

# MI

MOTION IMAGING JOURNAL

HEG MEDIA TRANSPORT

ENCODING EFFICIENT

BROADCAST SIGNALING

VA NMO5 specification

Fast Metadata Framework

## CASE STUDY

KEYWORDS CINEMA SOUND // LOUDNESS

Você já foi a uma sala de cinema e sentiu um desconforto por achar que o som estava alto demais? Pois é o artigo desta edição se fez a mesma pergunta ... A partir de medições em centenas de salas comerciais e testes controlados com o público, o estudo analisa como níveis de reprodução, trailers e picos sonoros impactam diretamente o conforto e a satisfação da audiência. O autor revisita um dilema de décadas com uma abordagem inovadora: em vez de confiar apenas em métricas objetivas como o Leq (Nível de Som Contínuo Equivalente, uma média de energia sonora ao longo do tempo), ele analisa a correlação direta entre os ajustes de Fader (controle de volume do processador de cinema) e a real satisfação do público. E assim define um novo *Comfort Threshold* (Limiar de Conforto, ou nível de reprodução onde as reclamações do público são minimizadas), revelando uma discrepância chocante nas operações atuais. Enquanto o *Reference Level* (que é o padrão industrial de calibração geralmente fixado em 7dB) busca a fidelidade artística, a realidade mostra que trailers frequentemente exigem uma redução drástica de até -11.4 dB para serem aceitáveis. Mais surpreendente ainda: o estudo derruba o mito de que baixar o volume global prejudica significativamente a inteligibilidade dos diálogos. Sem a automação correta, os cinemas vivem um “compromisso” destrutivo onde trailers ensurdecem ou o filme perde seu impacto. Bora ler o artigo e ajustar nossos níveis?

Tom Jones Moreira  
([tvdigitalbr@gmail.com](mailto:tvdigitalbr@gmail.com))



## Cinema Sound Playback Level and Audience Satisfaction

By Charles Q. Robinson



Results suggest that the fader settings in commercial cinemas provide a reliable measure of the discrepancy between program loudness and the levels acceptable to most of the cinema-going public. By this measure, trailer and feature levels are too high, with trailer program levels exceeding acceptable levels by more than 10 dB.



### Abstract

Over 25 years ago, Ioan Allen presented a paper asking, “Are Movies Too Loud?” He described the then-current situation, highlighting many issues and symptoms, including audience complaints, concerns reported in the press, theater operators playing films below Reference Level, trailer loudness wars, and excessive trailer loudness pushing down feature playback level. These issues remain decades later despite some progress (notably the Trailer Audio Standards Association [TASA] standard). Previous investigations focused on deriving objective loudness metrics (variations on weighted Leq) and associated thresholds. For this paper, we take a different approach; we use a subjective loudness metric and measure audience satisfaction to establish a Comfort Threshold for feature and trailer soundtracks. We present research in real-world and controlled environments to explore the causes and consequences of reduced playback levels and propose changes to current practice.

The superior experience of watching and listening to a movie in a cinema is made possible partly by the cinema ecosystem—a set of recommendations and standards that enable accurate translation from mixed stage to cinema and from cinema to cinema. Unfortunately, the goal of uniform presentation often conflicts with differences in playback environment and customer preferences, particularly regarding sound levels. Many movie-goers find cinema audio uncomfortably loud. In response to customer feedback, cinema operators rarely run programs at the industry standard Reference Level. However, the situation remains unsatisfactory to audiences, theater operators, mixers, and directors. Playing a movie below the intended level does more than reduce loudness.

The timbre, the spatial balance, and the balance between the audio elements within the mix are distorted. With a loss of Reference Level, directors can no longer be confident that their vision and artistry will be faithfully reproduced in theaters, moviegoers are less engaged and remain unsatisfied with playback levels, and exhibitors are caught in the middle.

For this study, we objectively measured the playback

level, noted the sound processor fader setting, and conducted a short survey of on-site operators to get their opinions on sound level issues at hundreds of commercial cinemas.

Separately, we conducted controlled screenings in a small, capable, well-calibrated cinema at various fader settings, ranging from Reference Level to 10 dB below. Exit surveys probed moviegoers’ opinions on sound levels and engagement.

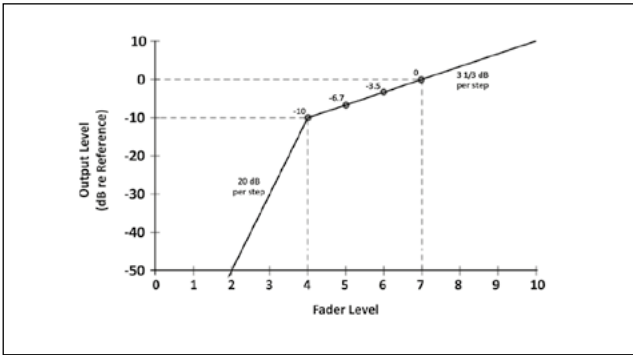
The results from this comprehensive approach suggest that the fader settings in commercial cinemas provide a reliable measure of the discrepancy between program loudness and the levels acceptable to most of the cinema-going public. By this measure, trailer and feature levels are too high, with trailer program levels exceeding acceptable levels by more than 10 dB. In this paper, we provide a detailed assessment of the scope of the problem and offer evidence-based recommendations for improvement.

