Audio Description Research Project

Investigating a Standardised Approach to Setting Audio Description Dip Values.

Kieran Lynch[1] Oran O'Neill[2] Tomás Ward[3]

kieran.lynchk@iadt.ie oran.oneill@rte.ie tomas.ward@dcu.ie

[1]Dun Laoghaire Institute of Art, Design and Technology, Dun Laoghaire, Dublin, Ireland

[2]RTE, Donnybrook, Dublin 4, Ireland

[3]Insight SFI Research Centre for Data Analytics, Dublin City University

Final Report to Coimisiún na Meán

June 2023

Table of Contents

Acknowledgments

Research discussed in this report was funded by Coimisiún na Meán (the regulator of broadcasting and online media in Ireland) as part of a research grant received through its Sectoral Learning & Development's Diversity fund. In addition, the National Council for the Blind of Ireland played a significant role in enabling the completion of this work. We express our sincere appreciation for the funding and support provided by these organisations.

Executive Summary

Audio Description (AD) production involves dipping the volume of the original programme material to facilitate speech intelligibility during the descriptive passages, while also maintaining the integrity of the original programme's narrative. Common technical standards for setting the AD Dip Value (DV) are not currently documented in industry, with each broadcast and streaming service providing separate technical recommendations to AD content creators.

In order to gain a better understanding of the issues and inconsistencies experienced by those accessing AD content, we carried out a service-user survey involving 24 participants across 5 countries. Examples of AD service issues highlighted by respondents included out-of-sync AD cues, missing AD cues, inconsistent AD cue levels, inconsistent and inappropriate Dip and Fade Values, and the mono delivery of stereo programme material. In an attempt to determine the production stages at which such issues occur, 42 professional AD content creators were also surveyed. The industry survey was carried out to further understand the roles, duties, and technical knowledge of those involved in producing content for AD services; the workflow, production practice, and DVs used; and whether a standardised set of recommendations and guidelines for setting DVs is welcome. On average, sound engineers are using a DV of -12dB and adjust this value in accordance with the programme loudness at the AD cue point. In contrast to this, non-sound engineers involved in the creation of AD content are using a wide variety of DVs ranging from +3dB to -18dB and are less inclined to adjust the DV over the programme duration. There is a strong appetite among non-sound engineer AD content creators for a standardised set of technical recommendations and guidelines on setting AD DVs for programme material of varying loudness.

This report also presents the findings from experiments used to benchmark the mixing practice of professional sound engineers when setting AD DVs. We present data that identifies the linear correlation between a programme's integrated average loudness (LUFS), its Loudness Range (LRA), and the integrated average loudness of the programme's AD narration track. This linear relationship is presented in Table 17 of Appendix 4 as a target LUFS look-up table for AD narration tracks. A *Zero DV Threshold* is also presented. That is, if the programme's measured LUFS value at an AD cue point is below this threshold value, then it is suggested that no DV be applied. For the 6 programme titles investigated here, the *Zero DV Threshold* was established at -48 LUFS. Analysis also presents the correlation between a programme's LUFS value at an AD cue point and the DV applied, yielding a set of logarithmic quantile regressor slopes used to determine target DVs.

Three programme loudness bands have also been established. The loudness bands of *Quiet*, *Normal*, and *Loud* have been determined using the normal distribution of programme loudness probability about the mean integrated loudness value of -26.6 LUFS for the 6 titles considered. The measured integrated loudness of the programme material at an AD cue point can now be categorised as either *Quiet*, *Normal*, or *Loud* allowing for specific target DVs to be applied.

9 DV conditions were assessed on an audience to determine the preferred DV for each of the programme types congruent with the *Quiet*, *Normal*, and *Loud* programme loudness bands. From our analysis, it is understood that as the integrated loudness of the programme audio increases at the AD cue points, so too does the listener's preferred Dip Value. Based on the industry survey carried out, the benchmarking of the mixing practice of professional sound engineers when setting DVs, and the listening test results, we recommend that a DV of -19 dB be applied to background programme audio that lies within the *Loud* loudness band, a DV of -9 dB be applied to background programme audio that lies within the *Normal* loudness

band, a DV of -3 dB be applied to background programme audio that lies within the *Quiet* loudness band, and that no dip be applied to programme material that has an integrated average loudness lower than -48 LUFS at the AD cue point.

Introduction

Audio Description (AD) is a form of verbal commentary used to provide detail in relation to important visual plot and character information in a media production, such as a TV Programme, for the benefit of blind and vision impaired audiences. AD describes scene landscapes, body language, expressions and movements, helping to make the programme's visual content and narrative clear through sound (see Image 1).

Image 1: AD production typically includes a video track for the TV programme, the original programme audio, and an AD narration track

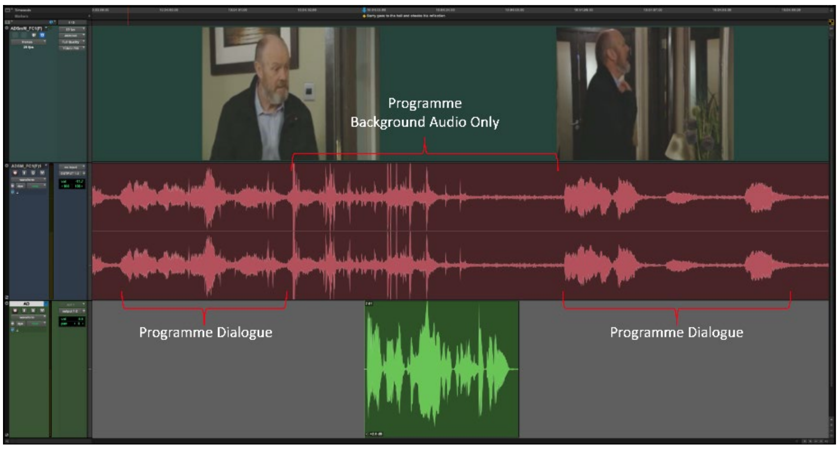


Image 2: The AD narration track is typically recorded in the gaps between programme dialogue.

During the production of AD, the narrator records descriptive passages in the gaps between the programme's dialogue (see Image 2 above). In order for the narration to be clear, the original programme material may necessitate a reduction, or dip in loudness so as to prevent masking of the audio descriptions (see Image 3 below).

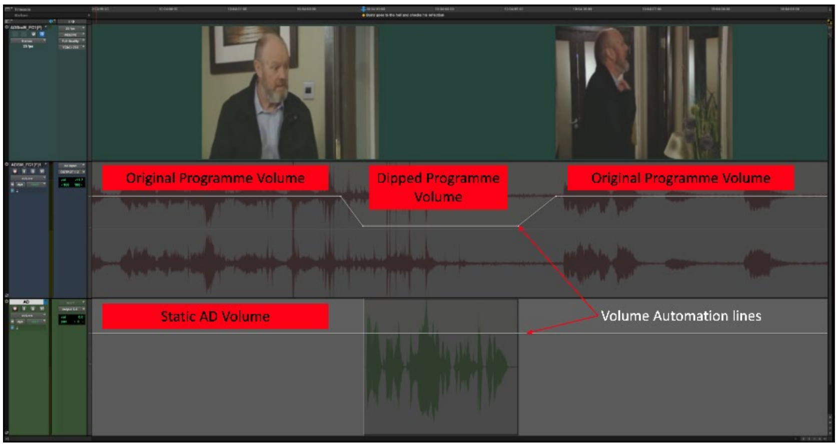


Image 3: The original programme audio may require a dip in volume to facilitate good speech intelligibility in the AD narration.

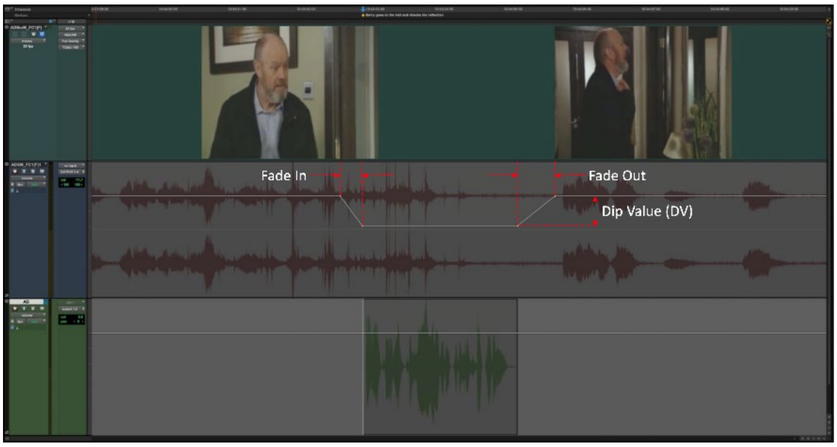


Image 4: Loud programme types will typically require larger DVs than quiet programmes.

The Dip Value (DV), or amount by which the programme material's loudness is reduced, is dependent on the programme's measured loudness value at the AD cue point (see Image 4 above). Loud programme passages, such as action or music scenes, require higher DVs than those containing quiet ambient background information, such as rural or domestic soundscapes. Currently, DVs are set by both sound engineers and non-sound engineers working as AD content creators. This has led to a wide variance observed in production practices and standards presented in commercial AD services.

This study aims to better understand the relationship between audio described TV programme loudness levels and the perceived quality of the programme's audio descriptions. The study is primarily focused on the end user experience, how this experience can be improved, and consistency brought to AD services.

This is an important area of study as it aims to assess the range of current practices on an audience, analyse and evaluate the assessment and, in an attempt to optimise the listening experience of the target audience, develop a standardised approach to setting DVs. This standardisation of the production process should benefit anyone accessing AD services as well as those wishing to bring quality and consistency to their AD content.

Background and Context

Audio Descriptions

Although AD services have been available on commercial TV broadcasts since its use at WGBH[1] in 1988[1] and the many techniques associated with the scripting and delivery of Audio Descriptions are well documented[2], there is little by way of technical standards or recommendation in setting DVs for TV programme audio of varying loudness.

AD service content is generally created in two stages:

Authoring: Is the offline process in which the AD narration is scripted, recorded and edited in segments to appropriate times in the programme title. The recorded commentary is typically delivered to the broadcaster in one of two ways.

1. The individual AD audio segments, their timecode reference, and dip and fade information are stored in an industry-standard file interchange format such as the Extended Subtitle Exchange Format (ESEF).
2. As a stereo WAV file in which the AD commentary has been pre-mixed with the original programme audio.

Broadcasting: Broadcasters have the choice to transmit the ESEF metadata and media to create the final audience AD service via user-adjustable controls in the local TV or set-top box (Receiver Mix). Or to transmit the pre-mixed version as an alternative audio channel that is user-selectable (Pre-Mix). Using the timecode reference and dip & fade information in the ESEF file, most broadcasters are capable of creating an in-house pre-mixed audio WAV file consisting of dipped program sound and its AD.

[1] WGBH is a Public Broadcasting Service (PBS) member television station licensed to Boston, Massachusetts, United States.

The broadcast stage also includes any additional audio processing (such as loudness normalisation) and delivery of the AD commentary track as part of the final broadcast channel output.

Through observations in the literature and discussions with broadcasters, it is evident that much of the broadcast content that includes AD is authored by third party service providers. It is this relationship between the content creator and the broadcaster that is under investigation here, and in particular, the processes, standards, recommendations, and guidelines used when setting and applying loudness dips to the original programme material at AD cue points.

For example, Ofcom does not offer any technical recommendations or guidelines in setting the DVs, but rather guidance on best practice in what and when to describe[3]. Ofcom's predecessor, the Independent Television Commission (ITC), in the past offered the following guidance on standards for AD production[4]:

> "*When a descriptive commentary is inserted into a programme, the background level of programme audio needs to be reduced so that the description can be clearly heard.*"

The recommendation here is helpful to AD content creators but lacks any real benchmark or standard and does not present target values.

Amazon Prime Video offer the following information in their Global Content Guide[5] but do not offer specific dip targets for when creating imbedded AD content:

> "*In our effort to further support blind and visually impaired customers, Prime Video fully supports descriptive audio or "narration."*"

The BBC has in the past also offered guidelines on the approach to AD production[6], but nothing that could be considered a technical recommendation for setting programme audio DVs.

> "*As the programme sound level varies throughout a programme it is also desirable to be able to change the relative level of programme sound during description passages as heard by the AD user, reducing the programme sound more during loud passages (e.g. during an action sequence such as a car chase).*"

Again, these recommendations offer excellent guidance to professional mixing engineers involved in the production of AD. Experienced professional mix engineers can achieve the optimal balance between dynamic programme material and static audio descriptions instinctually. However, for those less well versed in the art of audio mixing, setting the optimal DVs can be challenging, if not a daunting task. Without the recommendation of a specific set of DVs for *Loud*, *Normal*, and *Quiet* programme types, expressed in decibels (dB), the subjective interpretation of guidelines leads to a wide margin of inconsistency in both the production of AD and the experience of the service user.

In its technical specifications document[7], Sky television states the following regarding the creation of AD content:

> "*AD is not a requirement but if available it should be sent to Sky and preferable as BWAV file, though other formats are accepted.*"

Netflix does offer the following technical guidelines around AD programme dip values[8].

"The description should be mixed to sound as though it was part of the original content."

"For a 5.1 Printmaster (PM), dip center channel only for descriptive events. For very loud sections or for films with very wide dynamic range, it's acceptable to dip the Left and Right channels of a 5.1 PM as well, generally no more than -6db, and sparingly up to -12db when absolutely necessary."

"For a 2.0 Printmaster, dip both channels accordingly. Dip original version mix 6-12 dB, per mixer discretion."

The recommended Netflix values are welcome, but still broad in their scope. The term "*mixer discretion"* is broad in both its application and its possible outcomes, and if DVs are not set by a professional mix engineer, there is a possible risk that the applied DV will be less than optimal for the given programme type, leading to inconsistencies in the audience experience.

The following details were provided by a broadcaster:

"*Our default settings are fade depth: 15db, fade duration: 3 frames (beginning and end). If the programme/scene calls for it, we can also adjust to the following:"*

Fade Depth (Dip Value)
-0.3 dB
-3 dB
-6 dB
-9 dB
-12 dB
-13 dB
-15 dB
-18 dB
-21 dB
-25 dB
-30 dB
-35 dB
-40 dB
-85 dB

Fade Duration	
Frames (25FPS)	ms
0	0
3	120
10	400
25	1000

It is encouraging to know that the practice outlined here facilitates the ability to adjust the DV to a setting appropriate to the loudness of the programme audio at the AD cue point. It is understood that the choice of DV setting is left to the best judgment of those creating the Receiver Mix metadata or Pre-Mix embedded WAV file, but it is unknown if guidance or defined values are given for specific programme loudness types such as *Loud*, *Normal* or *Quiet* passages. Again, without an industry defined DV paradigm, inconsistencies may be introduced to the quality and consistency of AD services if those making the adjustments mentioned above are not experienced mix engineers.

Other Research in the Field

The only known study investigating the preferred dip values applied in broadcast was published in the Journal of Audio Engineering Society in December, 2019[9]. In that paper an overview of the literature regarding the favored Loudness Difference (LD) between background audio and foreground speech in TV programme mixes is presented. Table 1 below summarises the favored LD values in Loudness Units (LU) for the literature reviewed in [9]. It is noted that all of the research reviewed there is for programme excerpts where the background audio is at a static level with no dips applied.

Source	Loudness Difference	Condition
BBC	-4 dB	In addition to standard mix levels
ARD/ZDF (Germany)	-7 LU	Commentary over music and ambience
	-16 to -23 LU	Commentary over Voice
The Digital Production Partnership (UK)	-4 LU	Minimum requirement
NHK	-9±3 LU	Documentary programmes only
The Adjustment/Satisfaction Test	-7 LU	Average value

Table 1: Overview of the literature presented in [9] *in relation to the preferred loudness difference between background audio and foreground dialogue/speech/commentary.*

[9] also presents the commonly used LDs observed during the dipping of background audio in commercial documentaries broadcast in the UK, Germany and France, and are presented in summary here in Table 2.

Excerpt Type	Observed Loudness Difference (Dip Value)
Voice over Voice (VoV)	-10 to -15 LU
Commentary over Music (CoM)	-10 to -15 LU
Dialogue over Music (DoM)	-14 to -17 LU

Table 2: Observed Loudness Difference or Dip Values by in 12 documentaries broadcast in the UK, Germany and France.

Key findings from the listening test carried out in [9] show that "*non-expert listeners (such as consumers) prefer LDs that are, on average, 4 LU higher than the levels preferred* *by experts (a category that would include mix creators).*", that non-expert listeners prefer LDs between 6.5 and 15.4 LU (equivalent to DVs between -6.5 and -15.4 dB) for Commentary over Music (CoM) and between 11 and 20 LU (equivalent to DVs between -11 and -20 dB) for Commentary over Ambience (CoA), and that "*participants found a range of ±3 LU around the preferred LD acceptable*". This research recommends the following Loudness Differences (LDs) for the following programme types:

Programme Type	Loudness Difference (equivalent to -dB Dip Values)
Dialogue over Music (DoM)	10-15 LU
Commentary over Ambience (CoA)	15-20 LU

Table 3: LD ranges recommended in [9] *for the programme types of DoM and CoA.*

It is noted that the values presented here are for programme types with "*background signals with characteristics that do not vary largely over time*". No integrated loudness values are given for any of the programme types and therefore there is no way of knowing precisely how the loudness of the background audio corresponds to the loudness differences applied during the dialogue and commentary passages.

Service-user Survey

In order to better understand the issues and inconsistencies experienced by those accessing AD content, a survey of service-users was carried out. A questionnaire was created to gather information in relation to AD services currently available in broadcast, on streaming platforms, and other audio-visual amenities. It was designed to help better understand the current standards in AD services, and if a standardised approach to AD production is needed. The questionnaire is available in Appendix 7.

24 people with an average age of 44 participated in the survey across the UK (16), USA (5), Canada (1), India (1), and New Zealand (1). 62.5% of those surveyed reported to have *Excellent* hearing, 12.5% *Very Good*, 20.8% *Good*, and 4.2% *Poor*. 100% of respondents identified as *Vision Impaired*.

Figure 1 below outlines the types of AD services accessed by respondents. It is evident that TV and Streaming services dominate the viewing habits of those participating in the survey.

