Dolby v DTS - The academic viewpoint

by Guy Walker BSc and Shui-I-Shih PhD

By its very nature, *Cinema Technology* usually concentrates on the 'hands-on' practical side of the exhibition business, and technical / academic papers are fairly rare events, with the classic Ioan Allen paper 'Are movies too loud' being one of the relatively few within recent memory.

This paper, from *Guy Walker* at the Department of Design, Brunel University, and the Tivoli Theatre, Wimborne, Dorset, and *Shui-I-Shih* of the Department of Psychology at the University of Southampton, shows the fascinating results of applying academic theory and practice to cinema sound, a controversial topic that is usually regarded by most cinemagoers as purely subjective.

The commercial companies whose products have been involved in these 'tests' will be invited to put forward their comments in future issues of *Cinema Technology.*

Jim Slater

Corresponding author's contact details:
email: guy.walker@brunel.ac.uk
Tel:
+44 (0)1784 431341
Ext. 244

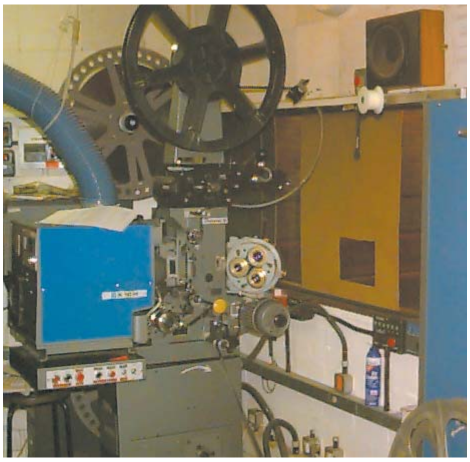
Abstract

Auditory masking describes the ability of one sound component to render another sound component inaudible to the listener. The effect of discarding to-be-masked sounds in a commercial application was evaluated by comparing Dolby Digital (DD) and DTS on the basis of perceived sound quality.

No sounds are completely discarded when recorded by DTS. Whereas sounds predicted to be inaudible due to auditory masking are completely discarded when recorded by DD.

On each trial, the listeners were presented with two paired sound samples in succession, with each half of the pair recorded by identical or different systems. They indicated which sample (first or second) they preferred on a rating scale. Signal Detection Theory (SDT) was employed as the analysis method.

This article demonstrates, firstly, the results that can be gained through SDT, and its applicability towards research concerned with subjective sound quality. In this case, the results indicated that DD was more favourably rated than DTS. Secondly, while fully accepting the limitations of the present experimental paradigm, it can be concluded that the amount of digital data used to describe an audio signal is not necessarily a reliable metric for perceived sound quality.



Introduction

The audio community's traditional stance on reproducing sounds with very high quality is to attempt to reproduce the original audio signal not only accurately, but in its entirety. This is termed a linear perspective, and infers that what you put in to any audio system is linearly related (as far as possible) to what you get out. However, the receiving apparatus, the human ear, is in fact far from linear. Furthermore, it is connected to the world's most powerful signal processor, the human brain. What the brain does with the sounds that we perceive is studied within the field of psychoacoustics, itself a branch of psychophysics. Psychophysics describes the relationship between objective, measurable, physical stimuli occurring out there in the world, and our perception of them as humans.

The sounds we encounter in daily life are usually very complex. But not all sounds transmitted to our inner ears can be detected. The detectability of a sound component (e.g., a tone) depends not only on its intensity (intensity is an objective measure that we experience as 'loudness') but also on the characteristics of other sound components. For example, a particular sound component can be detected (heard) when it is presented in isolation, but not when another tone or noise of similar frequency content is presented earlier, simultaneously, or later in time (e.g., Fletcher & Munson, 1937; Jesteadt, Bacon, & Lehman, 1982; Kallman & Morris, 1984). The reduced or blocked ability in detecting one sound component in the presence of another component(s) is called auditory masking.