The remainder of this paper is organized as follows. The background information provides the necessary context. The “Subjective Assessment Levels of Playback Levels” section describes the experiments conducted to better understand and assess movie-goer impressions of playback levels in cinemas. The next section describes a field survey of commercial cinemas to determine existing practices regarding cinema content playback (ads, trailers, features). In the final section, we summarize our conclusions and recommendations.

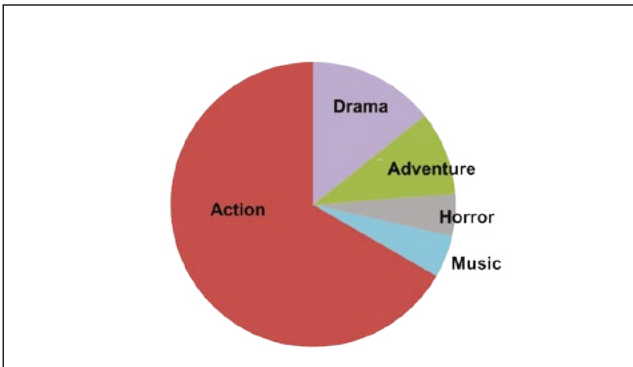
### Background

The movie soundtrack for each theater is reproduced by a cinema audio processor, which is calibrated during installation to meet industry standards for frequency response and sound level. The sound level for each auditorium is set with a volume control, commonly referred to as the fader. Many cinema processor faders use a 10-point scale, with Fader 7.0 established as the Reference Level. The following discussions describe the playback level setting using the fader value or the offset in dB from the Reference Level. The relationship between the 10-point fader scale and level is shown in **Fig. 1**. The procedure for measurement and calibration of playback level is defined by SMPTE RP 200,<sup>1</sup> and has been standardized internationally as ISO 22234.<sup>2</sup>

Most cinemas utilize a Theater Management System (TMS) to control many aspects of theater operation, including creating a playlist and initializing the playback of the playlist elements—commercials, trailers, and features—at each screen



**FIGURE 1.** Plot illustrating the 10-point Fader scale. The output level is shown as a function of Fader setting.

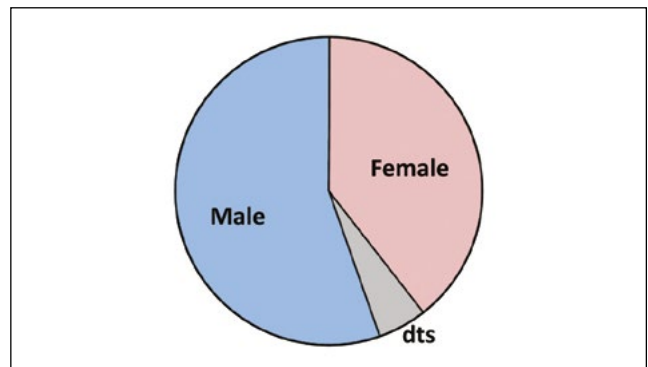
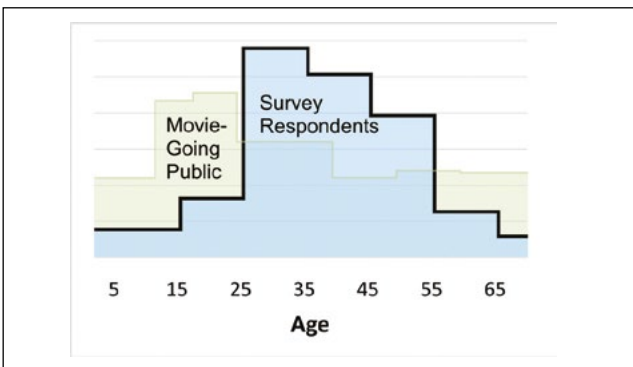


**FIGURE 2.** Distribution of movie genres included in the survey.

for each show time. The TMS can also be programmed to automatically alter the sound levels during the playlist.

For decades, it has been widely recognized that cinemas rarely reproduce cinema sound at Reference Level.<sup>3-10</sup> This practice has been adopted due to guest complaints about excessive sound levels. The selection and sophistication of equipment and operations policy enable a varying degree of level control. The modes of control include the following:

- Individual control of the playback level across content types: ads, trailers, and features.
- Individual control of the playback level from room to room.