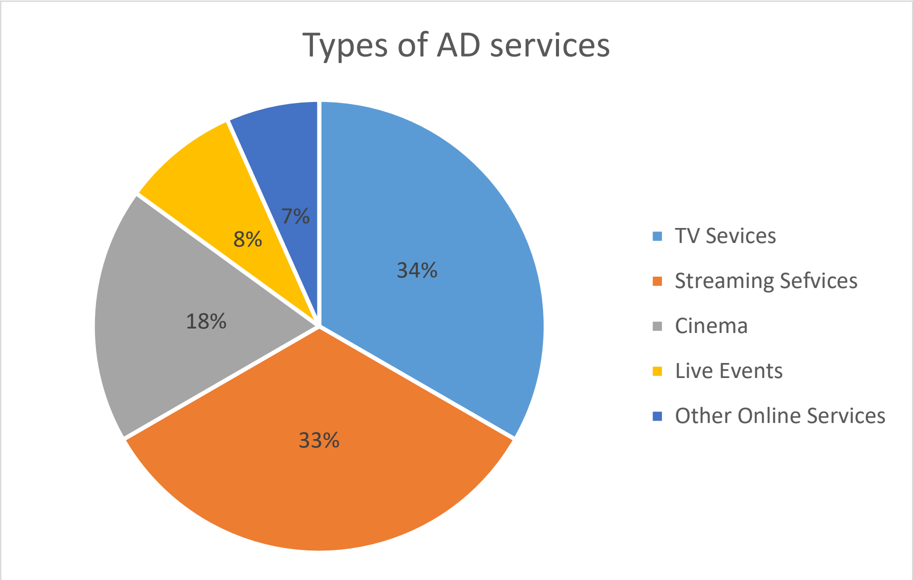


Figure 1: Distribution of the types of AD services accessed by respondents where N=24.

Figure 2 shows the prevalence of AD services accessed by the survey respondents. Participants were asked to select from a list of service providers and also to offer details of other AD service providers that they use. From these results it can be seen that Netflix is the most prevalent among responders, with the BBC, ITV, Channel 4, and Amazon Prime also showing strong popularity.

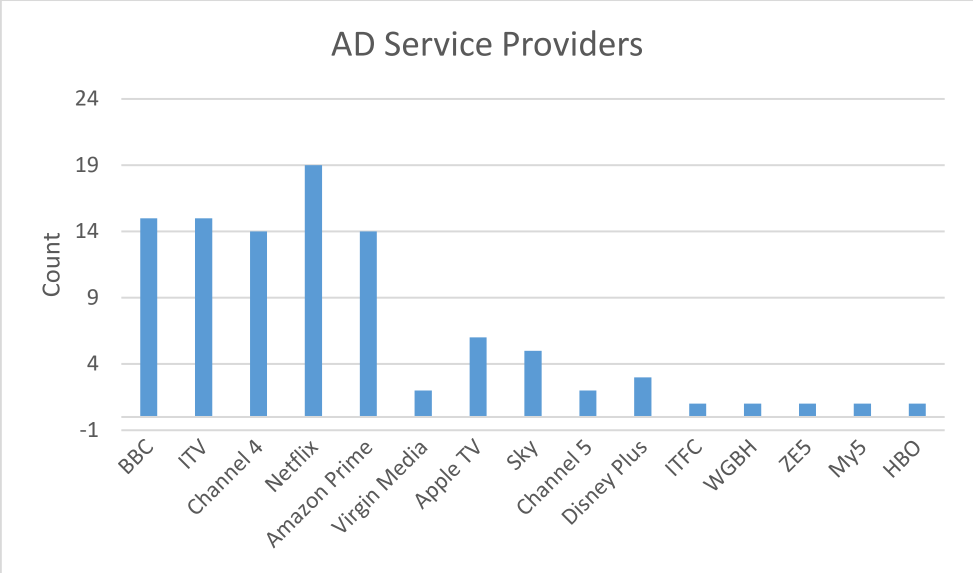


Figure 2: The prevalence of AD service providers among survey respondents where N=24.

When asked '*How would you rate the quality of AD services that you use?*' it is encouraging to see from Figure 3 that 62% of respondents rated the AD services that they use as either *Very Good* or *Excellent*.

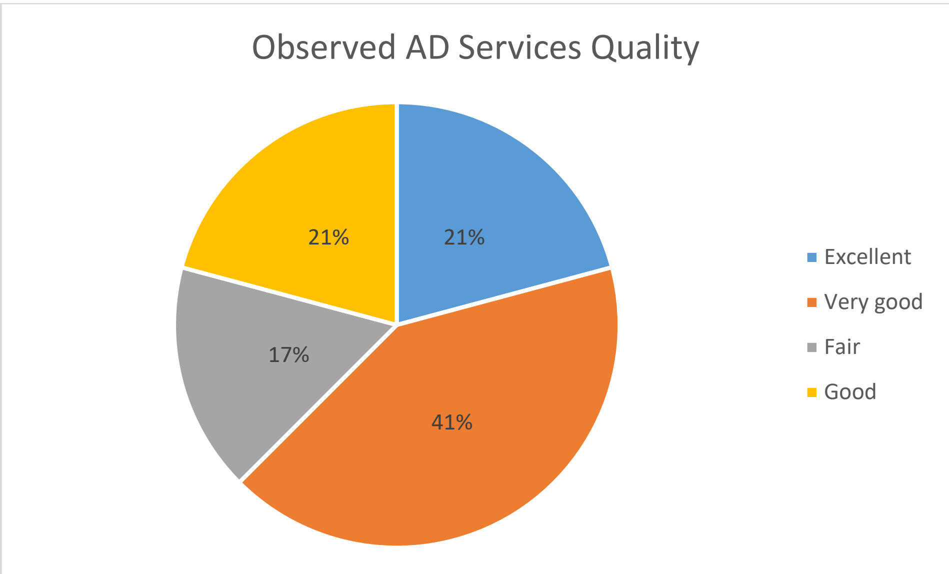


Figure 3: The distribution of results when participants were asked 'How happy are you with the AD services you use?'. The number of participants was N=24.

However, when asked '*How happy are you with the AD services you use?*', 29% of respondents said they were *Very Happy*, while 50% said they were *Somewhat Happy*. 21% of those surveyed indicated that they were *Somewhat Unhappy* with the AD services they use. There would indicate that the quality of AD services provided by major broadcast and streaming services falls short of the expected quality and standards. 79% of the 24 respondents indicated that they have experienced inconsistencies in the AD services that they use. It can be seen from Figure 4 that both the AD loudness relative to the programme dialogue and the programme loudness during the AD passages score high as production inconsistencies experienced by service-users. It is these relationships that are under investigation in this report.

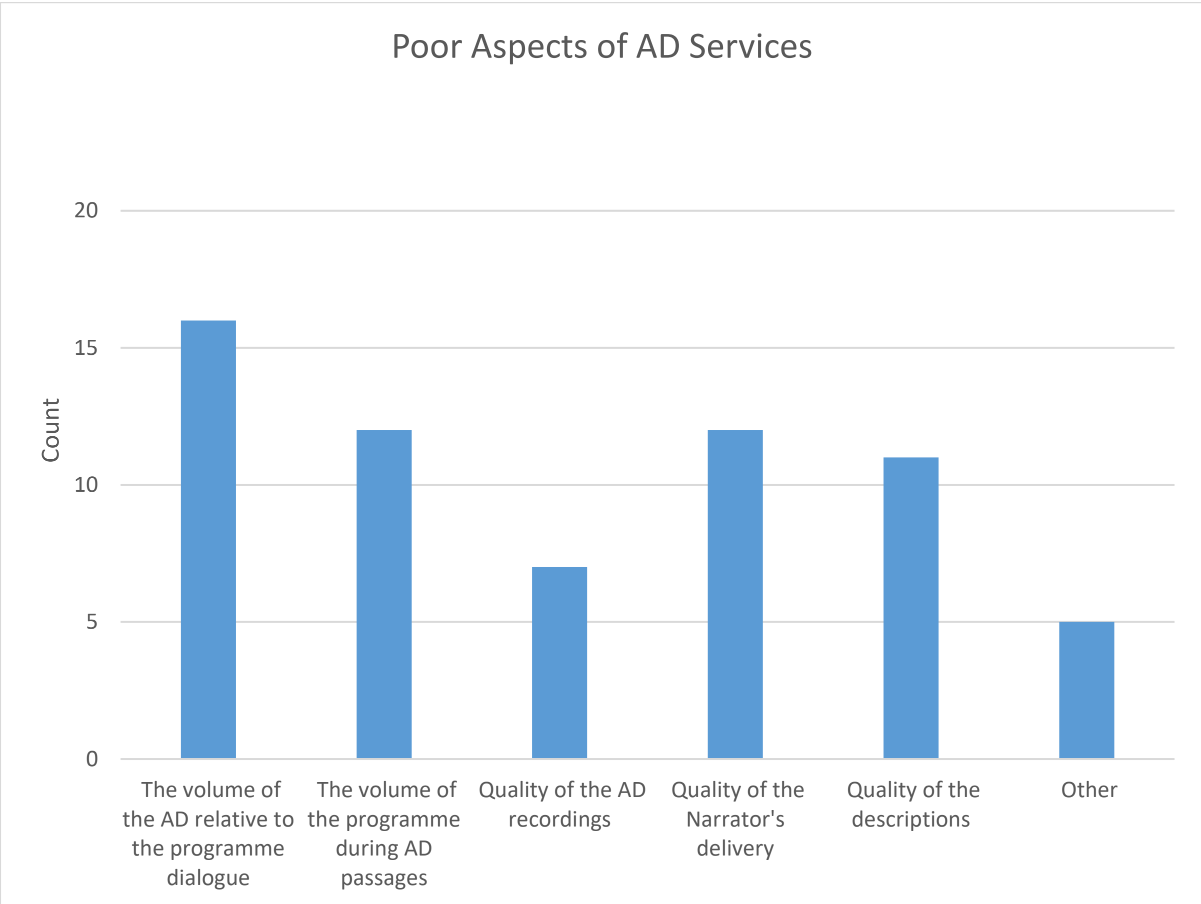


Figure 4: Distribution of responses to the suggested inconsistencies in AD services where N=24.

The qualitative data relevant to this research that was provided by respondents through the *Other* option include:

"1. AD content can sometimes be louder than actual program, meaning that it can boom through your house if using a subwoofer, for example.
2. Level of ducking on UK terrestrial TV is very different from say, Netflix, which I personally think, gets it right.
3. Relating to 2, the nasty step-step method used by UK broadcasters is incredibly jarring. It seems that it's not a linear fade, but a few DB step-down, then a few more, describer speaks, then the reverse. It's not smooth and it irritates me beyond belief."

"Variable volume levels."

"Individual descriptions differ [in] volume levels, alignment between dialogue can be variable"

"AD levels can vary from each provider - when loud music is played [and] dialogue, the AD can get lost in the mix and, the volume overall needs to be higher to hear it but, the music also gets louder too."

"…details of description varies, for example sometimes race is described and sometimes not, volume reduction of program not always consistent."

"They are just all inconsistent and often stop working"

“The most irritating issue I have with watching TV on any device is that there doesn’t seem to be any consistency in reliable provision of audio described programmes.

For live programmes:

- *Sometimes, a live programme is advertised as AD, yet the AD doesn’t play on any device.*
- *Worse, it sometimes doesn’t play on one device but does on another. Case in point was The Ypcress Files on ITV on Sunday. The AD did not work on the ITV player, but did on the Sky Q.*

On demand:

- *It is rare that a programme on catch-up will have AD.*
- *Sometimes, the latest catch-up episode of a programme will have AD, but previous episodes do not.*

Searching for AD content:

- *Trying to find an AD programme is hit and miss. For example, Sky has a voice control to find AD, but, depending on which command I use, I get two different listing. Also, when looking at the results, I have to go into the programme details before I can find out what channel it is on. This is irritating when I’m looking for a film to watch as I do not subscribe to Disney, yet some of the listings are for Disney films.*
- *On the BBC player, my wife has to check out the ‘Audio described’ listings to find AD content. However, sometimes, the latest programme in a series is not listed in that list so she has to go to some other listing to find it. Case in point is ‘Murder in Paradise’.”*

> *"Particularly in musicals or movies with music, there seems to be very mixed opinions on whether the volume of the music should be lowered for description. A bigger issue is that sometimes, the description happens before or after the event does."*

> *"sometimes when I am watching description on programs it is like the volume is cranked down the description is inserted and then the volume is cranked up after. it makes the film really choppy and un watchable. x 2 is a good example of this."*

> *"...American AD puts the program in mono, when targeting the group that actually listen, so will notice stereo the most"*

These comments provide strong evidence of a service that is not only inconsistent between providers but also within individual programmes. The observed inconsistencies span many of the production stages of AD content creation and it is the next stage of this report that will look to determine the details around these inconsistencies, where they lie within current industry practice, and if they can be ameliorated through a standardised approach to AD content creation along with a more robust Quality Control (QC) stage.

Of those surveyed, 42% of respondents said that they have complained to a provider regarding the quality of their AD service.

Objectives and Scope

This project attempts to obviate any inconsistencies observed in current AD production practices by benchmarking the current AD mixing practices of professional mix engineers, assessing the range of current practices on an audience, analysing and evaluating the assessments, and by developing a standardised approach to setting AD Dip Values. The focus of the research is to develop a set of DV recommendations for *Loud*, *Normal*, and *Quiet* programme types in order to optimise the listening experience of the AD service users.

For the purpose of this study the following questions will be addressed:

- Does an audience have a preferred set of AD Dip Values for *Loud*, *Normal*, and *Quiet* programme types?
- Will the perceived quality of the AD user experience correlate to a prescribed range of AD Dip Values?

As part of this study, investigations will include one research hypothesis:

- A standardised approach to setting Audio Description Dip Values for *Quiet*, *Normal* and *Loud* TV programme types will help optimise both the end user experience, and the AD service production process.

Methodology

The research design is centred around the following investigations:

- Assess and benchmark the current practices in AD content creation within commercial broadcasting and streaming services. This is done firstly by means of an industry survey to better understand the key roles and production practice of those creating content for AD services, and secondly through an observational study of the mixing practice of professional sound engineers when setting AD Dip Values.
- Utilizing a control content (the programme material), develop a range of Audio Descriptions for evaluation by a target audience.
- Assess the range of current practices on an audience using listening tests to elicit quantitative and qualitative feedback across a range of AD users and a broad variety of content types.
- Analyse and evaluate the assessment to determine if a set of DV preferences exists among listeners correlating to *Loud*, *Normal*, and *Quiet* programme types.
- Determine if a set of recommendations and guidelines for setting DVs can be established to optimise the listening comfort of the target audience.

Stage 1: Current Practice

Observations in broadcast AD

During the establishment and development of the AD department in RTE in 2018, we found that there was no industry consensus on AD Dip Values and in particular, DVs appropriate to programme material of varied loudness. Evidence of this can be found in AD services on multiple broadcast and streaming platforms, with DVs observed to exhibit a wide and inconsistent range of values. Table 4 below shows the default Dip Values in dB as observed in various AD content.

Observed Default DV
-6 to -12 dB
-13 dB
-15 dB
-18 dB

Table 4: A selection of default Dip Values as observed in a variety of AD content.

These figures suggest that inconsistent standards exist in AD production practices for broadcast and that a standardised set of recommendation and guidelines in setting Audio Description Dip Values would bring about a more consistent end user experience. To better understand the AD production practices in industry and where potential sources of such disparity might lie, an industry survey was created.

Industry Survey

In October of 2021, a questionnaire was circulated among the international AD production community to establish current production trends and duties, document the technical knowledge of those producing AD, and explore whether the automation of this process would be welcomed. The key focus of this survey was to better understand the roles, technical skills, production practices and workflow of those involved in the AD production process, and whether these factors contribute to inconsistencies in AD content across a variety of broadcast and streaming services. The survey was focused mainly on standards in setting AD Dip Values (DVs) and was aimed at those practitioners responsible for setting this parameter.

Participants: To establish a subject group reflective of the type of roles involved in the production of AD content, the industry survey was distributed among AD scripters, narrators, sound engineers, media producers and directors in Ireland, the UK, Europe, the USA, Canada, Argentina, and Australia. AD is still at an early stage of development, with a limited number of production specialists in each region. Recruitment in such a specialised area of audio production can prove difficult but, in total there were 42 respondents.

Participants were sourced via Audio Description associations and a number of key individuals and service providers were also recruited (see Table 5). Respondents have experience in the production of AD content for both broadcast and streaming services, TV and film, advertising, theatre and cinema, and across multiple genres.

Region	Association	Website
UK & Ireland	The Audio Description Association (ADA)	http://audiodescription.co.uk/
USA	Audio Description Associates LLC	https://audiodescribe.com/
USA	The Audio Description Project	https://adp.acb.org/index.html
Canada	Accessible Media Inc.	https://www.ami.ca/new_ami
Argentina	Textual Perceptions	https://www.percepcionestextuales.com.ar/
Australia	Vision Australia	https://www.visionaustralia.org/

Table 5: Audio Description associations and service providers used to source survey participants.

Survey Findings

The overall response to the survey was very positive with participants keen to contribute to the research. Many are heartened to discover that research is underway in this field and have also expressed interest in receiving the research findings and final report.

To better understand the various production workflows used by AD content creators, it is important to establish the roles and technical knowledge of those involved in the process, the standards used, and whether a common approach exists when setting DVs for programme material of varying loudness.

Defining the Sample Groups.

When asked to identify within the AD production titles of *Scripter*, *Narrator*, *Producer*, *Director*, and *Sound Engineer*, many participants identified with multiple roles. Only 14% of respondents classified themselves solely as *Sound Engineer*, a further 6% identified as *Sound Engineer* plus a number of the other AD roles presented, while 14% identified with all titles. In total 40% of respondents had various levels of sound engineering expertise. It is important to note that 60% of the AD content creators surveyed did not classified themselves in any way as sound engineers.

It is expected that participants identifying, in any way, as a *Sound Engineer* will have more expertise in the field of audio production than those respondents not identifying as such. Because of this, the main sample group has been partitioned into the three sample subgroups of *Sound Engineer*, *Sound Engineer with Other AD Roles*, and *Non-Sound Engineer*, as shown in Table 6 below.

Sample Group	Producer	Director	Scripter	Narrator	Sound Engineer	Other
Sound Engineer					✓	
Sound Engineer with Other AD Roles	✓	✓	✓	✓	✓	
Non-Sound Engineer	✓	✓	✓	✓		✓

Table 6: Participant sample groups and their constituents.

Figure 5 shows the number of respondents in each group as a percentage of the total number of respondents.