Psychophysical studies on auditory masking have enabled us to predict with great accuracy the masking effect that one sound component has upon another. The ability to predict which components in complex sounds are audible has significant applied value as far as sound reproduction is concerned. To reproduce a sound, the original sound has to be recorded or stored in a medium (e.g., an optical soundtrack) and then played back via a reproduction system (e.g. a projector sound head, processor, amplifiers, and loudspeakers). Any component in the reproduction system may introduce distortions, which in turn affect the fidelity and perceived sound quality of reproduction. But fundamentally the reproduction, of course, depends heavily on what has been recorded originally.

Taking a linear stance, and given a reproduction system, the highest achievable fidelity would be to faithfully record every component in the original sound signal (a data-lossless approach). If the entire sound signal is to be stored digitally using contemporary linear encoding methods such as Pulse Code Modulation (PCM), then it demands huge storage space, 650MB for 74minutes of two channel audio in the case of Compact Disc, for instance. In other audio domains this demand for storage space can be difficult to meet. In particular, the desire for multichannel digital cinema sound has always been in conflict with the ability of motion picture film to cope with storing the enormous amount of digital data that is normally required.

Employing sophisticated high speed digital signal processing techniques, coupled with an understanding of auditory masking and psychoacoustics, provides an opportunity to selectively discard sound components that are predicted to be inaudible to humans and to only record the audible components (a data-reduction approach). The data-reduction approach provides not only a potential saving in data storage resources, but also a direct test of our understanding of the workings of the human auditory system.

The present study evaluated the two approaches (data-lossless and data-reduction) with respect to their application in commercial cinema sound. Specifically, we compared DTS that employs the APT X-100 audio compression algorithm (something approaching a data-lossless approach) and DD, which employs their proprietary AC3 audio compression algorithm (a data-reduction approach). APT X-100 is a form of Adaptive Differential Pulse Code Modulation (ADPCM), and performs a *time based* predictive analysis of the sound signal. In DTS, data allocation is biased towards lower frequencies (where most of useful auditory perception occurs) but it should be emphasised that unlike DD, no sounds are completely discarded according to a model of auditory masking. This yields an amount of data too large to be printed onto the actual film print itself, and requires two Compact Discs to store the film soundtrack. To synchronise the film frames and sound samples, a simple and robust time code is printed on the film print, and at playback, the time code tells the decoder which samples of sound to play for any given film frame.

Dolby's AC3 audio compression algorithm is a form of Transform Coding (TC), and performs a *frequency based* analysis of the sound signal. This analysis allows it to selectively record sound components according to a sophisticated in-built model of auditory masking and sensitivity. Sound components predicted to be inaudible according to the model are completely discarded from the sample; whereas sound components predicted to be perceptually salient are often coded with greater resolution than they are in a linear audio system. Discarding sound components in this manner causes its own distortion in the form of quantisation noise. To overcome this negative side effect the same principles of auditory masking are applied. This ensures that the desirable sounds that are left can be shaped such that they cover up, or mask any resultant quantisation noise (Davidson, Fielder, & Link, 1994). This yields a reduction in digital data to the extent that it is possible to print it onto the actual film itself in the form of "pixels" located between the sprocket holes. As the film print passes through the cinema projector, an optical device "reads" the data pixels and converts them into an electrical bitstream to be decoded. Using DD the cinema audience is presented, on average, with around 20% of the original audio signal (APT, 1998).

Understandably this is a frightening statistic, if we take the quantity of digital data used to describe an audio signal as a reliable metric for perceived sound quality. Anecdotal evidence, coupled with enthusiastic industry (and public) debate implies that a reduced data set produces poorer sound quality than an entire data set (Dolby, 1999; Sun, 1998). Given the highly variable circumstances in which these judgments seem to be made, we perhaps need to firstly scrutinise how we go about measuring and analyzing perceived sound quality in the first place. At one end of the scale we have informal magazine publications who often present paradigms with virtually no experimental control. A cursory review of the scientific literature does not necessarily fare that much better. Highlights include experiments whereby participants are not blind to the fact that sound stimuli are from different sound reproducing apparatus (Gabrielsson & Sjogren, 1979; Gabrielsson, Rosenberg & Sjogren, 1974), or subjective methods such as factor analysis are employed (Eisler, 1966).