**FIGURE 3.** The demographics of survey respondents: left plot shows the distribution of ages of our survey respondents and of the movie-going public.<sup>11</sup> The gender of our subjects is shown on the right. The gender of movie-goers is nearly evenly split between male and female<sup>11</sup> “dts” indicates no response (decline to state).

- Control of playback levels across time. Specifically, changes in response to complaints may be temporary (current screening or current playlist) or permanent (until the next guest complaint).

We will describe how these types of control can affect playback levels and audience satisfaction.

**Subjective Assessment of Playback Levels**

In the first part of this study, we conducted an exit survey to measure moviegoers impressions of movie sound levels. The data collected for this study is based on “Employee and guests” screenings hosted at the Dolby Potrero Cinema, a well-calibrated and maintained theater in San Francisco. The room is relatively small, with variable seating of up to approximately 100 seats. For a year, we screened 21 titles and collected over 700 responses. Many titles were screened multiple times to different audiences. Various genres were included; most were “Action” movies (Fig. 2). The subjects voluntarily attended the screening and took the survey. The moviegoers were not informed or recruited in advance. Some details of the subject demographics are shown in Fig. 3.

Respondents were asked to assess three aspects of the movie’s loudness—the overall loudness, the loudness peaks, and dialog loudness, and respond using a 5-point Likert scale. The survey was implemented as a touch-screen app on a computer tablet (Fig. 4).

**Results**

While a majority (61%) found the overall level to be about right, the distribution is biased toward “too loud,” with very few/none responding slightly to quiet/too quiet, respectively. The subjects were less satisfied with the sound level of the loudest scenes. Over half (52%) found the peak level too high.

In contrast, most subjects (88%) were satisfied with the dialog level. We will return to the dialog level later.

When a movie title was screened to more than one audience, we presented the movie in various fader settings to measure subjective response to different sound levels. The results for the overall and peak levels are shown in Fig. 6. The fader settings are pooled into four ranges based on the 10-point fader scale described previously: 7.0 (Reference

**FIGURE 4.** The user interface for sound level exit survey.

	Overall Level	Peak Level	Dialog Level
Too Loud	8	20	2
Slightly Too Loud	29	32	4
About Right	61	46	88
Slightly Too Quiet	1	1	6
Too Quiet	0	0	1

**FIGURE 5.** Subject response for the overall level, peak level, and dialog level after attending a screening was presented at the reference level.

Level), 6.x (0.1 to 3.33 dB), 5.x (3.34 to 6.66 dB), and 4.x (10 dB).<sup>i</sup>

For both overall and peak levels, the subject responses reflect the playback level changes.

As the level is reduced from the Reference Level,

- There are fewer responses on the loud side of the scale and
- more on the quiet side of the scale.

Also,

- Maximum satisfaction with the *overall* level (largest number of “About Right” responses) is achieved with a slight reduction in playback level (~2 dB). Satisfaction with the overall level remains constant with further reductions (2 to 10 dB) in playback level;
- Satisfaction with the *peak* level increases continuously as the level is reduced, reaching the maximum number of satisfied listeners at a fader setting of 4.x (-10 dB!).

### Discussion

Observation #1: We have shown that level preference ratings for audiences attending a movie at their leisure reflect playback level differences as small as 3.3 dB. A larger sample, e.g., all the patrons that visit a commercial cinema day after day, would likely have much higher resolving power, perhaps approaching the just noticeable level difference of 0.77 to 1.24 dB for typical program material as reported in Ref. 12.

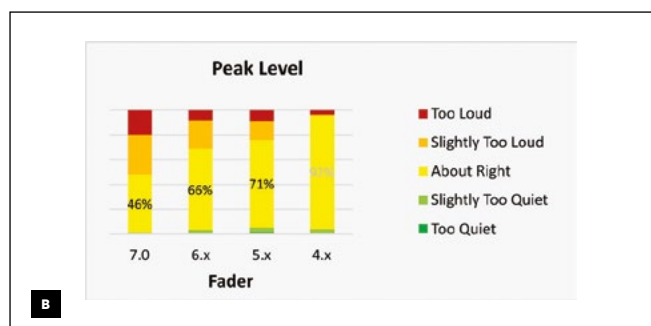
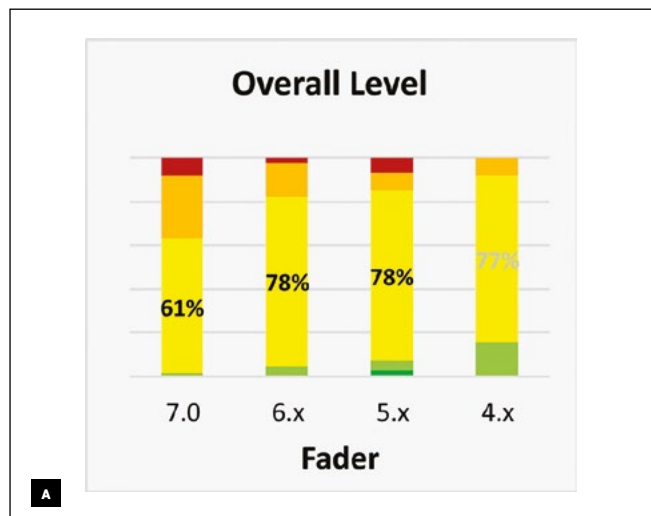
<sup>i</sup>Few movies were presented at less than 5.0, so the statistics associated with the 4.x fader setting are less reliable and should be viewed with caution. All the 4.x screenings used a fader setting of 4.0.