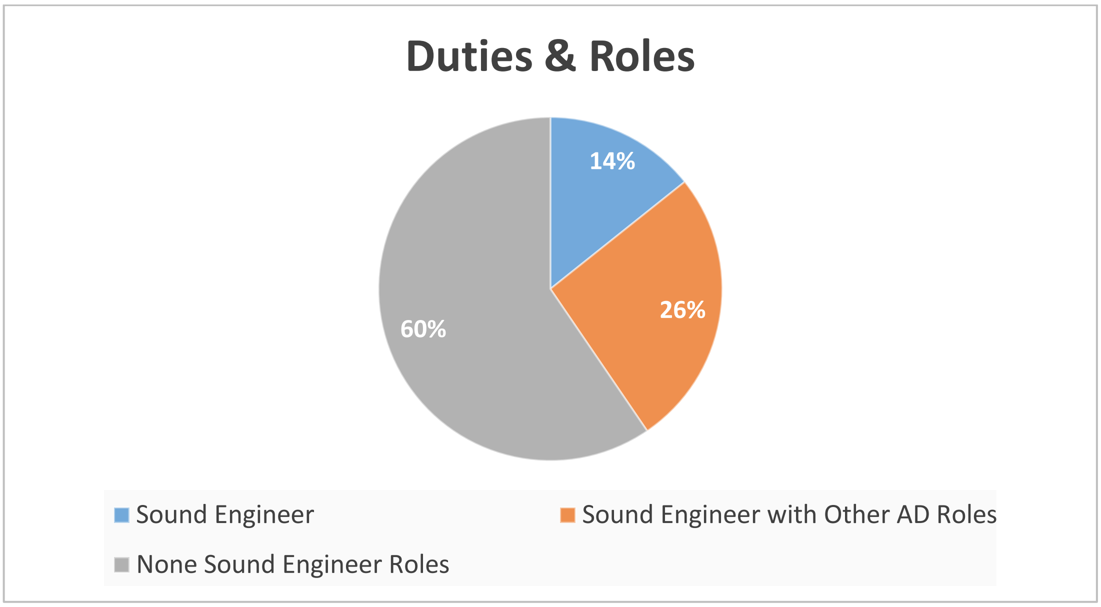
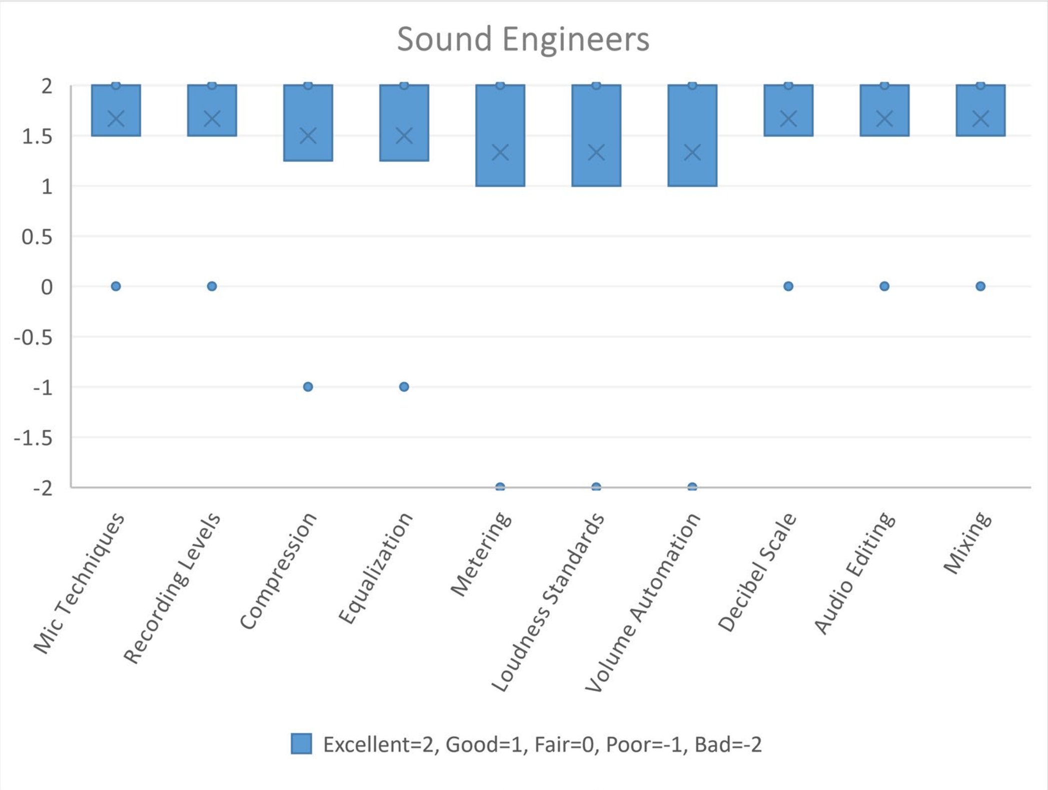


Figure 5: The distribution, where the number of respondents N=42, of 'Sound Engineer' (14%), 'Sound Engineer with Other AD Roles' (26%), and 'Non-Sound Engineer' (60%) roles among participants when asked "What role(s) best describe your AD production responsibilities: Scripter, Narrator, Producer, Director, Sound Engineer?"

Technical Knowledge.

When asked '*How much involvement do you have in the creation of AD content?*', 45% of all respondents chose *Lots* and 32% chose *All aspects*. This would suggest that respondents have considerable insight into the AD production process, are experts in their field, and are well placed to provide informed and reliable data.

When asked to rate their technical knowledge on audio production aspects such as microphone techniques, recording levels, compression, equalization, metering, loudness standards, volume automation, the Decibel scale, audio editing, and mixing, a high level of expertise was found to be as expected within the *Sound Engineer* group, with a knowledge mean between *Good* and *Excellent* (see Figure 6 below).



s

Figure 6: Technical knowledge in audio production among N=6 respondents classified as 'Sound Engineer'. Participants were asked to rate their knowledge in the above audio production categories.

Figure 7 shows the distribution of audio production knowledge among respondents classified as *Sound Engineers with Other AD Roles*. Here we can see that the knowledge mean is still high in the categories of Microphone Techniques (*Good*), Recording Levels (*Good / Excellent*), Audio Editing (*Good / Excellent*), and Mixing (*Good*), but drops in the categories having a stronger association with setting DVs, such as Compression (*Fair / Good*), Loudness Standards (*Fair*), Decibel Scale (*Fair*), and Volume Automation (*Fair*, with a lower quartile at *Poor* and values outside the quartile reaching *Bad*). Volume Automation is the audio production technology employed to dip the programme volume at the AD cue point. Along with a *Good* to *Excellent* understanding of Loudness Standards, the Decibel Scale and Mixing, a *Good* to *Excellent* understanding of Volume Automation is the preferred expertise when manually setting AD Dip Values.

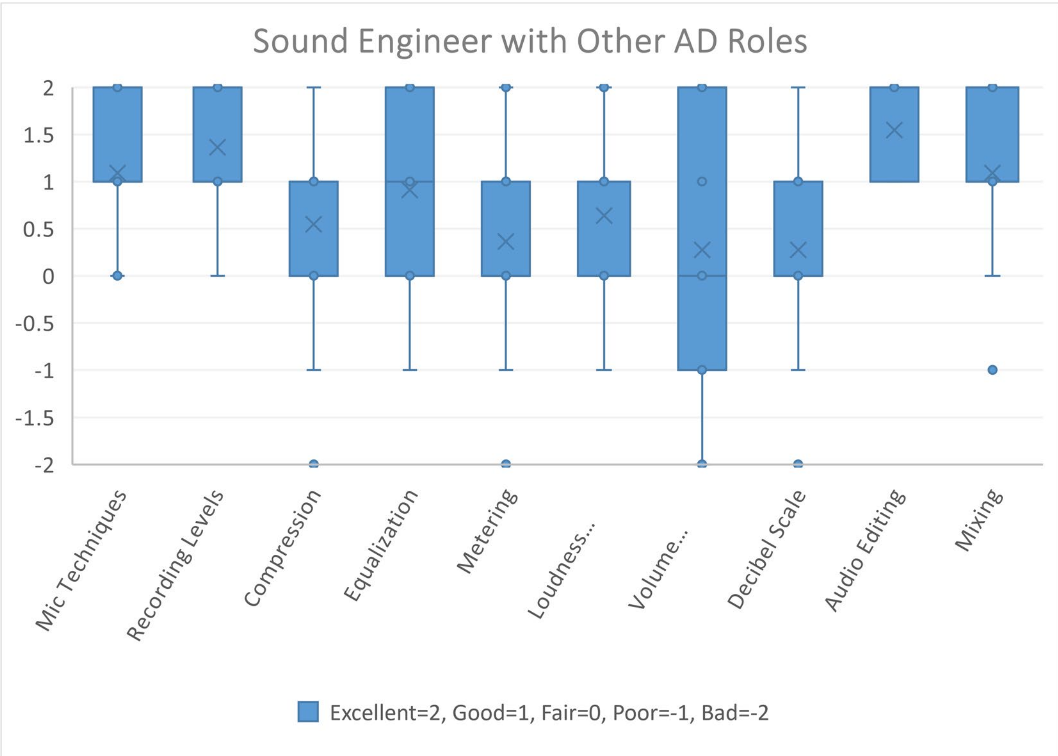


Figure 7: Technical knowledge in audio production among N=11 participants classified as 'Sound Engineer with Other AD Roles'. Participants were asked to rate their knowledge in the above audio production categories.

A further drop-off in knowledge in the categories of Compression, Metering, Loudness Standards, Volume Automation, the Decibel Scale, and Mixing is evident in Figure 8 among respondents classified as *Non-Sound Engineer*, with the mean for these audio production categories lying between *Fair* and *Poor*, and with values extending in many of the categories to *Bad*. Again, this is the largest subject group involved in the production of AD content.

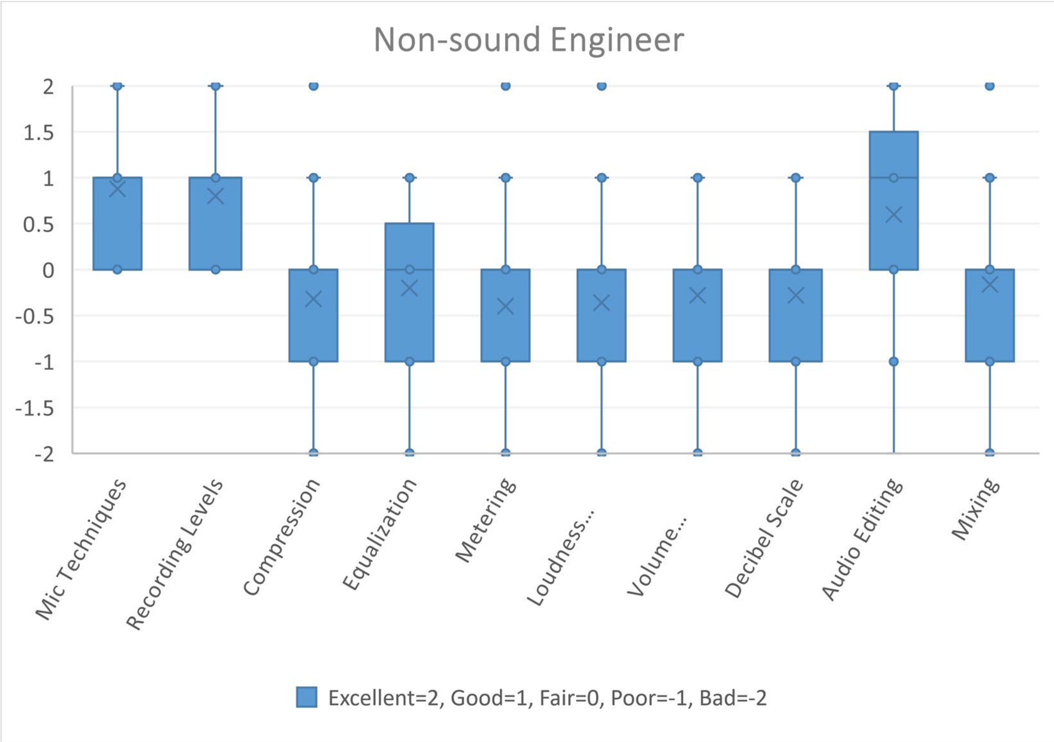


Figure 8: Technical knowledge in audio production among N=25 participants classified as 'Non-Sound Engineer'. Participants were asked to rate their knowledge in the above audio production categories.

Production Practices

The workflow for setting Dip and Fade Values is software dependent, but the practical application is similar in all Digital Audio Workstations (DAW). In order for the narration to be clear, the loudness of the programme material at the AD cue point is reduced, or dipped, by a user defined amount in Decibels (dB). The Dip and Fade Values, or amount and rate by which the programme material volume is reduced, is defined by the measured loudness of the programme material at the AD cue point. Programme material with a wide Loudness Range (LRA) requires Dip and Fade Values to vary in response to the programme loudness. Programme material with a narrow loudness range requires less variance in the Dip and Fade Values.

"*Loudness Range (abbreviated 'LRA') quantifies the variation in a time-varying loudness measurement. Loudness Range is supplementary to the main audio measure, Programme Loudness, of EBU R 128*[10]. *Loudness Range measures the variation of loudness on a macroscopic time-scale, in units of LU (Loudness Units). The computation of Loudness Range is based on a measurement of loudness level as specified in ITU-R BS.1770*[11]*, albeit with a different relative gating threshold. Loudness Range should not be confused with other measures like dynamic range or crest factor, etc.*"[12].

Dip and Fade Values

When asked "*Do you set the Dip and Fade values during AD production?*", Table 7 below outlines how this important responsibility is distributed among the three sample groups. Here it can be seen that 67% of the *Sound Engineer* respondents said they *Always* set the Dip and Fade Values during AD production. 24% of *Non-Sound Engineer* respondents said they *Always* set the Dip and Fade Values and 16% said they *Sometimes* do. This shows that 40% of AD production personnel not identifying as a sound engineer and having no formal training in the art of mixing still undertake the responsibility of setting AD Dip and Fade values. The *Other* category contains responses that are more ambiguous, such as:

"*I don't usually handle that side of things.*" – *Non-Sound Engineer*

"*no experience yet with dip and fade values*" – *Non-Sound Engineer*

"*I set my mix manually, so unsure what you mean.*" – *Sound Engineer with Other AD Roles*

"*On subtitled sections*" – *Sound Engineer with Other AD Roles*

"*Haven't had the opportunity yet.*" – *Sound Engineer*

Sample Group	Always	Sometimes	Never	Other
Sound Engineer	67%	0%	17%	17%
Sound Engineer with Other AD Roles	27%	36%	18%	18%
Non-Sound Engineer	24%	16%	40%	20%

Table 7: Responses to the question "Do you set the Dip and Fade values during AD production?".

When asked "*What determines the AD Dip and Fade value?*" participants chose from the following options: *Default Settings, Recommendations and Guidelines, The programme loudness at the Audio Description point, Your own judgment,* and *Other*. Only 7% of all participants said that *Recommendations and Guidelines* determine their Dip and Fade Values, while 21% said their *Own Judgement*, and 31% said *The programme loudness at the Audio Description point*. It is noted that 17% of respondents use *Default* settings for Dip and Fade Values. The 24% of respondents who chose *Other* do not set Dip and Fade Values. Table 8 shows how these determining factors are distributed between the *Sound Engineer*, *Sound Engineer with Other AD Roles*, and *Non-Sound Engineer* sample groups.

Subject Group	Default Settings	Recommendations and Guidelines	The programme loudness at the AD cue point	Your own judgment	Other
Sound Engineer	0%	0%	50%	33%	17%
Sound Engineer with Other AD Roles	9%	0%	45%	45%	0%
Non-Sound Engineer	24%	12%	20%	8%	36%

Table 8: Response to the question "What determines the AD Dip and Fade value?"

Further analysis shows that 40% of *Non-Sound Engineers* responded by choosing an option aside from *Default Settings* and *Other*. This result is consistent with the data presented in Table 7 and confirms again the significant proportion of *Non-Sound Engineer* participants responsible for setting AD Dip and Fade Values during AD production. Surprisingly, no participants from the *Sound Engineer* group are using *Recommendations and Guidelines*. It is encouraging to see that 83% of *Sound Engineer* and 90% of *Sound Engineer with Other AD Roles* participants are using an intuitive approach to setting Dip and Fade Values based on their *Own judgment* and the *Programme loudness at the AD cue point*.

Defined Dip Values

Figure 9 shows the ranges of DVs used by the participants that provided defined values. There is a wide range of DVs used within the *Non-Sound Engineer* sample group, while all *Sound Engineers* quote a DV setting of -12dB.

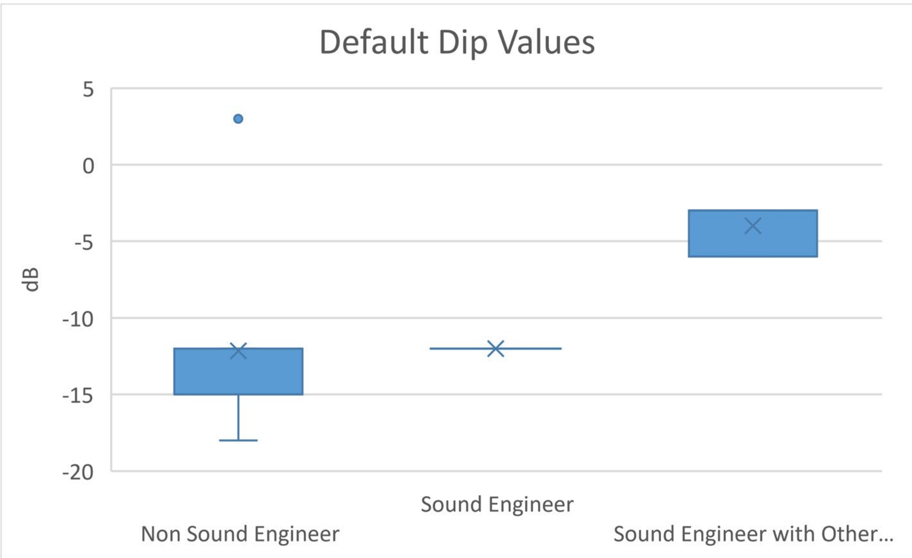


Figure 9: The distribution ranges for Dip Value settings presented by sample groups when asked: "What is your default Dip Value in dB?".

It is important to note that not all respondents provided a single value. Many presented a range as outlined in the following statements:

> *"A range based on loudness. I get videos from all over and there's not a solid one size fits all. What I get varies in levels." – Non-Sound Engineer*

> *"4-8 dB for quiet - 8-12 for louder sections" – Sound Engineer*

> *"12-18 during very loud music, but closer to 3-6 if it's less distracting environmental sound. 0 if the environmental sounds are quiet like room tone, traffic, or wind, for example." – Sound Engineer with Other AD Roles.*

> *"Between -2dB and -10dB depending upon the loudness of the source audio." – Sound Engineer with Other AD Roles.*

There is clear evidence here that *Sound Engineer* and *Sound Engineer with Other AD Roles* respondents are conscious of the need in their production practice to adjust DVs in accordance with the loudness of the programme material at the AD cue point.

Using Recommendations and Guidelines

When asked "*What recommended values or guidelines, if any, do you use for setting Dip Values?*" *Non-Sound Engineer* respondents provided answers such as:

> *"Our own guidelines for dip values, EBU guidelines for Loudness"*

"In-house guidelines recommending -13.0 as standard, increasing to -18.0 for louder moments e.g. music / loud background noise"

"-15dB was specified by one company I worked with so I set it to this level, then this level was approved by other companies."

"I've not been involved in deciding on the default. I do listen back and if the dip seems too little I will increase it (e.g. for foreign language subtitle dubbing) but also aim to ensure some of the original sound is still audible so as to not be too abrupt with the beginning / end of the dip."