Given that some debate exists as regards the relative sound quality of DD and DTS, and that the methods to analyse such differences are variable, the intention within this paper is twofold. Firstly, to introduce SDT as a particularly rigorous methodology for assessing the ability of human listeners to detect differences in perceived sound quality, and to present the kind of results that can be derived. The DD vs DTS debate is a convenient and topical vehicle upon which to meet this first objective. Secondly, although limited to some extent by the present experimental scenario, this paper presents an opportunity to put our understanding of auditory masking into a direct empirical test. This is achieved by examining whether human listeners can perceptually experience the difference between the two systems.

To evaluate whether the two systems can be perceptually discriminated, we asked the listeners to provide preference ratings on the basis of perceived sound quality. On each trial, listeners were presented in succession with two sound samples recorded by the same system or different systems. The listeners indicated which sample (the first or second) they preferred on a rating scale. We expected the rating would cluster around the "No Preference" category when the two samples were recorded by the same system. If perceptible differences exist in the reproduction between sound samples recorded by different systems, we expect one system would receive a higher preference than the other. We did not specifically predict which system would receive higher preference because the judgement was contingent on subjective experience and bias. A mere preference towards one system or the other is sufficient to demonstrate that the two systems induce a different perceptual experience.

METHOD

Design

On each trial a pair of sound clips (Clip1 and Clip 2) were presented in succession, without projecting a cinema picture. Two within-subject factors were manipulated: (1) the system used to record Clip1 (DTS or DD) and (2) the system used to record Clip 2 (DTS or DD). The two factors were crossed, yielding 4 conditions: DTS-DTS, DD-DD, DTS-DD, DD-DTS. There were 10 trials for all but one condition. Due to a technical problem, there were only 9 trials for the DTS-DTS condition. The 4 conditions were counterbalanced. Listeners were informed that the 4 conditions would occur with equal probability. On each trial, the task was to compare the two sound clips on the basis of overall perceived sound quality.

The result of the comparison was recorded as one of the following 7 response categories:

- Prefer Clip 1 with 3 categories ranging from strong to weak preference
- Equal preference for Clip1 and Clip 2
- Prefer Clip 2 with 3 categories ranging from strong to weak preference.

Subjects were blind with respect to which sound system was being played to them at any given point.

Procedure.

Prior to the experimental trials the first author provided the 'audience' with instructions and demonstrations. The demonstration illustrated the very apparent differences in sound quality between various analogue sound formats and digital sound. It was emphasised that on half of the trials the two clips were recorded by the same system, and the remaining half by different systems. It was further emphasized that the two clips within a pair were different in content, and that it was the overall quality of the sound that was to be evaluated (See Note 1). However, we did ensure that the nature of the content was matched within a pair. For example, if Clip 1 contained a male voice in conversation, Clip 2 would contain a male voice in conversation; if Clip 1 contained sounds related to a violent scene, Clip 2 would contain sounds related to a violent scene, and so on. It was anticipated that if there were any fairly marked, systematic differences in sound quality, then they would manifest themselves over the course of 39 trials (many more than is the norm), despite the unmatched content of the sound stimuli (See Note 1). The audience were advised that there was no right or wrong answer and that they should judge the sound quality using whatever personal criteria they choose.

On each trial, Clip 1 was presented for 15 seconds, and after a brief pause, Clip 2 was presented for 15 seconds. During the pause between Clip 1 and Clip 2, the system was switched, or made to appear to be switched by the projectionist. The audience was prompted for a response selection via an overhead projector after the presentation of Clip 2, and listeners recorded their judgments. At the end of the experiment, listeners were asked to fill out an optional questionnaire (see below).

Materials

A 'Dolby Tone' Test-Film, cat. 69T, Dolby Channel I.D. Film, and DTS 6-Track Set-Up DS1 Disk were used to check and calibrate the cinema sound system. The sound clips used in the present experiment were selected from Reels 5, 6 and 7 of the UK 35mm Distribution Print (Copy Number 406) of the motion picture "Lethal Weapon 4". The reels were in good clean condition.