Observation #2: The discrepancy between listener satisfaction with peak level (too loud) versus overall level (about right) suggests that for these listeners and these titles (major U.S. productions), program loudness is not as significant an issue as excessive peak levels.

Action movies routinely use the full scale of the digital audio format. As reported by others,<sup>7,13</sup> an audience member’s impression and dissatisfaction with peak level is more likely the result of sustained loud scenes (e.g., a car chase or battle) rather than an occasional sound event (e.g., gunshot or car crash). Short, dynamic sound events are associated with “punchiness”<sup>14</sup> (generally a desirable trait), while extended loud scenes are more fatiguing. Unfortunately, regardless of the exact nature of the dynamics and annoyance, the only correction an exhibitor can make to excessive peak levels is to turn down the entire feature. (At home, we can “ride the remote” to manually reduce the dynamic range.)

### About dialog ...

Dialog level and intelligibility in the home and the cinema are hot topics, with active investigations by SMPTE and the Audio Engineering Society. Many issues influence dialog intelligibility in cinemas, which is out of the scope of this paper. However, the results of this study strongly suggest that absolute dialog level is not among the issues. (A low dialog level relative to other elements in the mix, a “close” mix, may be



**FIGURE 6.** Subject response to feature loudness as a function of fader setting for (a) overall and (b) peak level. Movies in this study were rarely presented at less than 5.0. As a result, the response statistics shown for the 4.x playback level are somewhat suspect.

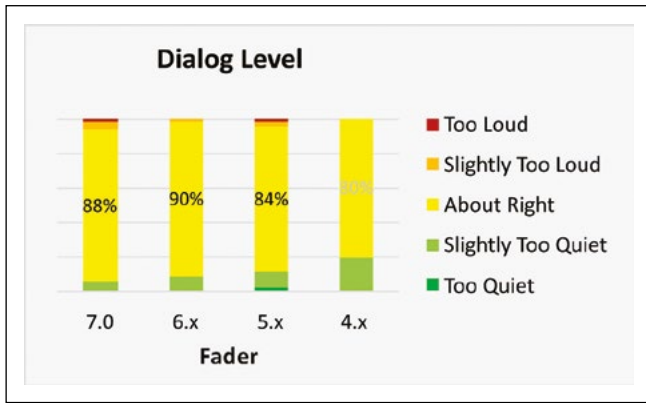


FIGURE 7. Subject response to dialog loudness as a function of fader setting.

an issue.) As seen in Fig. 7, listeners are very satisfied with the dialog level even as the fader is reduced to 4.0.

This is perhaps one of the more important revelations of this investigation as it is contrary to an often-voiced concern that reducing playback levels will negatively impact dialog level and intelligibility. This good news discovery, if true, also has an unfortunate corollary: one cannot count on complaints about dialog intelligibility to provide a lower limit on playback levels in cinemas. In the end, patrons rarely complain that playback levels are too low.

**Field Survey**

In a second independent study, we conducted a field survey of commercial cinemas to collect data on current playback environments and practices. The survey consisted of a short questionnaire for the on-site manager, objective measurement, and observation of the facility. Our survey spanned over 150 screens, 17 cinema complexes, and three metro areas in the western U.S.

A histogram of playback levels for trailers and features is shown in Fig. 8.

Few are likely surprised that trailers are played well below Reference Level, with a setting in the 4.x range the most common. Surprisingly, at some screens, the trailer playback level is more than 20 dB below reference. (Let’s hope they have automated the fader on these screens to enable a higher playback level for features!—Spoiler: they do.) A comparison of trailer setting and room size (using seat count as proxy) proves revealing (Fig. 9). While the trailer playback level is mostly independent of seat count, the lowest playback levels occur in cinemas with few seats.