"Guidelines from training with a bit of own judgement depending on the sequence"

"Recommend[ed] levels suggested in Stellar training"

Although some of the DVs presented here are approaching the mean *Sound Engineer's* DV of -12dB from Figure 9, there is also a lack of cohesion in the approach to setting DVs within the *Non-Sound Engineer* subject group.

Sound Engineer respondents provided answers such as:

"Mostly, I follow what I perceive as best, but often it is in line with specifications that Netflix and Amazon have sent. In general, not exceeding a 12 dB dip in program. I use a multi-band compressor on program, so it is not full-frequency. It only affects frequencies that overlap AD and program."

"It varies depending on the loudness of the scene"

"Have the AD be as subtle as possible against the material - but ensure the AD is always audible even in loud scenes"

"Allowing enough original audio to still be present to convey tone of sound track."

From these responses it can be seen that *Sound Engineer* respondents present a much more cohesive approach to the setting of DVs, paying close attention to both the requirements of the programme material at each AD cue point and defined recommendations and guidelines commensurate with the mean *Sound Engineer's* DV of -12dB from Figure 9.

Sound Engineer with Other AD Roles respondents provided answers such as:

"If I use them for subtitled sections, I do it so the English is clear over the subtitled language."

"I try to not dip the original audio if possible, to allow the audience to listen as it was intended. However, if it makes the AD unbearably loud or misunderstood, then I dip."

"I try to keep the AD levels steady between -10 and -7dB depending upon the peak points and overall volume of source film audio."

"Clarity of speech when I listen to it but also keeping the final piece at a steady volume throughout, whatever that volume is in the film when it's provided to me."

Again, inconsistencies in the approach to setting DVs are evident within the *Sound Engineer with Other AD Roles* subject group, with some respondents operating in line with the mean *Sound Engineer*'s DV of -12dB, while others tend not to adjust the programme material but unexpectedly boost the AD cue.

Adjustment of DVs over the Programme Duration

50% of all participants indicated that they *Never* or *Rarely* change the DVs over the programme duration while 50% said they either *Always*, *Often*, or *Sometimes* do.

With further analysis, Figure 10 shows that this statistic is reflected in the *Sound Engineer* and *Non-Sound Engineer* subject groups. The *Sound Engineer with Other AD Roles* group has a mean response between *Sometimes* and *Often*.

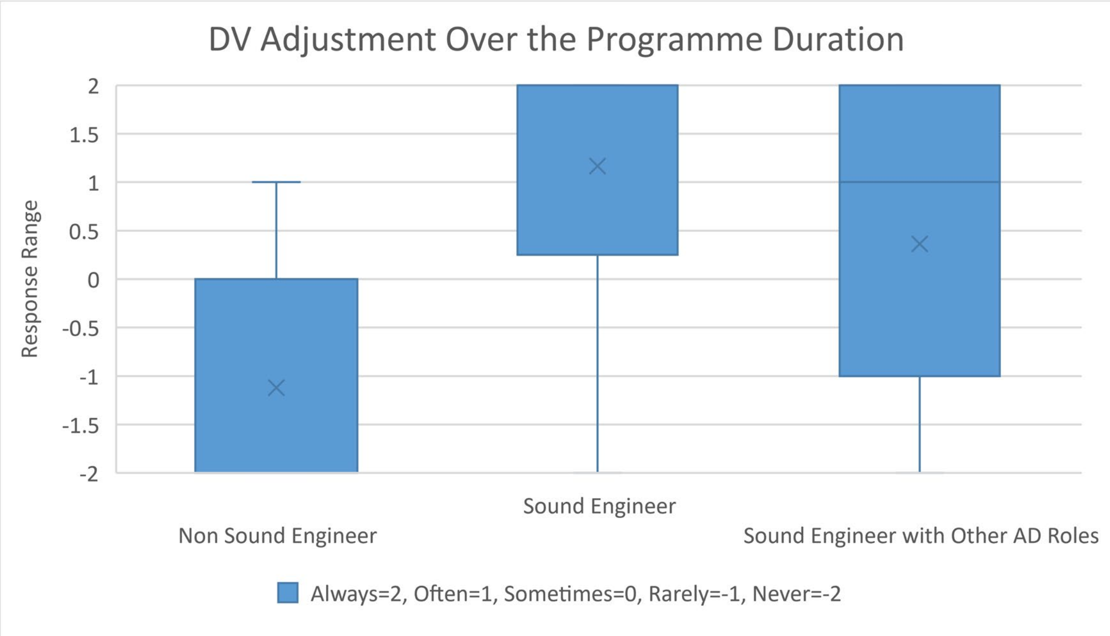


Figure 10: Responses to the question "Do you adjust the DV Value over the duration of the programme?"

This would suggest that the majority of *Non-Sound Engineer* respondents do not adjust the DVs in accordance with the programme loudness at the AD cue point.

Production Standards

Because of the wide range of DVs observed in AD content, it was felt necessary to determine; if Quality Control (QC) stages exist in AD production workflows; have AD content producers received feedback from end users or broadcasters; and have any complaints from end users been received in relation to the perceived quality of AD services.

Subject Group	Yes	No	I don't Know
Sound Engineer	83%	0%	17%
Sound Engineer with Other AD Roles	100%	0%	0%
Non-Sound Engineer	84%	8%	8%

Table 9:Responses to the question "Is there a Quality Control stage to your AD production workflow?"

Sample Group	Yes	No	I don't Know
Sound Engineer	67%	33%	0%
Sound Engineer with Other AD Roles	82%	18%	0%
Non-Sound Engineer	80%	12%	8%

Table 10: Responses to the question "Do you ever receive feedback from AD service end users or broadcasters?"

Sample Group	Yes	No	I don't Know
Sound Engineer	17%	83%	0%
Sound Engineer with Other AD Roles	55%	45%	0%
Non-Sound Engineer	28%	64%	4%

Table 11: Responses to the question "Have you ever received complaints from the end user or broadcaster about AD service quality?"

It is encouraging to see from Table 9 that a high percentage of those surveyed include a QC stage in their AD production workflow, while Table 10 reveals a strong dialogue between the service content creators and the end users or broadcasters. Complaints received by the content creators in relation to AD service quality are high, with 55% recorded among the *Sound Engineer with Other AD Roles* group, 28% of *Non-Sound Engineer* respondents, and 17% of the *Sound Engineer* sample group. *Other* responses include qualitative data such as:

> **Quality Control** - *"Scripters QC their own scripts and voicers QC their own recordings. Our voicers record to an onscreen levels meter and we process all individual recorded descriptions through a batch processer/leveler to process to uniform loudness EBU R128." – Non-Sound Engineer*
>
> *"most streaming and broadcast programs are listened to by producer and engineer before final files are made. Theatrical releases often are listened to by a 3rd party before being approved as well." – Sound Engineer*

> **Feedback** – *"Yes, we have a permanent user group for soliciting such feedback and we are also regularly in touch with our Playout provider (we are the broadcaster)." – Non-Sound Engineer*
>
> *"not directly, but certain narrators are more in-touch with AD audiences and I hear feedback occasionally that way." – Sound Engineer*

> **Complaints** – *"Occasionally there were some complaints about the AD being too soft or too loud. This was usually due to a tech glitch in the describer software. The software monitoring is much more robust now.;" – Non-Sound Engineer*

"Have received complaints on sound quality" – Non-Sound Engineer

"Paramount is very particular and requests that we do not record via remote means. Their programs are only recorded in-studio. This ensures consistency. Most broadcaster feedback is not about quality, it is sync related or other technical detail related that doesn't affect the quality of narration or mix." – Sound Engineer

"*A couple times the loudness values were wrong but soon learned." – Sound Engineer with Other AD Roles*

"*on dips that I have not had control over, i.e. a production company other than myself, I have had complaints that the AD was too low to hear, or that feedback was present*" - *Sound Engineer with Other AD Roles*

Standardisation of Dip and Fade Values

Recommendations and Guidelines: Because of the DV discrepancies observed in commercial AD services and the lack of standardised recommendations and guidelines in industry, participants were asked if a set of recommendations on setting Dip and Fade values would be useful to their workflow. Figure 11 outlines a strong response among the *Non-Sound Engineer* and *Sound Engineer with Other AD Roles* groups in favor of a set of technical recommendations and guidelines for the setting of AD Dip and Fade Values. Participants identifying themselves as purely *Sound Engineer* were less favorable towards the suggestion of a standard set of recommendations and guidelines, with 50% saying they would be *Somewhat useful*, 33% saying *Not useful*, while 17% of the group had *No opinion*. 19% of all participants had *No opinion*.

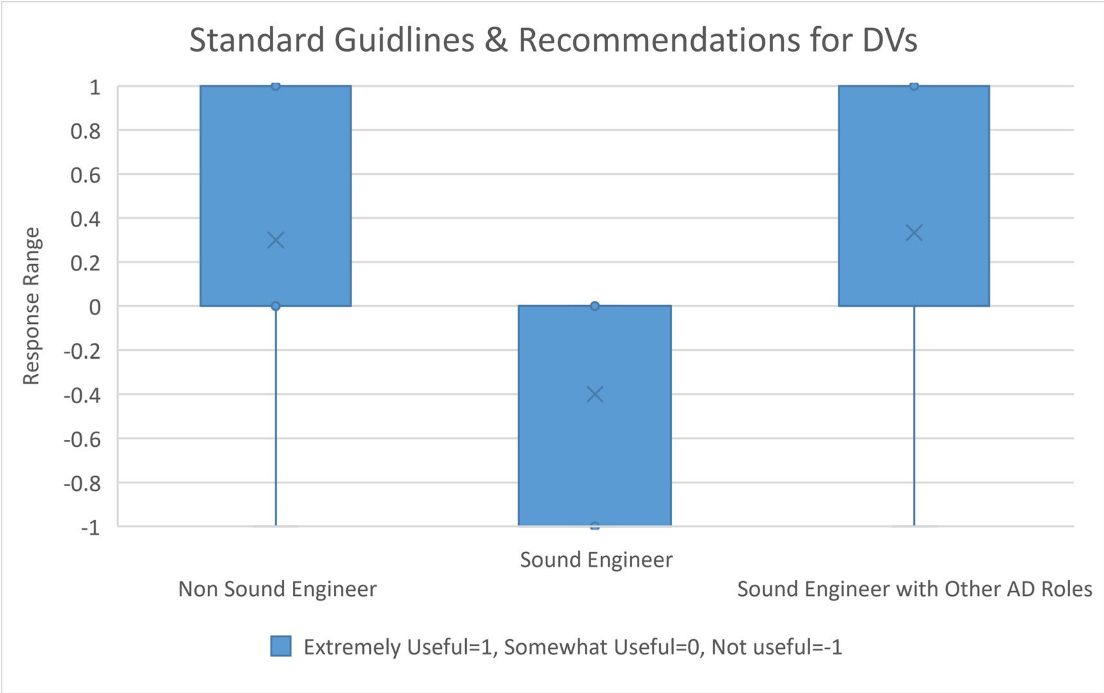


Figure 11: Responses to question "How useful would a set of recommendations on setting Dip and Fade values be to you?".

Dip Value Automation: One solution to the standardisation of DVs is through the use of automation. The setting of DVs can be controlled by software, with parameter values derived from the integrated average loudness of the programme material at the AD cue point.

It can be seen from Figure 12 that an automated setting of DVs is most welcome among *Non-Sound Engineer* participants, while the strongest opposition to the proposal understandably comes from the *Sound Engineer* group. Analysis of Figure 13 shows the response mean close to the *Maybe* option for the *Non-Sound Engineer* and *Sound Engineer with Other AD Roles* groups.

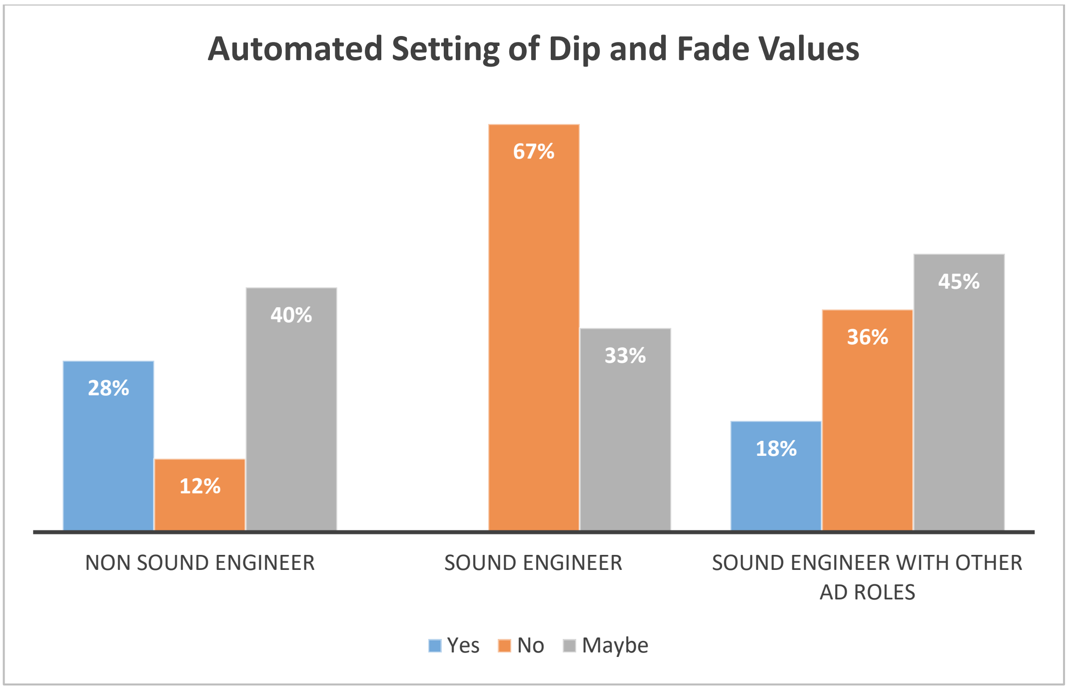


Figure 12: "Would you like the setting of Dip and Fade values to be an automated process?"

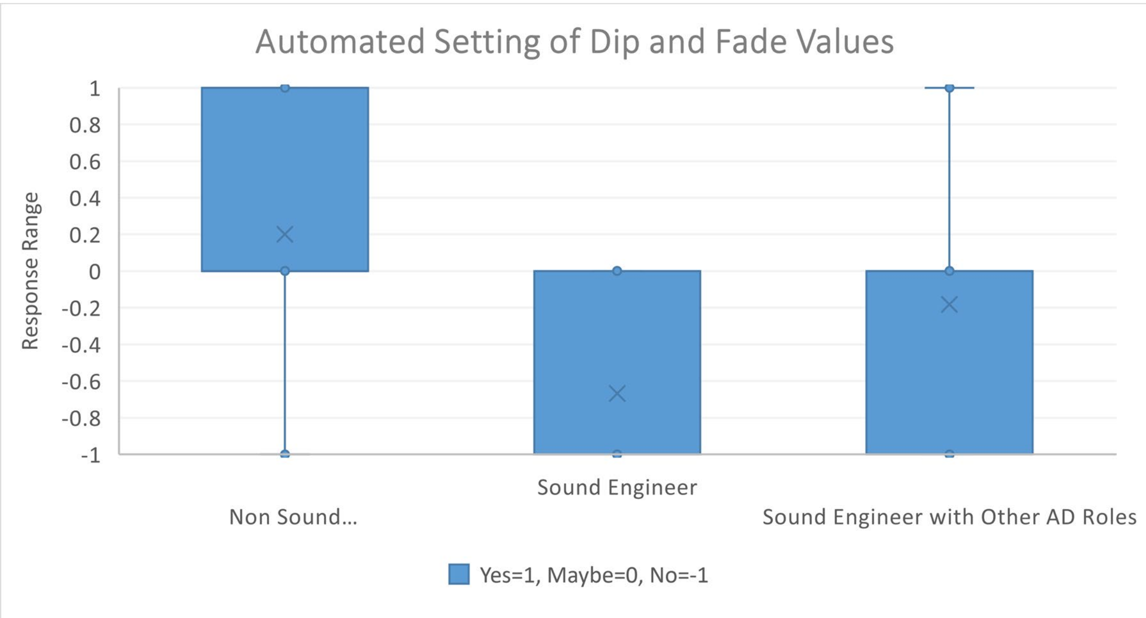


Figure 13: Responses to the question "Would you like the setting of Dip and Fade values to be an automated process?".

Creators as Users

As in Figure 14, it is clear that only a small percentage of AD content creators are users of AD services. This would suggest a low level of self-assessment and a disconnect may exist between those involved in the creation of AD content and the end user experience.

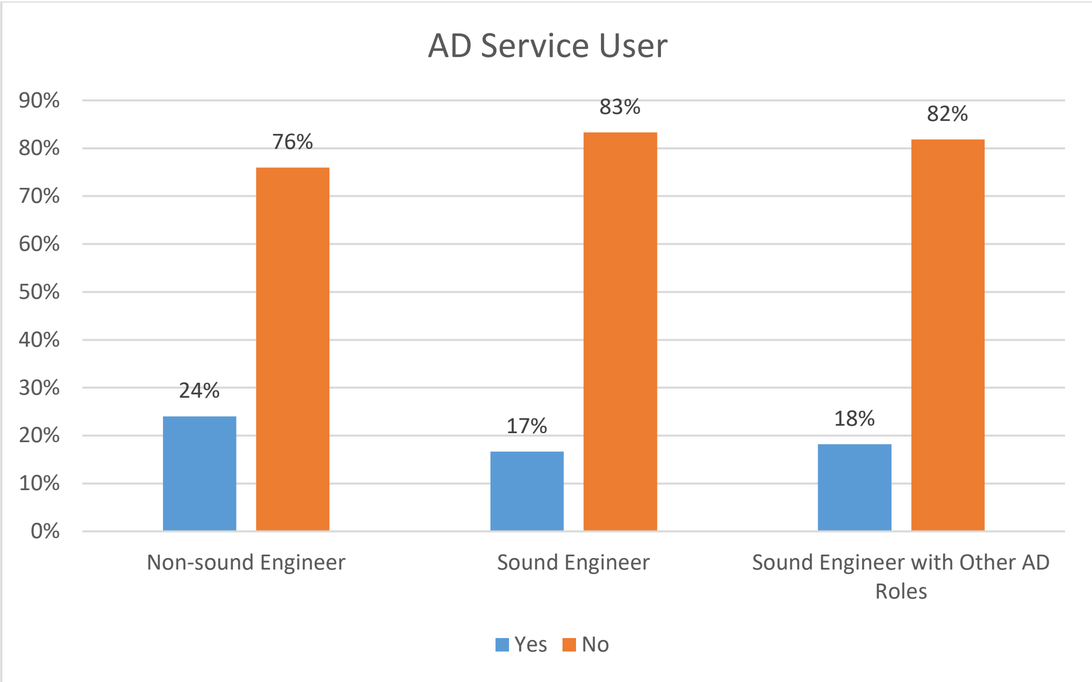


Figure 14: Responses to the question "Are you an AD service user?".