A booklet was provided for each observer. The booklet consisted of three parts. The first part provided brief instructions and illustrated the response categories in a seven-point categorical rating scale (see Design). The second part listed the response categories for each trial. The third part was a brief questionnaire in which the listeners could provide some demographic data (e.g., gender and age) and indicate whether they thought they would, or could, be able to tell the difference between the two sound systems. Space was also provided for general observations and comments.

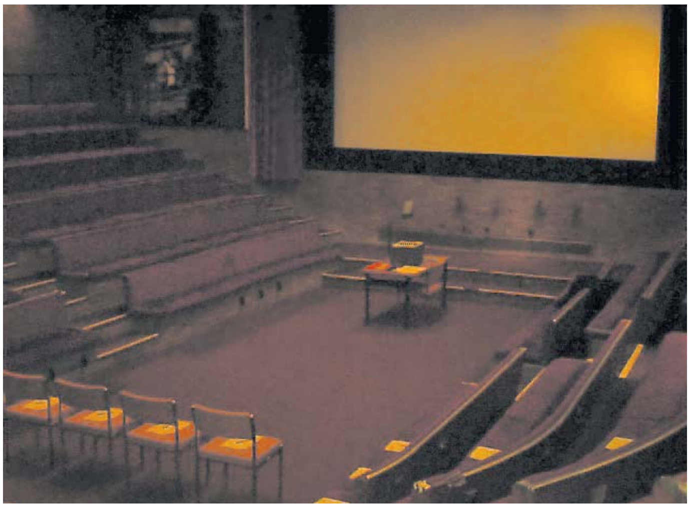
Apparatus.

The experiment took place in the auditoria of Southampton University's film society, Union Films. The cinema auditorium (views shown alongside) seats around 400 people, and is equipped with baffles within the roof to damp sounds from the air-conditioning plant, but the walls are undamped, bare brick. The cinema features a 5.2metre CinemaScope perforated Perlux screen. The experiment was conducted with the lights up, and the cinema screen set to its full width (CinemaScope) to ensure maximum sound throughput from the stage loudspeaker systems, although no cinema picture appeared during the test.

Behind the screen were three main stage loudspeakers located left, center, and right, respectively, to the screen. They were 150-300 watt, full range (40Hz - 20KHz) JBL4671 models, which have a wide dispersion 'Bi-Radial' horn loaded high frequency unit, and a reflex loaded 15) bass/midrange cone drive unit. The fourth loudspeaker behind the screen was a KCS C-118-APL low frequency 'sub-woofer' of 500/1000 watts, featuring a single 18) high power reflex loaded bass cone driver, with a limited low frequency response of 30 to 500Hz. Two (stereo) QSC USA850 200 watt per channel amplifiers powered these four loudspeaker systems. Also equipped were eight 'surround' loudspeakers, three on each sidewall, and two on the rear wall. These were JBL models, featuring a small horn loaded high frequency driver, and a 10 inch Bass/Midrange cone driver (reflex loaded) with a frequency range of 45Hz to 18KHz. The front baffle of these units is inclined 15 degrees downwards towards the audience. The eight surround loudspeaker units were divided into two arrays of four, forming the left and right hand surround sound channels. Two Quad 405 power amplifiers, rated at 100 watts, drive each surround channel.

The projection equipment used for the experiment consisted of a Cinemeccanica Victoria 5, 35mm cinema projector fitted with a Dolby cat.700 reader, and DTStech Timecode reader. Dolby Digital data was read from the moving 35mm film print by the cat.700 unit, and was processed by a Dolby DA20 Digital Sound Adapter. The DTS timecode was also read from the moving film print, and was processed by a DTS6D Unit. To enable an instantaneous changeover between the two digital sound formats, DTS UK provided a specially modified 'changeover board' through which the six channel analogue audio signals derived from both the Dolby DA20 and the DTS6D were passed. A Dolby CP65 Cinema Sound Processor accepted the analogue inputs from both digital units, and outputs the final calibrated and equalized signal to the appropriate power amplifier and loudspeaker system.