No prior public studies are known to have included direct observation and tabulation of fader settings. In 2017, John Fithian, then president and CEO of the National Association of Theatre Owners (NATO), conducted a “rough, anonymous survey” asking cinema companies for their best estimate of the fader settings used in their theaters.<sup>8</sup> In Fig. 10, we compare the results of the cinema owner’s responses with the data from this survey (re-binned to match Fithian’s data format). We see the same trend: the fader is set lower for trailers than features. The Fithian results show fader settings for trailers and features at or above five significantly more often

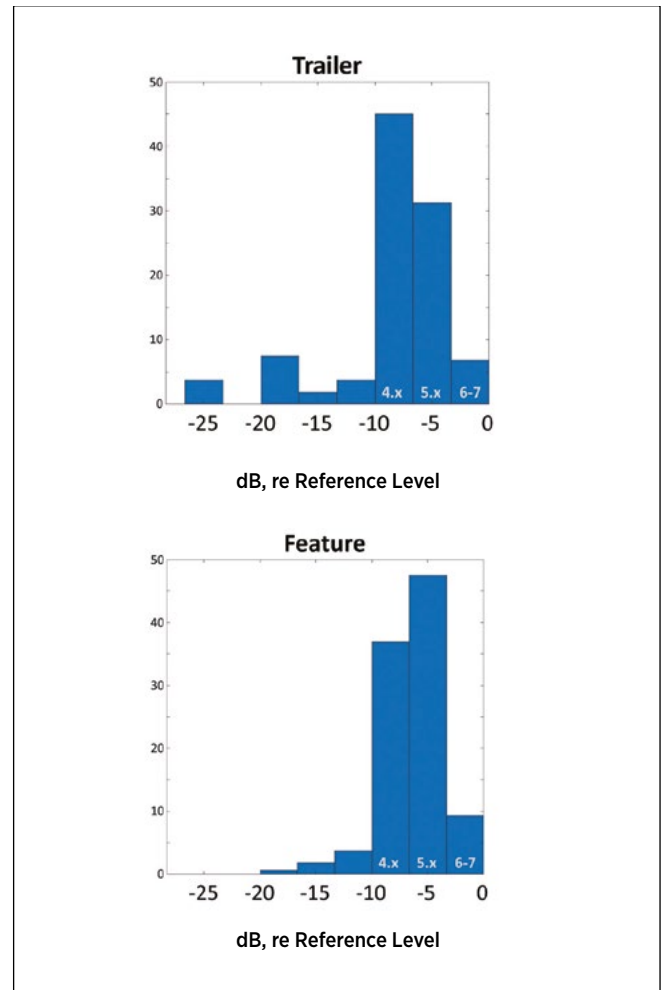


FIGURE 8. Histogram of fader settings for trailer and feature playback. The frequency of playback within eight ranges of playback level is shown in percent. The right-most bar in both plots represents fader settings of 6.0 up to 7.0 (Reference Level).

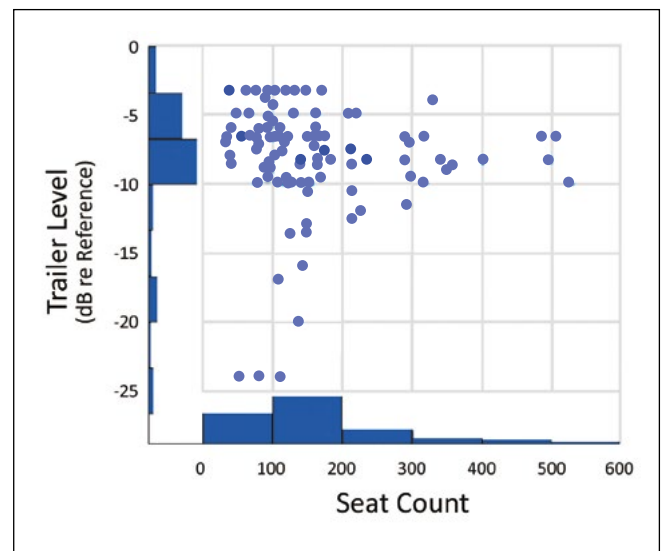
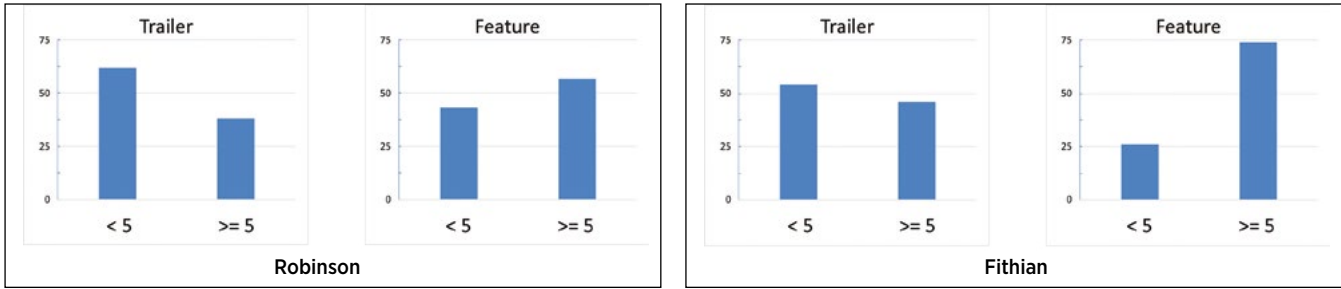


FIGURE 9. Plot of trailer playback levels as a function of auditorium seat count. (Note: this plot includes only those screens for which seat count data was collected, so the y-axis histogram does not exactly match the trailer histogram of Fig. 8.)