Conclusion

The scripting and narration of Audio Description cues is a very creative and expressive part of the AD production process, and one that should benefit and enhance both the programme narrative and the audience experience. However, the recording and mixing of the AD cues against the programme material involves a considerable amount of technical knowledge, experience and skill found only in the expertise of trained sound engineers.

The data from the subject groups surveyed here indicates that the majority of professional AD content creators are not sound engineers, have little or no training in the art, and only *Fair* to *Poor* technical knowledge in the essential areas of audio production required for the successful delivery of high and consistent standards to AD service users.

The practice of setting AD Dip and Fade Values requires skill in the art of mixing, particularly if the programme material loudness varies from cue to cue. The data gathered in this survey shows that AD Dip and Fade Values are being set by operators who are not sound engineers. There is also evidence to suggest that inconsistencies exist in the quantifiable Dip Values applied to the programme material at the AD cue points, and that DVs are often not adjusted in response to the programme loudness. The most consistent approach comes from *Sound Engineer* respondents, where a standard DV of -12 dB is applied but, outside of this group, Dip Values range from +3 to -18 dB. It is encouraging to note that 52% of respondents in some way vary the DVs according to their *Own Judgement* or "*The programme loudness at the Audio Description point*". However, 36% of those surveyed state that they never change the DV over the duration of the programme, and 17% are using a static *Default* setting.

It is not surprising that, given the lack of international standards, only 7% of all participants are using technical recommendations or guidelines during their production practice. There is evidence of a general consensus among AD content creators for the need to bring clarity to the descriptive passages, while also retaining the integrity of the programme narrative. This is most evident in the workflow presented by *Sound Engineer* respondents, in which a standard -12dB DV is applied at AD cues, and that this value should be increased (larger negative value) for louder programme sections and reduced (smaller negative value) for those that are more quiet. There is however no consensus on the amount by which DVs should be adjusted for loud and quiet programme material. Moreover, a mensurable definition of *Loud*, *Normal*, and *Quiet* programme types is missing from the available data. This report will later introduce the concept of loudness bands, with clearly defined upper and lower loudness values, used to quantify programme material as either *Loud*, *Normal*, or *Quiet.*

There is a strong appetite among *Non-Sound Engineer* and *Sound Engineer with Other AD Roles* respondents for a standardised set of technical recommendations and guidelines to setting AD Dip and Fade Values. Both groups have also responded in strong favor of automating the process using software.

From this we can conclude that inconsistencies will exist in AD services if there is a lack of standardised workflow and approach to setting DVs for programmes material of varying loudness. It is the mixing practice of professional sound engineers that is the benchmark when setting DVs and it this practice that will be under investigation in the next section of this report.

Stage 2: Benchmarking

Loudness Measurements

The terms used when measuring loudness are: LKFS, LUFS, LU and LRA [13].

LKFS: Loudness K-weighted referenced to digital Full Scale is the unit term used in the ITU-R BS.1770 [11] standard for measuring loudness. One unit of LKFS is equal to one decibel (dB).

LUFS: Loudness Units referenced to digital Full Scale is the term used by the European Broadcasting Union (EBU) when measuring loudness [10]. One unit of LUFS is equal to one dB.

LUFS and LKFS are identical and both are absolute measurements.

LU: Loudness Unit is the relative measure of loudness. Again, one unit of LU is equal to one dB.

LRA: Loudness Range is defined as a quantifier that "measures the variation of loudness on a macroscopic time-scale, in units of LU (Loudness Units)." [5]

A programme's integrated average loudness is measured in LUFS using the standardised measure of loudness level introduced in ITU-R BS.1770 [11] and implemented in the technical recommendation EBU R128 [10]. The EBU integrated loudness target for broadcast is -23 LUFS ±0.5 LU. Along with the integrated average loudness value of the AD narration track, it is of the opinion that Audio Description DVs are dependent on the programme's LRA and more specifically, on the measured integrated LUFS value of the programme material across an individual AD cue point. This stage of the research aims to present:

- The correlation between the integrated LUFS value of an AD narration track (as set by a professional mix engineer) and the programme material's LRA value.
- The correlation between the integrated LUFS value of the programme audio at an AD cue point and the applied DV.

Benchmarking Experiment

Method: As part of the benchmarking of current practices in commercial broadcasting and streaming services, investigations were made on the practice of professional mixing engineers when creating AD content. These investigations set out to determine:

1. The statistical mean integrated loudness of audio descriptions mixed by professional sound engineers against R128 compliant programme material.
2. The statistical mean Dip Values applied by professional sound engineers when mixing audio descriptions against R128 compliant programme.
3. If there is a correlation between the above and a programme's LRA value.

The experiment was an observational multiple stimuli test in which each stimulus (item) consisted of a TV programme and an accompanying AD narration. The experiment was designed to elicit the participants' response to the items and it was considered critical that participants carried out the experiment using a familiar mixing environment and professional technical setup. After consultation with the participants, it was understood that a standardised, but unfamiliar, setup and environment would introduce uncertainty to their mixing practice. Therefore, participants carried out their mixing duties in their place of work. The experiment was designed to benchmark the mixing practices of experts in the field of audio mixing and to determine if a statistical trend exists.

Items: Six test items were presented. Each consisted of stereo programme material (WAV file with a sample rate of 48 KHz and 24 Bit resolution), coupled with a separate mono, unprocessed, undipped, centre panned AD narration track (WAV file with a sample rate of 48 KHz and 24 Bit resolution). All programme material was extracted from its transmission MXF[2] as supplied by RTE. Each item was presented as an excerpt created by concatenating segments selected from a specific programme title.

Programme Types: Programmes were chosen to represent a wide range of genre types, a good blend between dialogue and music focused productions, and LRAs typical of both TV and feature film productions. Segments of each programme were chosen to illustrate a broad variety of integrated loudness levels and to yield as wide a range of DVs as possible. For example, one segment chosen from 'The Bourne Identity'[14] involves an action sequence that begins with dialogue, general sound effects and background atmospheres (medium integrated loudness), progresses into a fight scene involving loud dialogue, very loud sound effects and loud music (high integrated loudness), and ends with a transition into a new scene with quiet sound effects and background atmospheres (low integrated loudness). Other examples include segments from the film 'The Greatest showman'[15] and the TV series 'Reeling In The Years'[16] that involve continuous music scenes and montages. Other segments were taken from the RTE soap opera 'Fair City'[17] giving good examples of dialogue-heavy scenes incorporating general sound effects without music. A segment from one episode of the children's TV series 'What Makes My Day'[18] was chosen as an example of a TV production for young people. This title incorporates a comprehensive blend of dialogue, production sound effects (PFX), Voice Over (VO) and music. It is important to

[2] The Material Exchange Format (MXF) is a professional container format for video and audio files as developed by the Society of Motion Picture and Television Engineers (SMPTE).

understand that, in order to establish a standardised loudness reference point, all programme titles were loudness normalised to be R128 compliant (-23 dB LUFS ±0.5 LU).

AD Narration: English language Audio Description cues were recorded for each programme by professional AD scripters and narrators at the Audio Description Department in RTE. A total of 317 AD cues were scripted and recorded across the six programme types. The AD narration track for each title was recorded at an arbitrary level determined by the narrator but following the recording guidelines and targets as set out in the AD authoring software [19] and by the Audio Descriptions Department at RTE. Narrations were delivered by an equal balance of female and male narrators using identical AD recording workstations (see Appendix 2 for details).

Table 12 below outlines the selected items and their more salient attributes.

Title	Prog. Integrated Loudness (LUFS)	Prog. LRA (LU)	Genre	Total Excerpt Duration	No. of AD Cues	Narration
The Bourne Identity (**TBI**)	-23	22.4	Feature Film	00:22:31:05	133	Male
The Greatest Showman (**TGS**)	-23	24	Feature Film	00:23:05:07	91	Male
What Makes My Day (**WMMD**)	-22.9	9.4	Children's' TV	00:07:59:18	24	Female
Reeling In The Years (1979) (**RITY**)	-23	6	Factual TV	00:09:14:11	34	Female
Fair City (**FC Male**)	-22.9	10.7	TV Drama	00:12:11:07	22	Male
Fair City (**FC Female**)	-23.1	11.8	TV Drama	00:07:11:12	13	Female

Table 12 Benchmarking control content.

Participants: As in [9], 8 experts (between the ages of 38 and 49) in the field of audio mixing were recruited from the Irish sound engineering community to benchmark mixing practices when setting Audio Description DVs. Participants had high testing ability, they were all professional sound engineers with more than 5 years' experience mixing content for broadcast and digital media. All 8 participants verified not to have any known hearing impairments. All participants were reimbursed for their time in taking part in the experiment.

Instructions: Each participant was given a set of instructions (see Appendix 1) and asked to set the static playback level for the AD narration against the unadjusted TV programme. Participants were asked to then apply DVs to the TV programme audio as they felt appropriate to the programme's perceived loudness at each AD cue point, while still maintaining the integrity of the programme's narrative. Participants repeated the experiment for each item. During the experiment, participants had full control over the listening levels, i.e. the overall playback volume, and were encouraged to carry out their mixing in the same manner and timeframe as they would when undertaking commercial productions.

Locations: Participants carried out their mixing duties using professional studios that they were most familiar with. Each studio consisted of high standard stereo playback equipment including high-end digital to analogue (DA) converters, studio monitors, and acoustically treated low-reverberant rooms.

Benchmarking Findings

AD Narration Track

The Sound Engineers were asked to establish a single, static playback level for the AD narration track in relation to its programme material. This level was to be established at a quiet section of the programme such that the AD cues are clear, in keeping with the average dialogue levels within the programme, and at AD cue points where the programme loudness does not necessitate dipping. This approach established the AD narration loudness level corresponding to a 0 dB DV reference. The participant's adjusted fader level, measured in dB, for the AD narration track were converted to Clip Gain[3] and the fader was then returned to 0 dB. This allowed the Sound Engineer's adjustment to be rendered into the fundamental playback level of the AD narration track and therefore the adjusted integrated loudness level to be measured in LUFS. The process was repeated for each item.

Figure 15 shows the distribution of the integrated LUFS values of the AD narration tracks delivered by the narrators for each of the items along with their mean. When compared to Figure 16, it is evident that the Sound Engineers have made only minor adjustments to the playback level (see Table 13) of the supplied AD narration tracks. This would suggest that the narrators (RTE's professional AD narrators) have excellent judgment when recording and setting the initial record/playback level of the AD cues for programme material with a variety of LRA values.

[3] Clip Gain is an AVID Pro Tools (software) feature defined as "Clip-based gain applied premixer (pre-fader and before any plug-in processing)."[24]

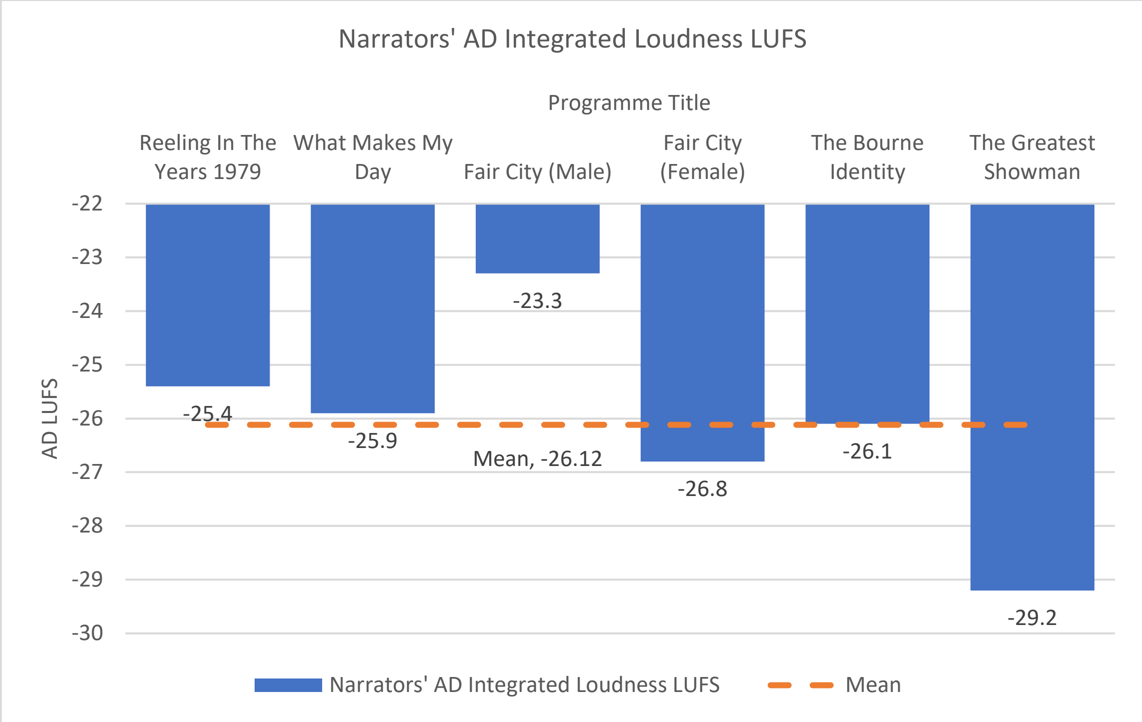


Figure 15: The integrated loudness of each programme title's AD narration track in LUFS, as recorded by the narrators. The mean integrated loudness for all six AD narration tracks is also shown.

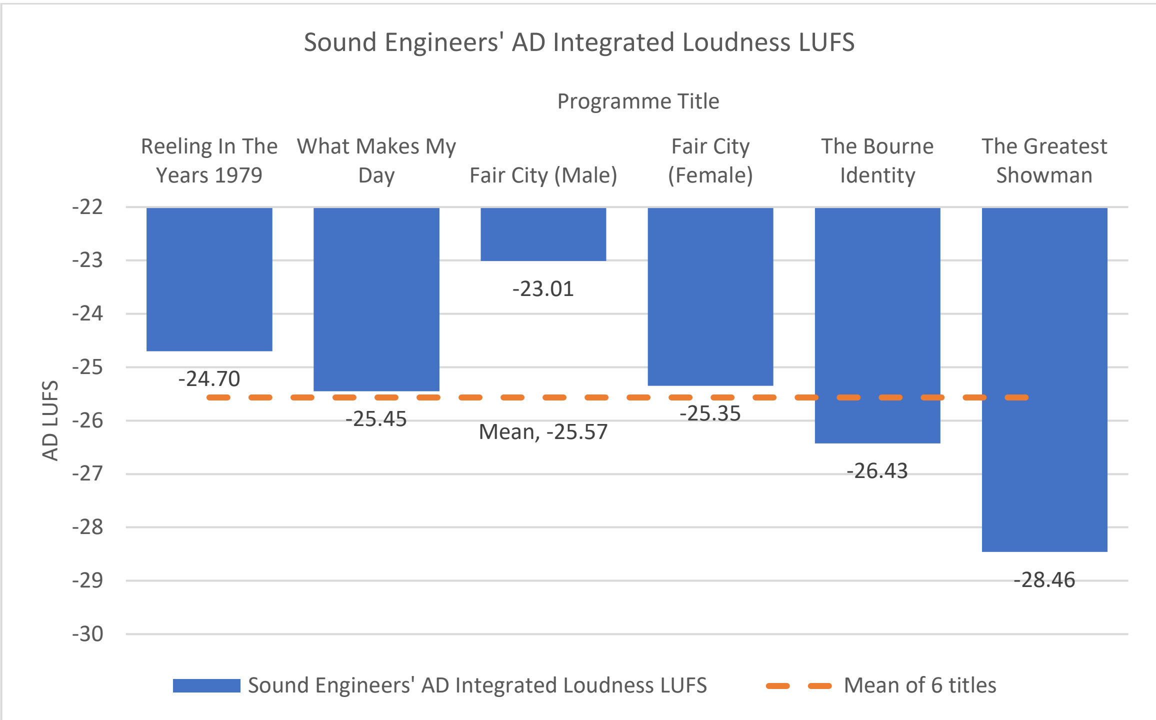


Figure 16:The mean integrated loudness of each programme title's AD narration track in LUFS, as adjusted and set by the Sound Engineers. The mean integrated loudness for all six AD narration tracks is also shown.

Programme Title	Narrator AD LUFS	Mean Engineers AD LUFS	Mean Engineers Adjustment (dB)
Reeling In The Years 1979 (RITY)	-25.4	-24.70	0.68
What Makes My Day (WMMD)	-25.9	-25.45	0.45
Fair City FC (Male)	-23.3	-23.01	0.56
Fair City FC (Female)	-26.8	-25.35	1.33
The Bourne Identity (TBI)	-26.1	-26.43	-0.33
The Greatest Showman (TGS)	-29.2	-28.46	0.69

Table 13: Comparison between the integrated loudness (LUFS) of the AD narration tracks as delivered by the Narrators, and the mean Sound Engineer values. The adjustment amount in dB is also shown.

It is clear from the data that a single mean integrated loudness value for the AD narration track will not suit all programme material, as a loudness difference of 5.45 LU exists between the loudest AD narration track (Fair City (Male), -23.01 LUFS) and the quietest narration track (The Greatest Showman, -28.46 LUFS). It is evident that the integrated loudness of the AD narration track, and the playback level it is set to, is dependent on the LRA of the programme material. Figure 17 below outlines the correlation between programme LRA and the mean integrated loudness value of the AD narration track set by the Sound Engineers.

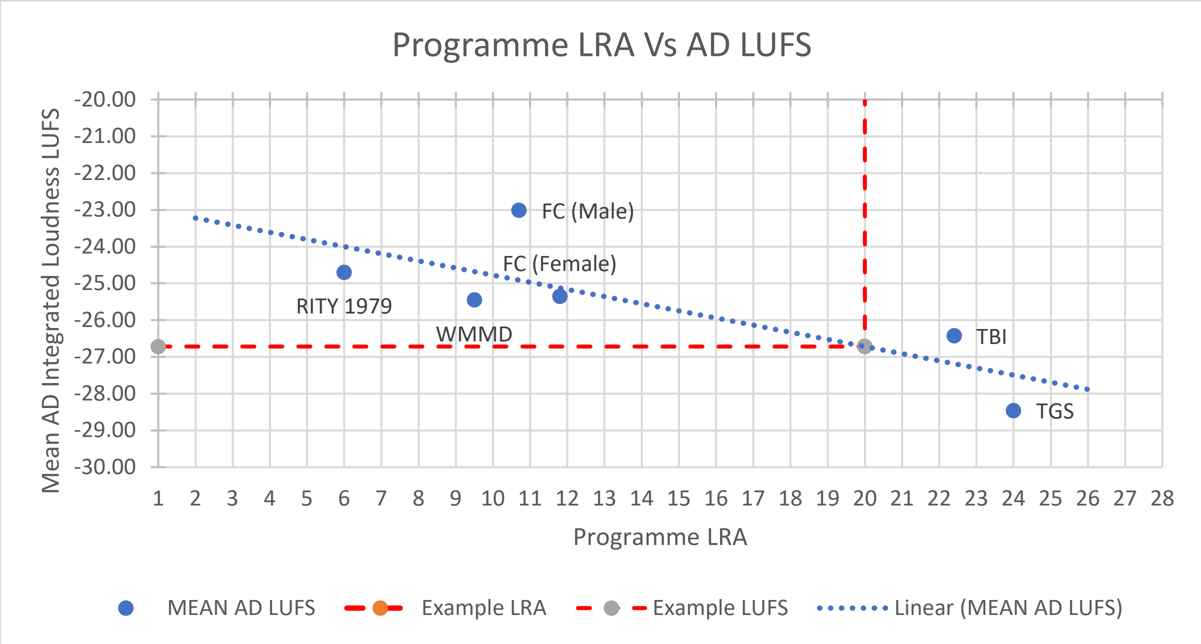


Figure 17: The relationship between programme Loudness Range (LRA) measured in LU (Loudness Units) and the mean integrated loudness of the AD narration track as set by the Sound Engineers and measured in LUFS.