The cinema sound system was equalized professionally by a sound engineer to exacting industry specifications a few months prior to the exper-



iment, and was continually monitored. Any degradation since was corrected by the on-site theater technician. During the experiment the master volume level was set at 7, which corresponds to a precise, equalized, Dolby Reference level of 85dBA with pink noise.

The auditorium was originally designed as a 'Debating Chamber' with a horseshoe shaped, amphitheater-seating plan to suit. Although this makes viewing and listening angles awkward with a 'full house', the smaller number of listeners in the present study allowed us to ensure that they were all seated in optimal listening positions.

Participants.

Thirty-four volunteers (5 females) took part in the study. The majority of the listeners were undergraduates at the University of Southampton. They responded to advertisements via posters, e-mail, and the University's Daily Bulletin.

Results and Discussion

Subjective report on the ability to discriminate between the two systems.

Eighteen percent (6/34) of the listeners indicated that they would be able to distinguish DD from DTS on the basis of perceived sound quality before taking part in this evaluation. The majority, 43% (15/34) had no idea whether they would be able to hear any difference and 24% (8/34) indicated that they would be unable to tell the two systems apart.

Response bias.

Each of the volunteer's responses was tabulated into 3 categories (with preference rating): Equal Preference (rating 0), Prefer Clip 1 (rating 1-3: weak to strong preference), and Prefer Clip 2 (rating 1-3). Curiously, independent of experimental condition, there was a strong tendency towards choosing Clip 2 as the response (average 47%) with the tendency of choosing Clip 1 (24%) and Equal Preference (29%) approximately the same. Furthermore, independent of experimental condition, Clip 2 received a higher preference rating (average 1.56) than Clip 1 (1.40). (By default, the preference rating for the "Equal Preference" category is zero.) Apparently the listeners did not notice such a response bias, as no one noted such a trend in the questionnaire. The bias is consistent with findings that performance in auditory detection and discrimination is typically best for the last interval or components, and worst for the first interval or components (e.g., Johnson, Watson, & Kelly, 1984; Leshowitz & Cudahy, 1973; Watson, Wronton, Kelly, & Benbassat, 1975). This fact might reflect limitations in central processing, such as memory and attention. Given that comparisons of an A versus B nature have this inherent problem, presenting longer sound stimuli is unlikely to improve differential sensitivity towards perceived sound quality. In fact, a larger quantity of shorter sound stimuli may be more useful, in addition to suiting the needs of SDT analysis very well.

Signal vs. noise; hits vs. false alarms.

Under the SDT paradigm the trials can be classified into two categories: (1) signal trials: clips within a pair were recorded by different systems (i.e., DTS-DD and DD-DTS conditions), and (2) noise trials: clips within a pair were recorded by the same system (i.e., DTS-DTS and DD-DD conditions). SDT is designed to assess individual's sensitivity (d') and decision criteria (ß) when they are attempting to discriminate/decide if a 'signal' is present within a background of 'noise'. Hence, choosing Clip 1 or 2 as a response was a Hit (H) for the signal trials, but a False Alarm (FA) for the noise trials. SDT compares the distributions of these signal and noise trials in order to statistically analyse the degree to which observers really can discriminate between the two.

Individual sensitivity.

We wished to identify any listeners who could distinguish between the two systems on the basis of perceived sound quality. According to SDT (Green & Swets, 1966), independent of response bias, the Hit rate is expected to be the same as the FA rate if the observer can hardly sense the difference between signal and noise. In other words, mere guessing would show no systematic effect in favour of either DD or DTS. Table 1 presents, for each observer, the Hit rate, FA rate and the Chi Square (c^2) value for evaluating whether or not the Hit rate differed from the FA rate. c^2 is a statistical test that permits us to say with a specified level of confidence (say, 95%) whether someone can indeed tell the difference between the two systems, or whether the result could have arisen due to random error. Given one degree of freedom and an a level of .05, the critical c^2 value to be compared with is 3.84 (see table). In other words, to provide a mere 5% chance of achieving a result due to random error, the c^2 value must meet or exceed 3.84. Clearly, none of the listeners could be considered as being able to detect the difference in sound quality between DTS and DD based on this analysis.