**FIGURE 10.** Fader setting data from this study compared to prior survey of theater owners by John Fithian. The y-axis indicates the frequency with which faders are set below Fader 5, and at or above Fader 5 (-6.7 dB re Reference).

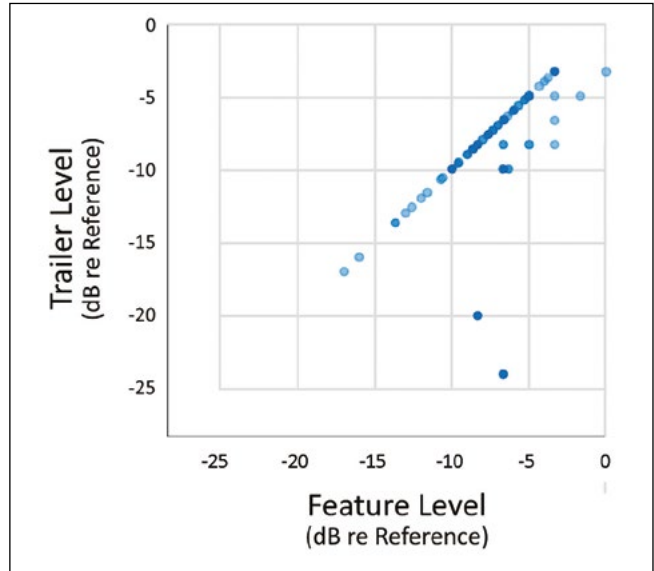
than we found in this survey. This may reflect an aspirational outlook from the cinema companies.

**Comfort Threshold (CT)**

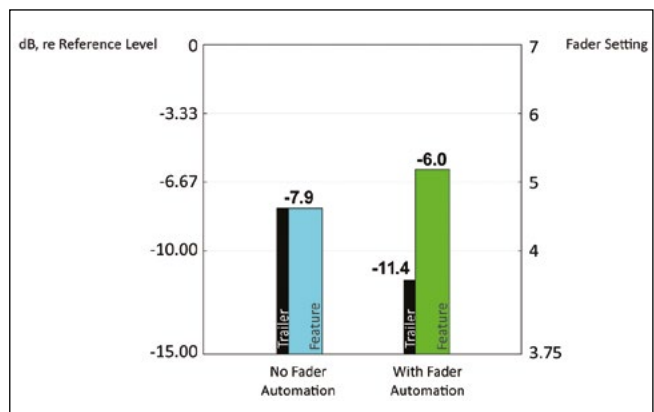
In the first part of this study, we showed that guests readily perceive and properly assess the sound levels in the cinema. In the second part, we observed the playback levels in use in cinemas. The playback levels are typically derived as follows: the playlist for a given show (ads, trailers, features) is presented, a dissatisfied guest responds, and the playback level is reduced. Over time, further presentations and adjustments are made until there are fewer complaints.

In essence, cinema operators have used a technique called *method of limits*<sup>15</sup> to measure cinema guest playback level satisfaction or comfort thresholds. We can quantify the results of this inadvertent measurement regime by taking an average of the derived fader settings. We call this measured value the audience Comfort Threshold (CT), expressed as either dB re Reference Level or using the 10-point fader scale.

The data presented so far (Figs. 8 & 10) suggest that the audio level for features is typically higher than for trailers. In fact, in most cases, 62.5% of the auditoria surveyed in this study, audio levels are not adjusted during the playlist presentations, and the trailer and feature levels are the same. This is illustrated in Fig. 11, which shows trailer level as a function of feature level. Each dot represents one cinema. The dominant diagonal line represents all the screens with equal trailer and feature playback levels. The points to the right of the line indicate screens where the feature level is higher. This suggests that for most screens, a compromise is being made to find a single fader setting suitable for trailers and features. To determine the CT for trailers and features independently, we need to look at the average fader settings in those cinemas that independently control levels. This is shown in Fig. 12, which delivers an important message. The CT for trailers is achieved with a fader setting of more than 10 dB below reference. The CT for features is achieved with a fader setting of just over 5. The fader setting required to meet the CT for features is more than 5 dB higher than for trailers. Comparing the independent CT fader settings with the compromise settings when automation is not used, we can see that using a single setting for trailers and features is a compromise: the trailers are too loud, and the feature is not loud enough.



**FIGURE 11.** Trailer vs. Feature fader settings. The x- and y-axis represent the feature and trailer playback levels in dB re Reference Level.