The linear trend line in Figure 17 enables us to predict a programme LRA-dependent integrated loudness level for the AD narration track. Using such an approach, the AD narration track can be adjusted post recording to have a particular LUFS value that is suitable to the programme audio's LRA. Table 17 in Appendix 4 presents a look-up table based on this trend line. For example, as indicated by the red and green dashed lines in Figure 17, if programme material has a measured LRA value of 21 LU then the AD narration track should be set with a 0 dB undipped playback level and integrated loudness value of -26.9 LUFS. Future work should investigate further this relationship through an expanded set of programme types and LRA values.

AD Dip Values

After setting the playback level of the AD narration track and establishing a 0 dB DV reference, sound engineers were asked to apply DVs to the programme audio as they felt appropriate for the perceived programme loudness at the AD cue points. DVs were to be applied as uniform automation moves across each AD cue point and to include a fade in and fade out as deemed appropriate by the sound engineer. See Figure 18 for an example of one sound engineer's work on 'The Bourne Identity'. This procedure was repeated by each participant for the six items.

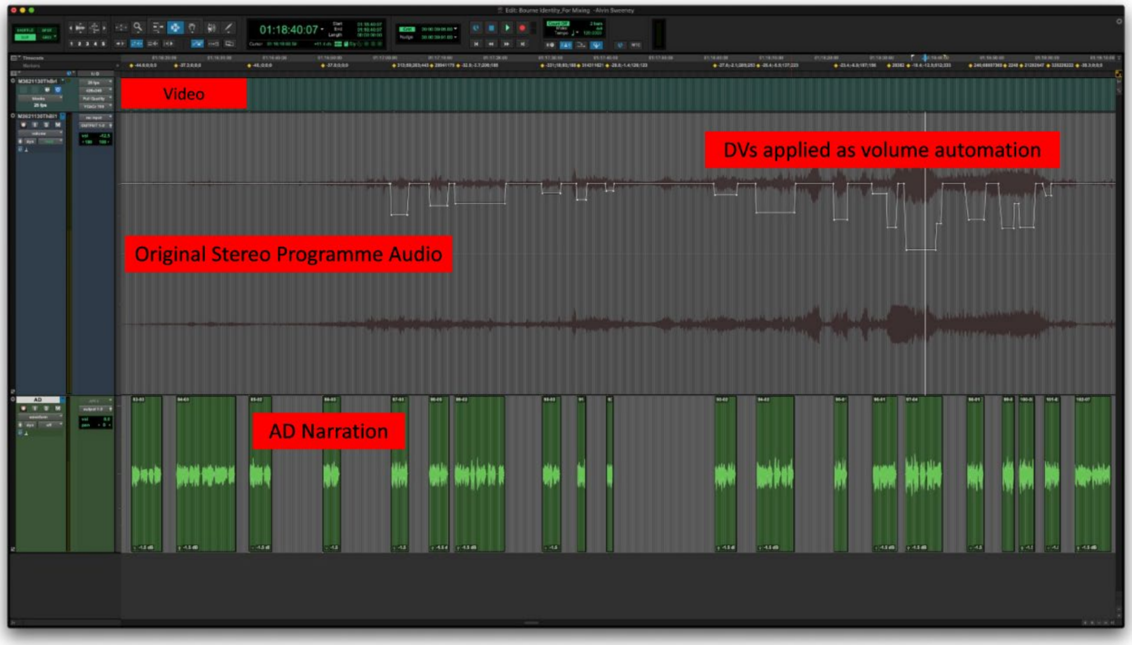


Figure 18: A Pro Tools[4] screenshot showing a participant's application of uniform automation DVs across AD cue points for the programme title 'The Bourne Identity'.

[4] Pro Tools is a professional Digital Audio Workstation (DAW) and the industry standard DAW in film and TV post-production sound.

Figure 19: A close-up on a section of Figure 18 showing the fade-in and fade-out, applied DV (-12.5dB at cursor) at the AD cues, and universal Clip Gain (-1.5dB) applied to the entire AD narration track.

Programme Loudness and DV Measurements

It is important to outline that it is the measured programme loudness at an AD cue point that determines the amount by which the programme material must be dipped. The programme loudness is measured in LUFS (integrated). The Dip Value (DV), or the amount by which the programme loudness is reduced, is measured in dB and should be such that the AD narration remains clear and unmasked by the programme material, while also preserving the integrity of the programme's narrative. It is the expertness of a group of professional sound engineers in setting DVs that is being benchmarked here.

The programme loudness was measured across the duration of each AD cue point for all 6 items, yielding a range of programme loudness values for each programme title. Figure 20 shows the range of programme loudness values measured at the AD cue points across each of the 6 programme titles. Figure 21 outlines the corresponding ranges of DVs applied by participants at the same AD cue points for each title.

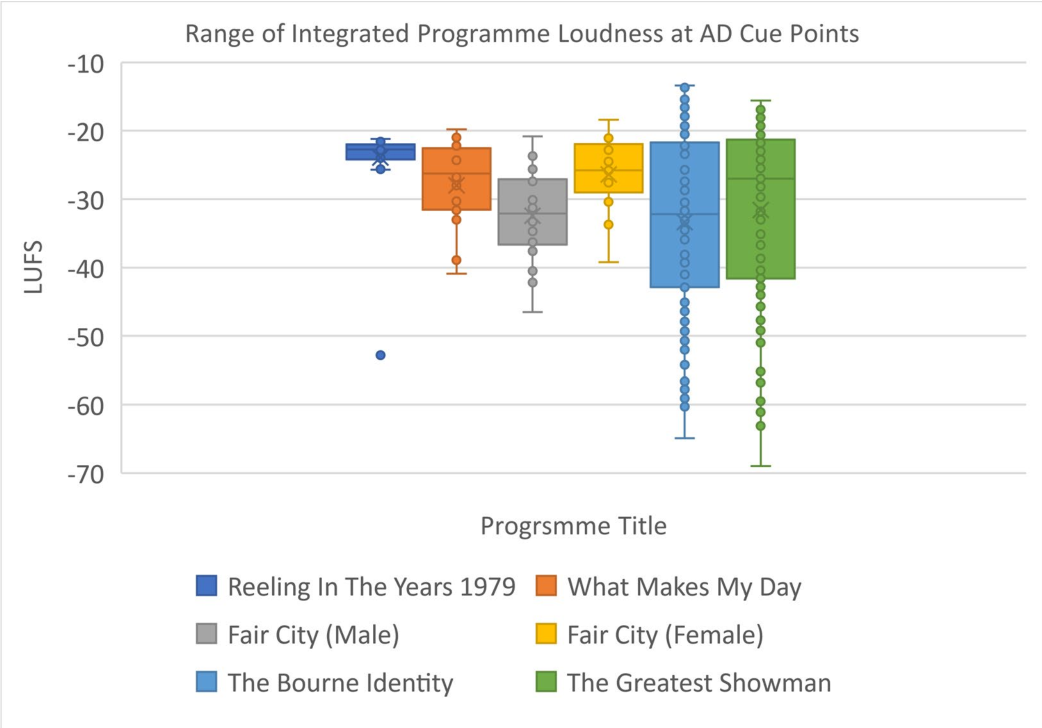


Figure 20: The range of integrated programme loudness (LUFS) measured at the AD cue points in each of the programme titles.

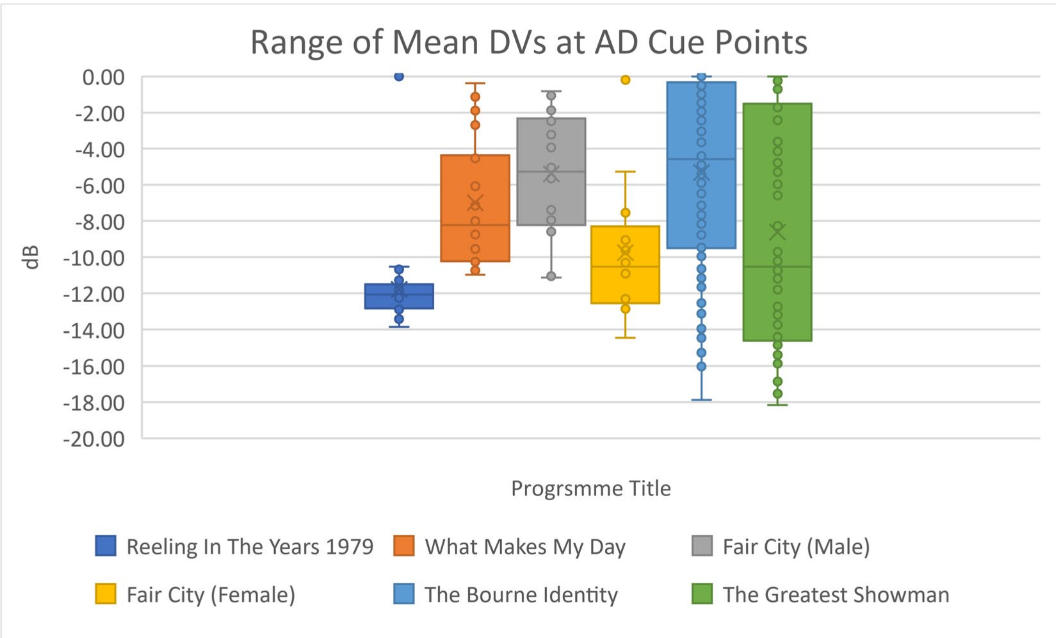


Figure 21: The range of DVs applied by participants at the AD cue points for each programme title.

From these figures we can see that each title has its own unique loudness features but also that there is a strong correlation between the range of applied DVs and the range of integrated LUFS vales measured across the AD cue points for each title. The data presented here reflects best practice in AD production as presented by the *Sound Engineer* subject group participating in the industry survey carried out in Stage 1: Current Practice. And contrasts strongly with the static approach to setting DVs outlined by many of the 'Non-Sound Engineer' subject group participating in the same survey.

7 of the 8 Sound Engineer participants returned a Pro Tools mix session for each of the 6 programme titles, while 1 participant returned an audio consolidated AAF[5] file. The DVs applied by the 8 participants were measured at each of the 317 AD cue point across all 6 programme titles. Each AD cue point yielded 8 unique DV settings, one for each participant. 2,536 (8x317) DV data points in total were returned by the sound engineers. As indicated in Table 12, DVs were applied and measured across a wide variety of programme types and LRAs, reflecting typical day-to-day content across many media delivery platforms.

Taking 'The Bourne Identity' as an item example with the highest number of AD cues (133) and subsequently the highest number of DV data points (1064), each participant's DVs have been plotted against the integrated programme loudness (LUFS) measured at the AD cue points (see Figure 22 below).

[5] The Advanced Authoring Format (AAF) is a professional file exchange protocol for the video post-production industry.

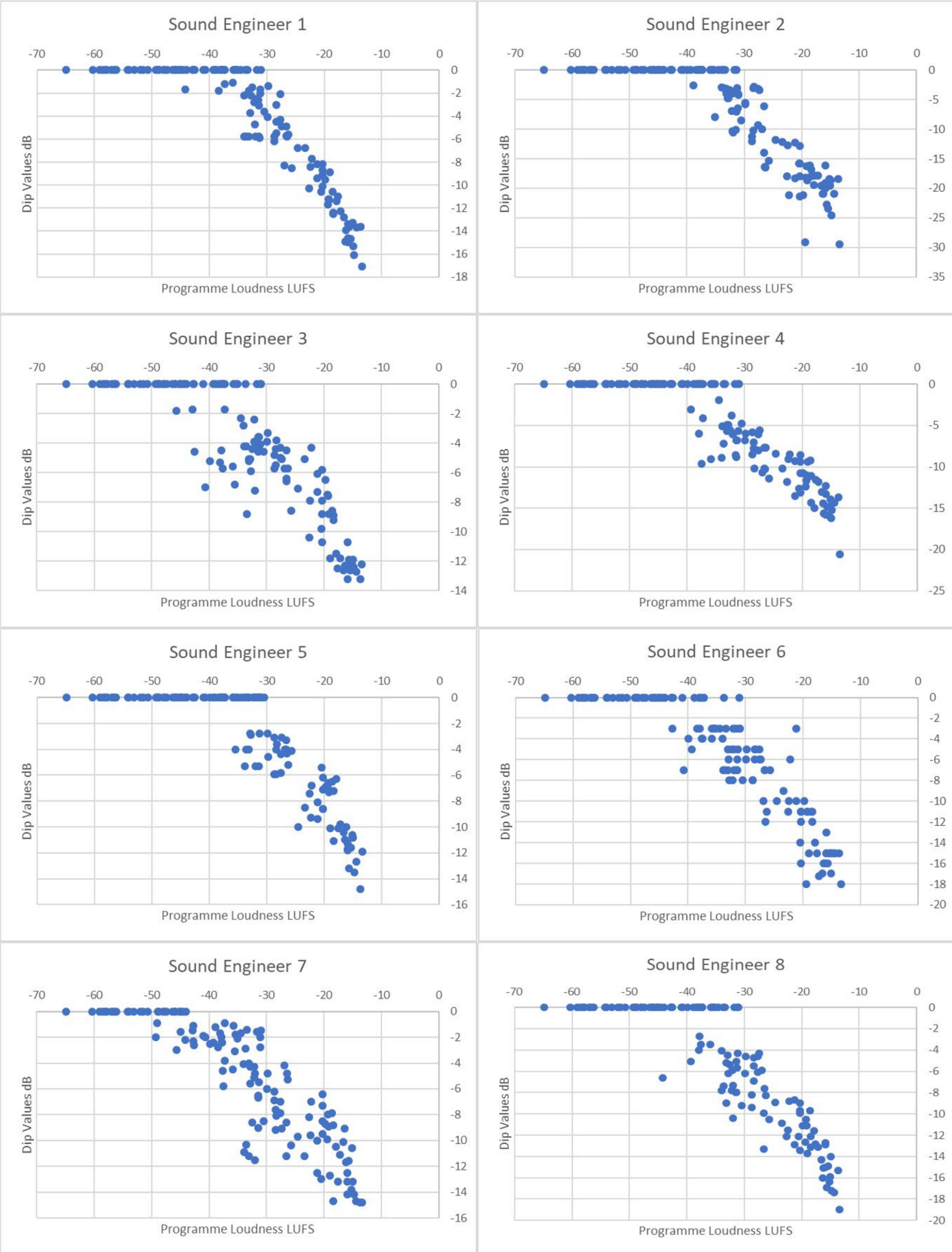


Figure 22: Scatter plots of the DVs applied by 8 Sound Engineers to 'The Bourne Identity' excerpt.

The scatter plots in Figure 22 show that there are nuanced micro differences in the DVs applied by the 8 participants, but also that a strong overall trend in the data exists. The scatter plot of Figure 23 shows the distribution of all 1064 DVs applied by the 8 participants at the 133 AD cue points for ‘The Bourne Identity’ excerpt. Since the measured integrated loudness of the programme material at each AD cued point was the same for each participant, the mean of the 8 DVs applied at each cue point was calculated and is displayed in Figure 24.

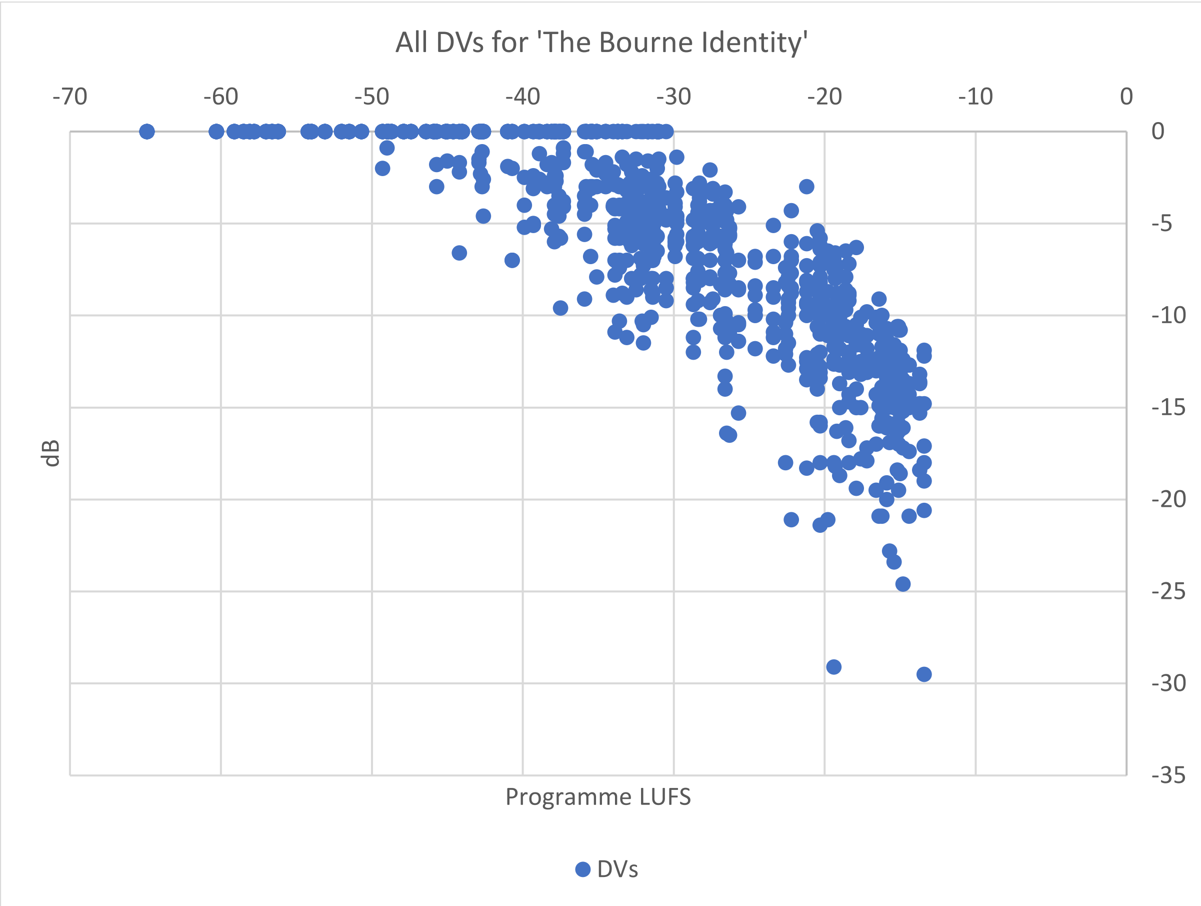


Figure 23: DVs applied at 133 AD cue points by 8 Sound Engineers for the 'The Bourne Identity'.