For a more fine-grained analysis, we examined whether there was a tendency towards selecting sound clips recorded by a particular system for the Hit trials. The two systems would be equally likely to be selected if no difference could be detected. For the Hit trials, the probability of preferring sound clips recorded by DTS, P(DTS/Hit) is presented for each observer in Table 1 (at end of this paper). Note that the probability of preferring sound clips recorded by DD, P(DD/Hit), is simply
1 - P(DTS/Hit). Also presented in Table 1 is the c^2 value for each observer testing whether there was a preference bias. Only Observer 4 revealed a significant preference bias toward one system (i.e., DD). The majority of the participants (97%) could not distinguish the two systems, and none of the six participants (indicated by # in Table 1) who thought they would be able to discriminate the two systems at the outset of the experiment showed a sign of being able to do so.

Preference rating.

The above analyses focused on the individual level by examining H and FA. We now turned to the group level by examining preference ratings. A statistical test demonstrated that there was no statistically significant difference in preference rating between signal trials (mean=1.12) and noise trials (1.11), paired $t(33)$ = 0.199, p » .8. However, mean preference rating for H (1.60) was slightly higher than it was for FA (1.53); preference rating for DD (mean = 1.64) was higher than it was for DTS (1.49). In other words, listeners were providing a higher preference rating on H trials for DD. A Response (H vs. FA) x System (DTS vs. DD) repeated-measure ANOVA showed significant main effects of Response ($F(1,33)$ = 4.39, p » .04) and System ($F(1,33)$ = 13.93, p » .001), but no interaction effect (F < 1). Overall, these results seem to tentatively suggest that DTS and DD induce a different perceptual experience, as evidenced by listeners preference ratings. In this instance a slight preference towards DD was detected.

General Discussion and Conclusions

The different approaches (data-lossless vs. data-reduction) in reproducing cinema sounds (DTS vs. DD) provide a topical vehicle upon which to apply the principles of SDT, as well as an opportunity to validate our understanding of the mechanisms of auditory perception. Although the sensitivity measures did not suggest perceptible differences between DTS and DD, the preference ratings did. The results of the preference ratings suggested that (1) listeners tended to give higher preference rating for the signal trials than for the noise trials despite their reluctance to choose "Equal Preference" as a response, and (2) DD was more preferentially rated than

DTS. In the present study, a sensitivity measure might not be sensitive enough to detect the difference because (1) the content of the sound clips was different, (2) the response bias towards selecting the last clip in a trial and (3) in terms of normal psychophysical studies a comparatively small number of trials was used. These factors inform the design of future research that could answer the DD vs DTS debate more powerfully using the principles of SDT.

Overall there seemed to be some subtle, but perceivable differences in sound quality between DTS and DD. In general, listeners rated DD more favourably than DTS. Taken at face value this seems odd. How could the quality of sound reproduced from a reduced data set (i.e. DD) exceed that for a more or less intact data set (i.e. DTS)? While recognizing the limitations imposed by the current experimental paradigm, it is still possible to speculate on this apparent paradox. Differences in perceived sound quality could be explained with reference to inherent unwanted noise in the original recorded sound. As an example, imagine someone is whispering in the background when you are attending a lecture. Although the whisper is barely audible, its mere presence can sometimes make the lecture a little difficult to follow or even induce unpleasant feelings. Consider sound reproduction. No matter how perfect a sound is, it is often accompanied by extraneous background and harmonically related noises. A data-lossless recording not only records the sounds intended, but also the extraneous noises; whereas in a data-reduction recording the noise is more likely to be removed from the intended sounds. However, removing possible extraneous noise introduces its own unwanted (quantisation) noise. To take care of the latter signal distortions, the data-reduction approach again applies the predictions of auditory masking to ensure that the desired sounds can cover up the noises related to quantisation. This two-stage process of noise removal and noise masking in DD might be contributing to the listener's perceptual experience of subjective sound quality when compared to DTS. But in whatever case, this paper does show that the quantity of digital data used to describe an audio signal is not necessarily a reliable metric for perceived sound quality, and that the real life application of psychoacoustics in cinema sound provides excellent validation of 140 years of work in the field of psychophysics.

