern University.
- Kerins, M. & Lipscomb, S.D. (2003). Study of audience perception of sound modes in cinema. Paper presented at the conference of the Society for Music Perception & Cognition. Las Vegas, NV. (June, 2003).
- Lipscomb, S.D. (1995). *Cognition of musical and visual accent structure alignment in film and animation*. Unpublished doctoral dissertation, University of California, Los Angeles.
- Lipscomb, S.D. (in press). The perception of audio-visual composites: Accent structure alignment of simple stimuli. *Selected Reports in Ethnomusicology*, 12.
- Lipscomb, S.D., Kendall, G., Moorefield, V., & Tolchinsky, D.E. (2003, June). Immersive sound: Does 5.1 surround sound really make a difference in the affective experience of music listening? Paper presented at the conference of the Society for Music Perception & Cognition. Las Vegas, NV.

- Lipscomb, S.D. & Kendall, R.A. (1994). Perceptual judgment of the relationship between musical and visual components in film. *Psychomusicology*, 13(1), 60-98.
- Marshall, S.K. and Cohen, A.J. (1988). Effects of musical soundtracks on attitudes toward animated geometric figures. *Music Perception*, 6(1), 95-112.
- Osgood, C.E., Suci, G.J., & Tannenbaum, P.H. (1957). The measurement of meaning. Urbana: University of Illinois Press.
- Thayer, J.F. and Levenson, R.W. (1983). Effects of music on psychophysiological responses to a stressful film. *Psychomusicology*, 3(1), 44-52.
- Thompson, W.F., Russo, F.A., and Sinclair, D. (1994). Effects of underscoring on the perception of closure in filmed events. *Psychomusicology*, 13, 9-27.

Appendix I

- Dance Music:** “Bootylicious” by Destiny’s Child (0:00 to 1:30; Columbia CS 61063) and “Rhythm’s Gonna Get You” by Gloria Estefan (0:00 to 1:35; Epic ES 86729)
- Classical:** “Oh, Fortuna,” the opening of Orff’s *Carmina Burana* (0:00 to 1:55; TelArc SACD-60575) and “1812 Overture” by Tchaikovsky (14:00 to 15:45, including cannon fire; TelArc SACD-60541)
- Jazz:** “Freddie Freeloader” by Miles Davis (0:00 to 1:52; Columbia CS 64935) and “Dancing in the Dark” by Diana Krall (0:00 to 1:36; Verve 314 589 597-2)
- Alternative Rock:** “Man in the Box” by Alice in Chains (0:00 to 1:45; Columbia CS 85922) and “(What’s the Story) Morning Glory” by Oasis (0:38 to 2:20; Epic ES 67351)
- Adult Contemporary:** “Africa” by Toto (0:00 to 1:40; Columbia CS 86544) and “Differences” by Genuwine (0:00 to 1:50; Sony ES 69622)

Appendix II

- Action-Adventure:** *Star Wars Episode I: The Phantom Menace* (1:05:56 to 1:09:21; 20th Century Fox 2002391) and *Terminator 2: Judgment Day* (1:46:09 to 1:48:41; Artisan 10967)
- Drama:** *American Beauty* (1:48:37 to 1:51:22; Dreamworks 85382) and *Se7en* (1:56:48 to 2:00:03; New Line N4997)
- Suspense-Horror:** *What Lies Beneath* (37:34 to 40:52; Dreamworks 86406) and *Joyride* (58:25 to 1:01:20; 20th Century Fox 2003627)