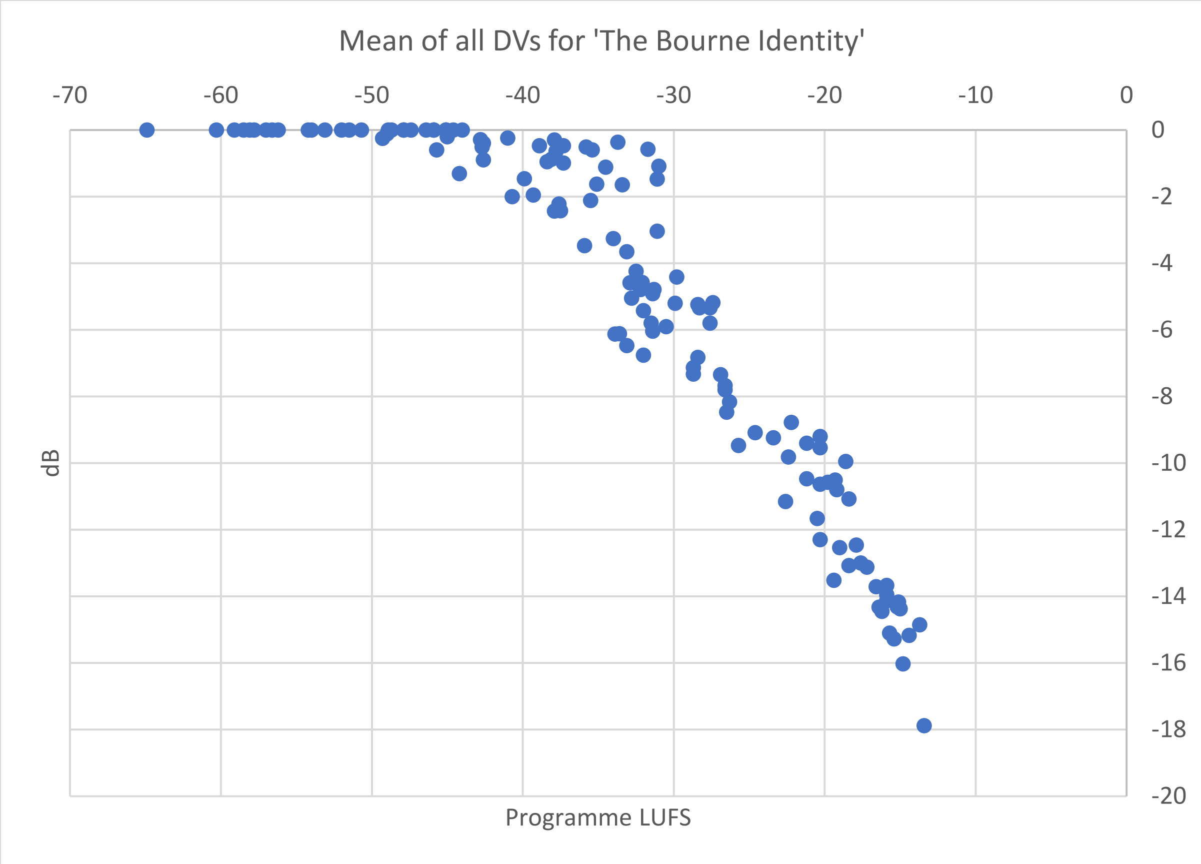


Figure 24: The mean DVs applied at each AD cue point for 'The Bourne Identity'.

The data points displayed in Figure 24 represents not only a distillation of the DV data applied to ‘The Bourne Identity’ excerpt but also the collective trend in the participants’ decision-making process. Raw data such as this is not always easily deciphered and often needs to be rearranged into visual representations that better convey meanings, patterns, and structure.

Statistical Analysis

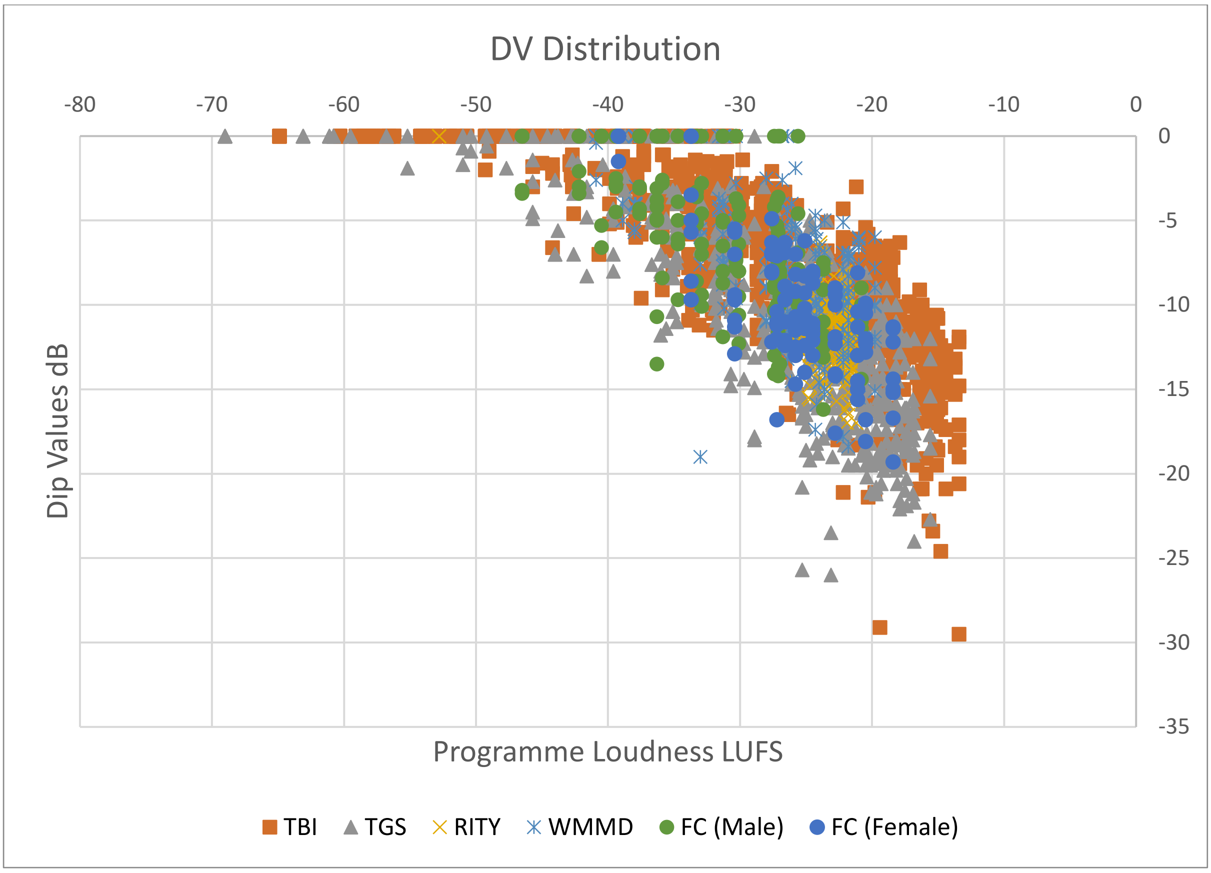


Figure 25: The distribution of the 8 participants' DVs for all 6 programme titles.

As can be seen in Figure 25, the complete dataset for the test does display a form or overall trend, but it is often difficult to precisely predict the exact relationship between the variables. The organising and interpreting of data sets using quantitative modeling allows us to identify predictive relationships between variables in situation where the relationship is not always clear, such as with the broader human response (dependent variables) to audio stimuli (independent variables). Regression, and more specifically Quantile Regression, is a type of statistical analysis that uses the values of multiple response variables to estimate their median regression slope. The main advantage of Quantile Regression over the Ordinary Least Square method is that it makes no assumptions about the distribution of the response variables and is much more robust against outlying observations. Quantile Regression has been chosen as the modelling method for the analysis of this particular dataset.

Assumptions

In order to predict an accurate and reliable median slope for the DV variables, the following assumptions have been made:

Assumption 1: It is assumed that there is a particular integrated programme loudness threshold below which the programme material is sufficiently quiet to negate the need for it to be dipped at the AD cue. Consequently, we assume that AD cues with observed DVs between -1 and 0 dB applied to programme material below this threshold are outliers and their inclusion in the data set will bias the median slope towards a lower programme LUFS value for the lower limit of the 'Zero DV' intercept, and therefore can be ignored. A *Zero DV Threshold* has been determined using a combination of:

a) The histogram in Figure 26 to determine the integrated LUFS value corresponding to the highest frequency of the mean DV events in Figure 27 with values between -0.99 and 0 dB
b) And the observations highlighted by the red dashed line in Figure 27 around the dataset's lower limit 0 dB DV intercept.

For the 6 programme titles investigated, representing multiple genres and a wide range of LRA values, the *Zero DV Threshold* is chosen as -48 LUFS (see Figure 28) which is a reasonable selection given the distribution in Figure 26.

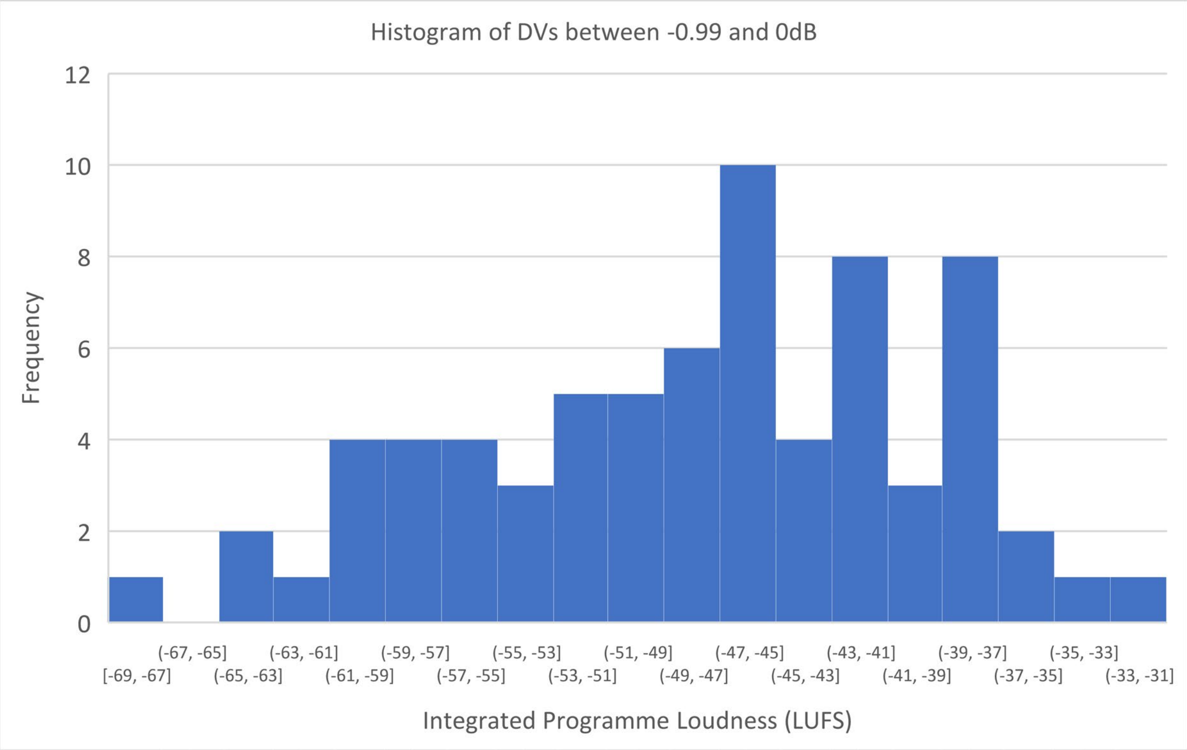


Figure 26: The distribution frequency from Figure 27 of the participants' mean DVs having values between -0.99 and 0 dB over all 6 programme titles.

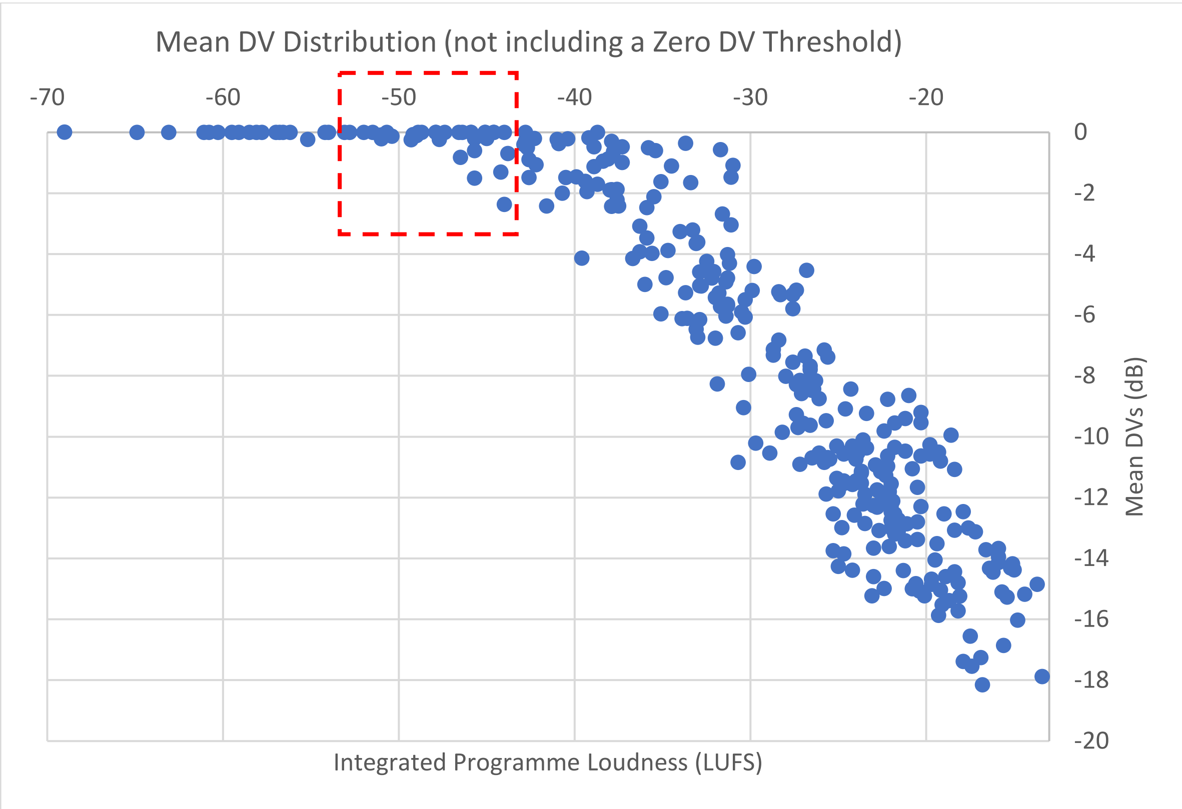


Figure 27: The distribution of the participants' mean DVs applied to all 6 programme titles. The lower LUFS 'Zero DV' intercept lies within the area of the plot highlighted in red.

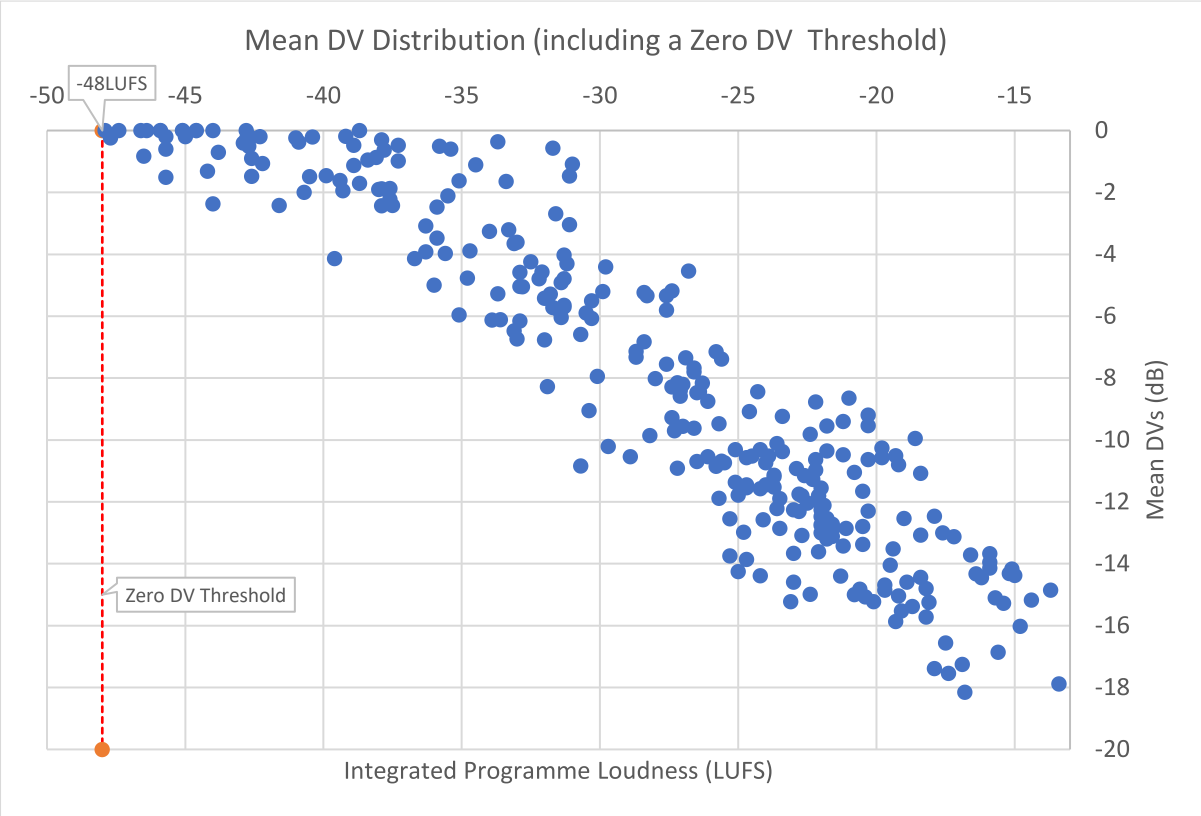


Figure 28: The distribution of participants' mean DVs (excluding those below the -48 LUFS 'Zero DV Threshold') applied to all 6 programme titles.