As an end note it should be emphasised that Dolby and DTS are both to be congratulated on designing excellent systems that bring cinema sounds to cinema audiences, as the film director intended. The first objective of this experiment was to apply SDT methodology within this domain, as opposed to offering the last word in sound quality differences between the two systems. In this respect it is the authors' view that this paper succeeds in its purpose, by offering research into differential sensitivity towards perceived sound quality a novel and powerful tool.

Note 1

It would of course be ideal if the same sound sample were used for each trial. This would undoubtedly increase the listener's sensitivity towards sound quality differences. However, with time and technical limitations, such an ideal control was difficult, if not impossible to achieve. First of all, cinema film projectors only run in one direction (i.e., forward). Rewinding the film requires unthreading the film from the projector, rewinding by hand, re-threading, and then playing it again. This is a very awkward procedure that is difficult to arrange and time consuming. Alternatively, one could use two copies of the same film, where each copy is played by a separate projector and the attendant digital sound equipment. However, only a handful of cinemas are so equipped, and were not available for us to use. Cutting or splicing two copies of the same film so that the sound clips could repeat themselves every 15 seconds (the duration of sound clips used in the present study) was also considered. Due to the high cost of a cinema film it is understandable that the owners do not permit them to be treated in this way. The first author further pursued the possibility of cutting and splicing by contacting MGM to obtain old obsolete films that were to be disposed of. Unfortunately, all old/worn out films are destroyed, and using worn out film with attendant inconsistencies is not ideal in any case. Also considered was cutting/splicing copies of inexpensive film trailers to obtain an identical sound sample for each trial. Because a film trailer only lasts 2 minutes, 15 to 20 copies would be required for 40 minutes of sound stimuli. Not all film trailers are DTS DD encoded. For those that are, obtaining up to 15-20 copies of DTS trailer discs would be difficult. It would also be problematic for the projectionist to arrange the discs because a DTS decoder can only hold 3 discs at a time. More importantly, film trailer soundtracks are mixed and recorded somewhat differently from the actual film itself. Given the constraints, matching the content nature of sound samples was perhaps the best we could do in the present study.

Acknowledgements

The assistance of Mr Mike Rendall and Union Films is gratefully acknowledged.

The authors wish to thank DTS UK for providing a specially modified circuit board that enabled an instantaneous changeover between the two digital sound formats (DTS and DD).

References

APT. (1998). Technical Overview: Apt - x 100 Digital Audio Data Compression System. Belfast: Audio Processing Technology Ltd.

Davidson, G. A., Fielder, L.D., & Link, B. D. (1994). *Parametric Bit Allocation in a Perceptual Audio Coder.* Paper presented at the 97th convention of the Audio Engineering Society Inc.

Dolby. (1999). Frequently Asked Questions about Dolby Digital [On-Line].
Available: http://www.dolby.com

Eisler, H. (1966). Measurement of Perceived Acoustic Quality of Sound Reproducing Systems by Means of Factor Analysis. *Journal of the Acoustical Society of America, 39(3),* 484-492.

Fletcher, H. & Munson, W.A. (1937). Relation between loudness and masking. *Journal of the Acoustical Society of America, 9*, 1-10.

Gabrielsson, A., & Sjorgren, H. (1979). Perceived Sound Quality of Sound Reproducing Systems. *Journal of the Acoustical Society of America, 65(4),* 1019-1033

Gabrielsson, A., Rosenberg, U., & Sjogren, H. (1974). Judgements and Dimension Analyses of Perceived Sound Quality of Sound-Reproducing Systems. *Journal of the Acoustical Society of America, 55(4),* 854-861

Green, D. M., & Swets, J. A. (1966). Signal detection theory and psychophysics. New York: Wiley.