**FIGURE 12.** Average level setting with and without fader automation. In each case, the narrow black bar and the wider bar show the average trailer level and the feature level settings, respectively.

As part of our survey of on-site managers, we asked how often they receive complaints. Most reported that complaints were rare, indicating that the fader settings at each screen were at or near the CT. However, Fig. 13 shows that complaints, while rare, are significantly more common on the compromised screens, those that do not take advantage of fader automation.

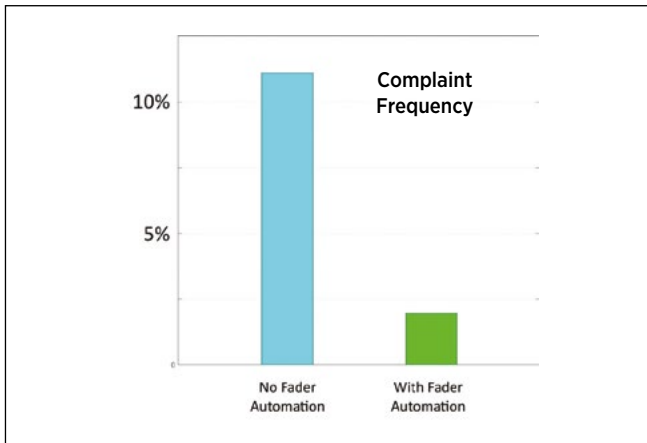


FIGURE 13. Portion of cinemas that receive complaints “Often.”

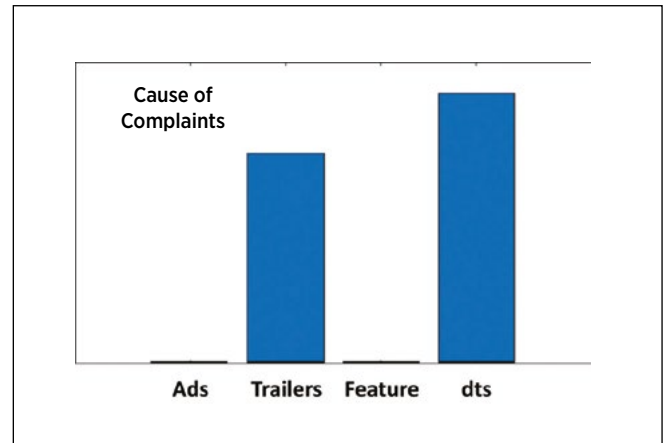


FIGURE 14. Source of complaints according to cinema managers. “dts” indicates no response (decline to state).

We also asked the on-site managers what causes the most complaints. Their answers are shown in Fig. 14. In most cases, the managers did not identify a specific source because they didn’t know or were not inclined to answer. Of those who responded, 100% identified trailers as the source of complaints. This is consistent with our results to this point. Features are generally presented at or below the CT. Trailers are most commonly presented above the CT. Our data does not clearly explain why ads do not seem to generate complaints. The playback fader setting for ads is the same as for trailers in almost all cinemas we surveyed.

**Conclusion**

The experiments described above lead to the following conclusions:

- Moviegoers are mostly satisfied with the *overall* loudness of movies played at the Reference Level.
- Moviegoers are mostly dissatisfied with the *peak* levels of movies when played at the Reference Level.
- Moviegoers’ satisfaction with the dialog level is largely unaffected by the playback level.
- The Comfort Threshold, the playback level required to reduce complaints to an acceptably low rate of occurrence, are as follows:
  - Trailer Comfort Threshold = -11.4 dB re Reference Level
  - Feature Comfort Threshold = -6.0 dB re Reference Level

There are several important caveats to this research, particularly the derived CT.

The CT is likely accurate and applicable to trailers. Trailers are monitored and controlled by the TASA-defined measurement and limit<sup>16,17</sup> and are consistent from trailer to trailer. Feature sound, on the other hand, can and does substantially vary from title to title. The peak levels of action movies likely dominate the CT for features.

The size of the cinema auditorium likely influences the CT: some small venues had very low playback level settings. More data is required to quantify this effect adequately.

It is important to note that we did not include any family/children’s movies in our exit surveys and could not measure



FIGURE 15. Movie rating as a function of fader setting averaged across all titles.



FIGURE 16. Movie rating as a function of fader setting for one critically acclaimed title.

the influence of family films in the field survey. Therefore, the CT for family-oriented movies may be lower than for movies in general.

Finally, it is important to note that Comfort is not the only consideration when establishing the optimal playback level setting. We have shown that reducing the playback level increases audience comfort and satisfaction with cinema loudness. However, reducing the playback level also changes the soundtrack mix, spatial balance, timbre, and impact. Intuitively, one may expect that a well-mixed movie becomes less engaging as the playback level is reduced. This is a legitimate

concern for directors; reducing playback levels interferes with the objective of delivering creative intent.