Jesteadt, W., Bacon, S.P., & Lehman, J.R. (1982). Forward masking as a function of frequency, masker level, and signal delay. *Journal of the Acoustical Society of America, 71*, 950-962.

Johnson, D. M., Watson, C. S., & Kelly, W. J. (1984). Performance differences among the intervals in forced-choice tasks. *Perception & Psychophysics*, 35, 553-557.

Kallman, H.J., & Morris, M.D. (1984). Backward recognition masking as a function of mask presentation. *Perception & Psychophysics, 35*, 379-384.

Leshowitz, B. & Cudahy, E. (1973). Frequency discrimination in the presence of another tone. *Journal of the Acoustical Society of America*, 54, 882-887.

Sun, P. L. (1998). DTS Digital Surround vs. Dolby Digital: Issues and Impressions [On-line].
Available:
http://www.moviesoundpage.com

Watson, C. S., Wroton, H. W., Kelly, W. J., & Benbassat, C. A. (1975). Factors in determination of tonal patterns. I. Component frequency, temporal position, and silent intervals. *Journal of the Acoustical Society of America*, 57, 1175-1185.

Table 1. Hit Rate (P(Hit)), False Alarm Rate (P(FA)) and $\chi^2_{(1)}$ Value Testing P(Hit) = P(FA). And Probability of Preferring Clips Recorded by DTS (or DD) Given the Hit Trials (P(DTS / Hit) or P(DD / Hit)) and $\chi^2_{(1)}$ Value Testing P(DTS / Hit) = P(DD / Hit).

	P(Hit) vs. P(FA)			[c]P(DTS / Hit) vs. P(DD / Hit)	
[a]Observer	P(Hit)	P(FA)	[b]$\chi^2_{(1)}$	P(DTS / Hit)	[b]$\chi^2_{(1)}$
1	0.55	0.68	0.29	0.55	0.09
2	0.70	0.58	0.49	0.50	0.00
3	0.75	0.74	0.00	0.40	0.60
4	0.75	0.79	0.02	0.20	5.40*
5	0.55	0.53	0.01	0.64	0.82
6	0.60	0.74	0.28	0.75	3.00
7	0.60	0.84	0.81	0.50	0.00
8	0.75	0.74	0.00	0.47	0.07
9	0.70	0.68	0.00	0.29	2.57
10	0.70	0.68	0.00	0.71	2.57
11	0.90	0.89	0.00	0.50	0.00
12	0.60	0.79	0.52	0.50	0.00
13	0.45	0.74	1.39	0.33	1.00
14#	0.60	0.58	0.01	0.42	0.33
15#	0.70	0.68	0.00	0.43	0.29
16#	0.60	0.84	0.81	0.42	0.33
17#	0.85	0.79	0.04	0.71	2.88
18	1.00	0.95	0.09	0.45	0.20
19	0.75	0.84	0.11	0.53	0.07
20	0.55	0.68	0.01	0.64	0.82
21	0.85	0.84	0.00	0.47	0.06
22	0.75	0.89	0.25	0.53	0.07
23	0.70	0.84	0.26	0.64	1.14
24	0.90	0.84	0.04	0.67	2.00
25	0.75	0.53	0.78	0.73	3.27
26#	0.50	0.68	0.57	0.60	0.40
27	0.60	0.79	0.62	0.58	0.33
28	0.70	0.68	0.00	0.57	0.29
29	0.75	0.74	0.00	0.53	0.07
30	0.70	0.68	0.00	0.50	0.00
31	0.60	0.68	0.18	0.56	0.25
32	0.75	0.68	0.06	0.40	0.80
33	0.60	0.58	0.01	0.75	3.00
34#	0.55	0.53	0.01	0.45	0.09

a. #: observer indicated that he/she was able to discriminate the two systems in the questionnaire

b. $\chi^2_{(1, .95)} = 3.84$.

c. P(DD / Hit) = 1 - P(DTS / Hit)

*: $p < .05$