To shed some light on this, let's briefly revisit the results from the exit survey. One of the prompts on the survey was to "Rate the Movie" on a 5-point Likert scale ranging from "Bad" to "Excellent." The results expressed for playback level are shown in **Fig. 15**.

On average, across all movies screened, listeners liked the movies more when presented with a slight reduction of level (~2.5 dB). This is in line with guest response to loudness: a slight reduction in increased listener satisfaction with the loudness and a greater appreciation for the movie. Correlation does not ensure causation, but it is consistent with the idea that listeners are more likely to enjoy the movie if they are comfortable and not annoyed by the playback level. Larger reductions in playback level did not further improve their appreciation of the movie.

The reviews for one particular title, a critic and audience favorite that won multiple Oscar awards, including Best Picture, Best Sound Mixing, and Best Sound Editing, are shown in **Fig. 16**.

In this case, a slight reduction in playback level results in a slight improvement in movie ratings, but further reductions have an increasingly adverse effect on them. This title is best presented at or near the Reference Level.

### Recommendations

The TASA standard has significantly improved the trailer sound level. Trailer levels are now much more uniform than they once were, but they are still too high. To play trailers at Reference Level while minimizing guest complaints, the TASA limit would need to be reduced by approximately 11 dB to 74 dB LeqM. An alternative, more moderate, and useful change would be reducing the TASA limit by 5 dB to 80 dB LeqM. This would align the trailer and feature CT, allowing a single fader setting for both features and trailers without compromise. Without this change, the sound level for features in many cinemas will be dragged down to mitigate customer discomfort during playback of trailers.

Failing a reduction in the TASA limit ...

Cinema operators are strongly encouraged to automate their fader-level settings if they have not done so already.

With automated settings, trailer playback levels can be reduced, complaints reduced, feature playback levels increased, and we can better deliver the creative intent.

### References

1. SMPTE, RP 200:2012, "Relative and Absolute Sound Pressure Levels for Motion-Picture Multichannel Sound Systems."
2. International Organization for Standardization (ISO), ISO 22234:2005, "Cinematography—Relative and absolute sound pressure levels for motion-picture multi-channel sound systems—Measurement methods and levels applicable to analog photographic film audio, digital photographic film audio, and D-cinema audio," Nov. 2005.
3. I. Allen, "Are Movies Too Loud?" *SMPTE Mot. Imag. J.*, 107(1): 30-35, Jan. 1998.
4. I. Allen, "How Loud is a Movie?—A New Measurement Procedure" *SMPTE Mot. Imag. J.* 125(4): 57-62, May/June 2016.
5. I. Allen, "A Primer on the Levels of Features, Trailers, and Commercials as Heard in the Cinema," Unpublished Technical Paper, Jan. 30, 2009.
6. R. Thom, "Are Movies Getting Too Loud?" 1996. [Online]. Available: <http://www.filmsound.org/randythom/loud-movies.htm>.
7. F. Rumsey, "Loudness Revisited," *J. Audio Eng. Soc.*, Vol. 62, No. 12, Dec. 2014.
8. J. Fithian, "Is the Cinema Sound Reference Dead?" Boxoffice, Nov. 2017.
9. Film-Tech Forum, (2002). "Volume Levels, is level 7 too high?" [Online]. Available: <http://www.film-tech.com/cgi-bin/ubb/ft/t003026/p1.html>.
10. E. Grimm. "Movie Dynamics Evaporated in Recent Years. Can we turn the tide?" AES 134th Convention. May 2013.
11. Motion Picture Association of America (MPAA). "Theatrical Market Statistics 2016." [Online]. Available: <https://www.motionpictures.org/wp-content/uploads/2017/03/2016-Theatrical-Market-Statistics-Report-2.pdf>.
12. G. Soulodre, M Lavoie, and S. Norcross, "The Subjective Loudness of Typical Program Material," AES 115th Convention, Oct. 2003.
13. C. Watson and R. Gengel, "Signal duration and signal frequency in relation to auditory sensitivity," *J. Acoust. Soc. Am.*, vol. 46, no. 4, pt. 2, pp. 898-997, Oct. 1969.
14. S. Fenton, H. Lee, "A Perceptual Model of 'Punch' Based on Weighted Transient Loudness," *J. Audio Eng. Soc.*, 67(6): 450-451 Jun. 2019.
15. G. A. Gescheider, "Psychophysics: The Fundamentals," Taylor & Francis, U.K., 2013.
16. SMPTE, RP 2054:2010, "Method of Measurement of Perceived Loudness of Short Duration Motion Picture Audio Material."
17. International Organization for Standardization (ISO), 21727:2016, "Cinematography—Method of measurement of perceived loudness of short duration motion-picture audio material," Edition 2, 2016.

### About the Author



Charles Robinson has conducted audio research in acoustics, audio coding, loudness, and spatial audio for media. He is a SMPTE Fellow and multiple Emmy winner recognized by the Academy.