Assumption 2: It is also assumed, through observation, that the median DV distribution profile in Figure 28 has a logarithmic characteristic and that we are modelling the data such that the target (dependent) variable, is of the form $Y = A + B(logX)$, where Y is the DV, A is the intercept, B the quantile regressor coefficient and $X = -1 * (LUFS)$. The linear transformed data is shown in Figure 29 for which we can apply quantile regression.

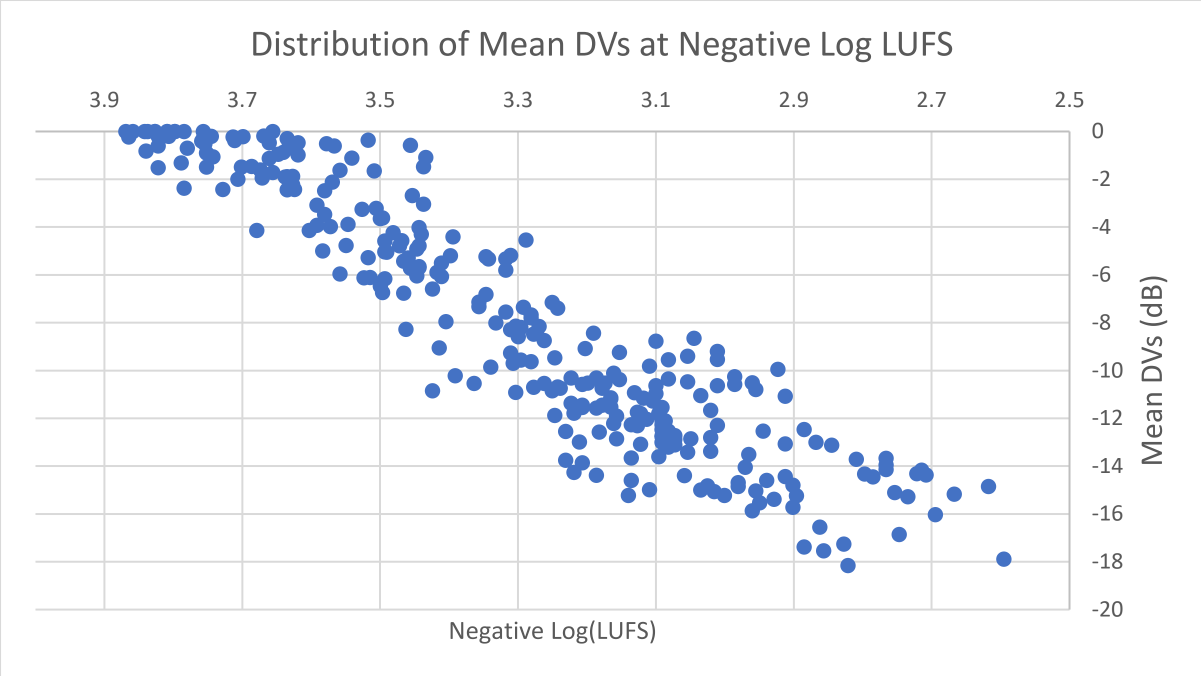


Figure 29:The distribution of the participants' mean DVs in relation to the negative Log of the programme LUFS in preparation for the fitting of a Quantile Regressor.

Fitting A Quantile Regressor

The following three conditional quantiles can now be plotted for the linear models: DV median (0.5 quantile, i.e. the median); the upper DV limit (0.9, i.e. the high quantile); and the lower DV limit (0.1, i.e. the low quantile) (see Figure 30). After the LUFS data is transformed back to its original observed unit and values, the logarithmic nature of the median slope, along with the 10^{th} and 90^{th} percentiles of the target variable are now evident (see Figure 31).

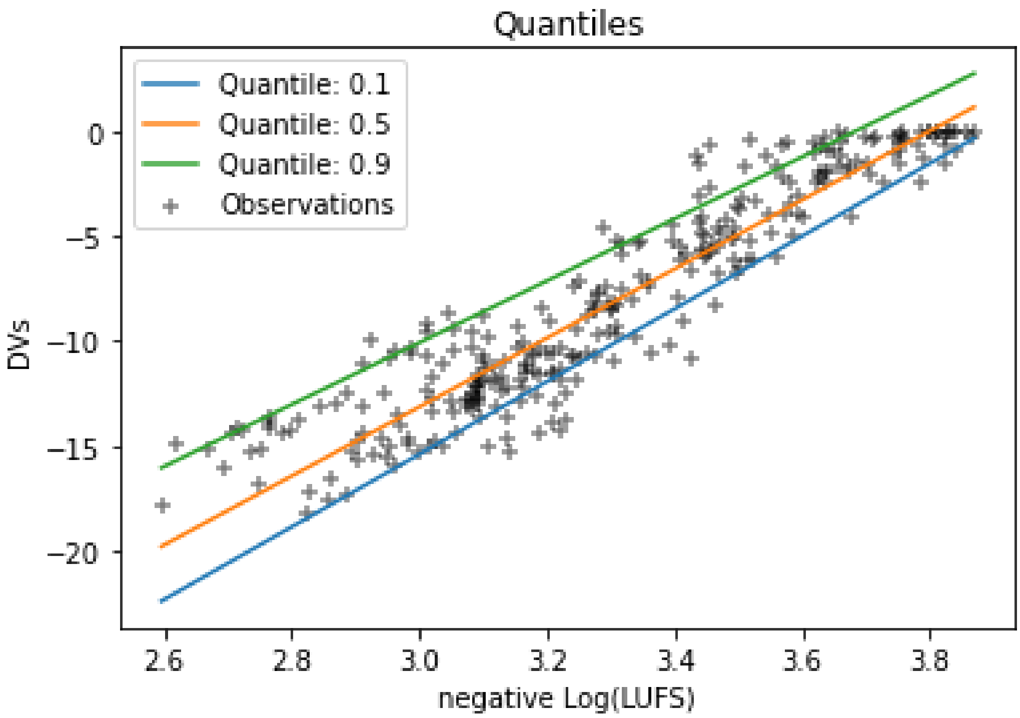


Figure 30 :The linear model predictions for the three fitted Quantiles of the conditional median (0.5 Quantile), the 90th percentile (0.9 Quantile) and the 10th percentile (0.1 Quantile)

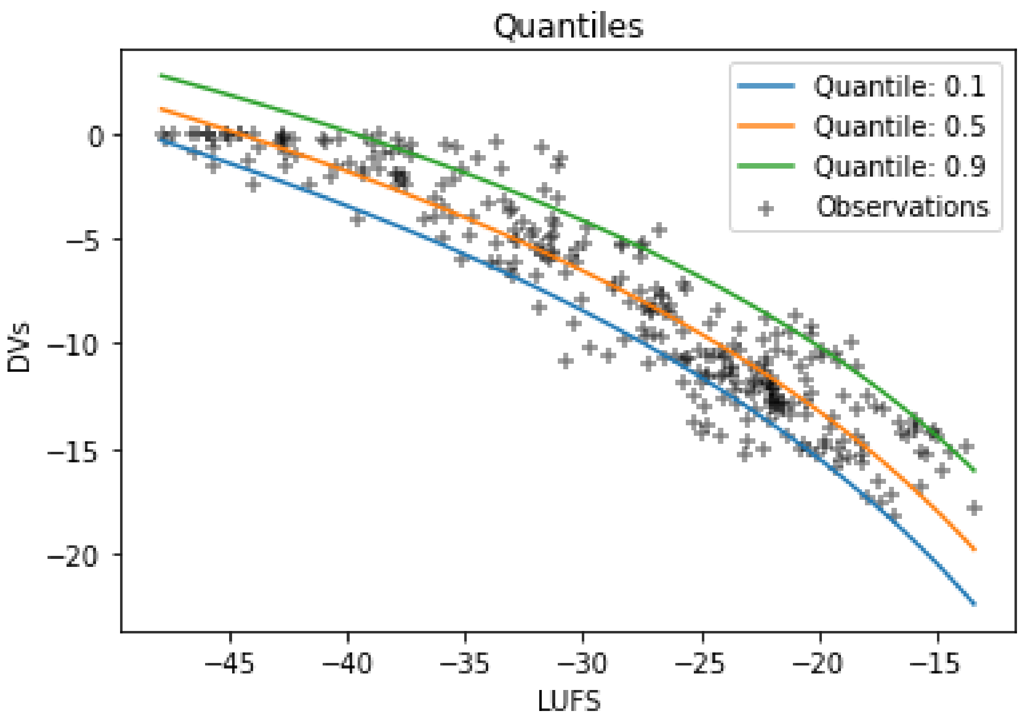


Figure 31: The transformed logarithmic expression of the three fitted Quantiles of the conditional Median (0.5 Quantile), the 90th percentile (0.9 Quantile) and the 10th percentile (0.1 Quantile).

It is these three quantile regression slopes that will be used to determine the DVs for assessment on an audience.

Determining Loudness Bands

This research project aims to determine an audience's preference for the Dip Values applied to programme material and specifically the preferred DVs for programme types classified as *Quiet*, *Normal*, and *Loud*. In order to assess an audience's preference, a criterion for these three loudness bands must be set. Working with the data presented in Figure 27, the Normal Distribution of the integrated programme loudness at the AD cue points from the mean was calculated. Figure 32 shows the probability density for any integrated programme loudness value occurring at an AD cue point. As can be seen from the plot, the median for the 6 programme titles considered lies at -26.60 LUFS. This implies that the majority of the AD cues occur at programme locations having an integrated loudness value of -26.60 LUFS and therefore it is this LUFS value that establishes the centre of the *Normal* programme loudness band.

The Inverse Normal Distribution method was used to determine the LUFS values corresponding to the 0.25 and 0.75 probability density points. They have been calculated to lie at -32 and -21 LUFS respectively. It is these values that determine the lower and upper limits for the *Normal* programme loudness band as the area below the curve between these two values represents 50% of the total probability density.

Subsequently, the area below the curve between the *Zero DV Threshold* (-48 LUFS) and the 0.25 probability density point (-32 LUFS) is considered the *Quiet* programme loudness band, and the area between the 0.75 probability density point (-21 LUFS) and the 1.0 probability density point (0 LUFS) is considered the *Loud* programme loudness band (see Figure 32).

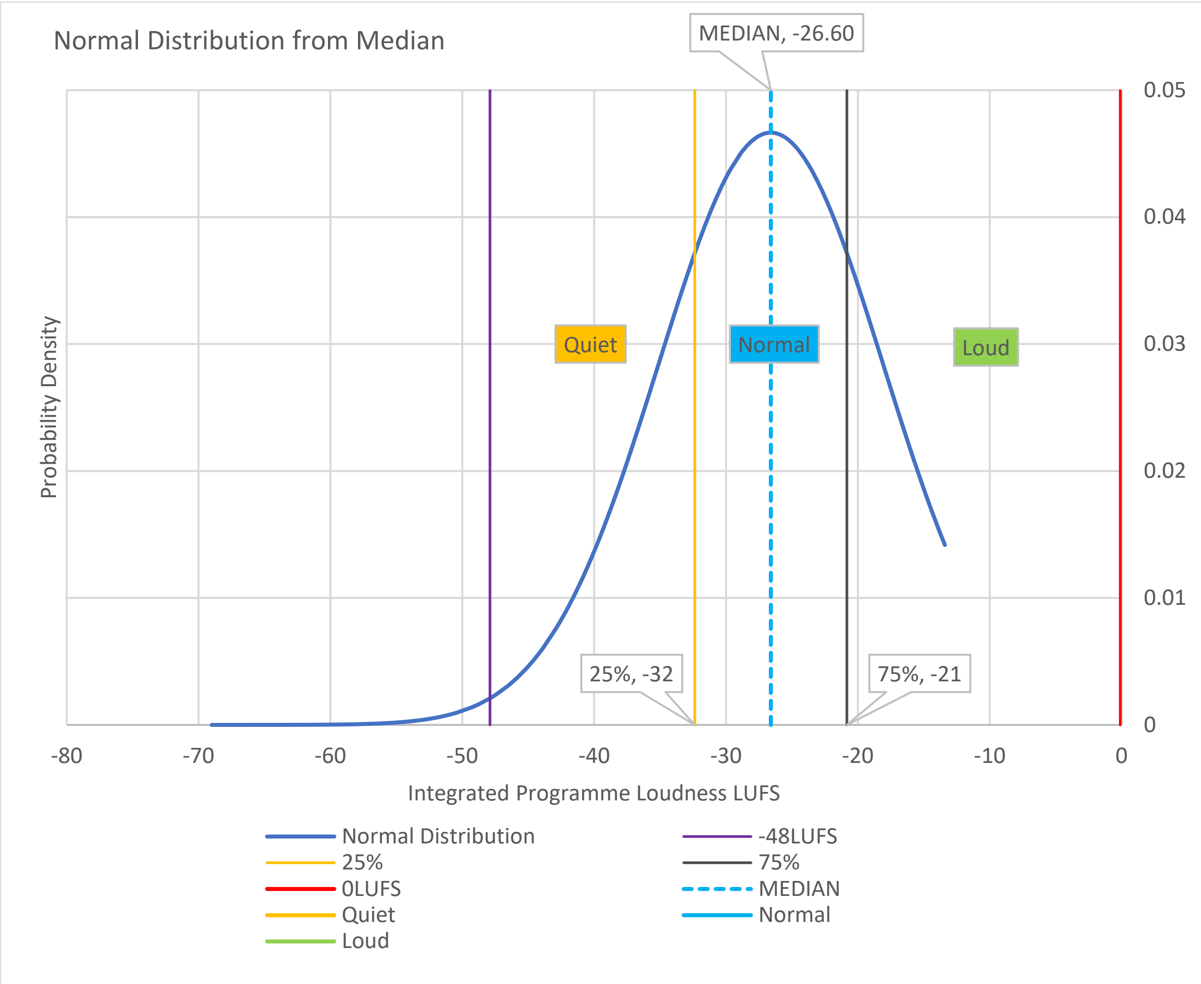


Figure 32: The Normal Distribution from the median for the integrated programme loudness (LUFS) at the AD cue points. The three programme loudness bands Quiet (-48 LUFS to -32 LUFS), Normal (-32 LUFS to -21 LUFS) and Loud (-21 LUFS and above) are also shown.

As in Figure 33 below, the LUFS value corresponding to the 0.125 probability density was calculated to be -36 LUFS. This value is the half-way point of the *Quiet* loudness band and therefore will be used to calculate the mean DV value for *Quiet* programme types. The LUFS value corresponding to the 0.875 probability density was calculated to be -17 LUFS. This value is the half-way point of the *Loud* programme loudness band and therefore will be used to calculate the mean DV value for *Loud* programme types.

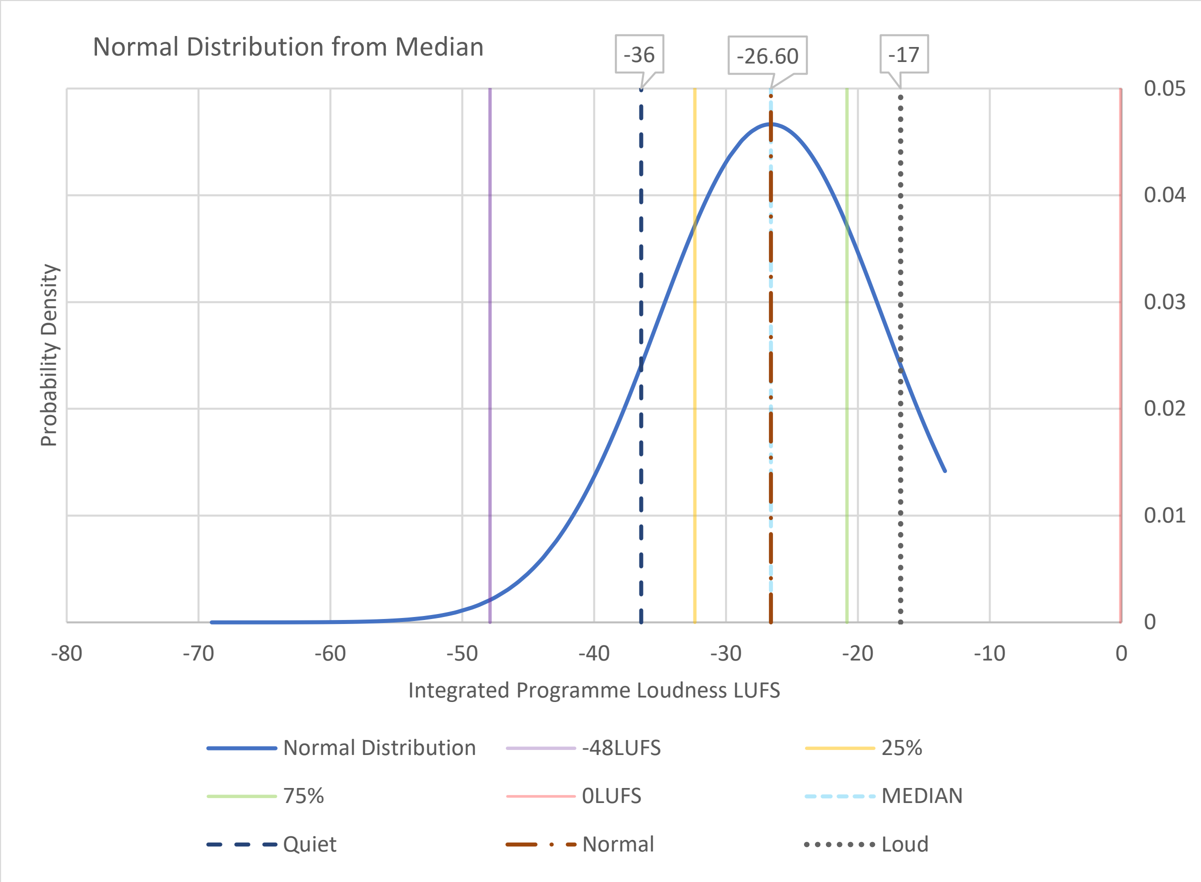


Figure 33: The Quiet, Normal, and Loud programme loudness bands will be represented by the LUFS values -36 LUFS, -26.6 LUFS and -17 LUFS respectively. The corresponding DVs will be determined by the quantile regressor slopes in Figure 31.

Target DVs

Using the LUFS values derived in Figure 33, three corresponding DVs for each of the *Quiet*, *Normal*, and *Loud* programme loudness bands can be determined using the three quantile regression slopes presented in Figure 26. The intersection points shown in Figure 34 yield the nine Dip Values presented in Table 14. This is the set of DVs used for assessment on an audience in Stage 3: Listening Tests.

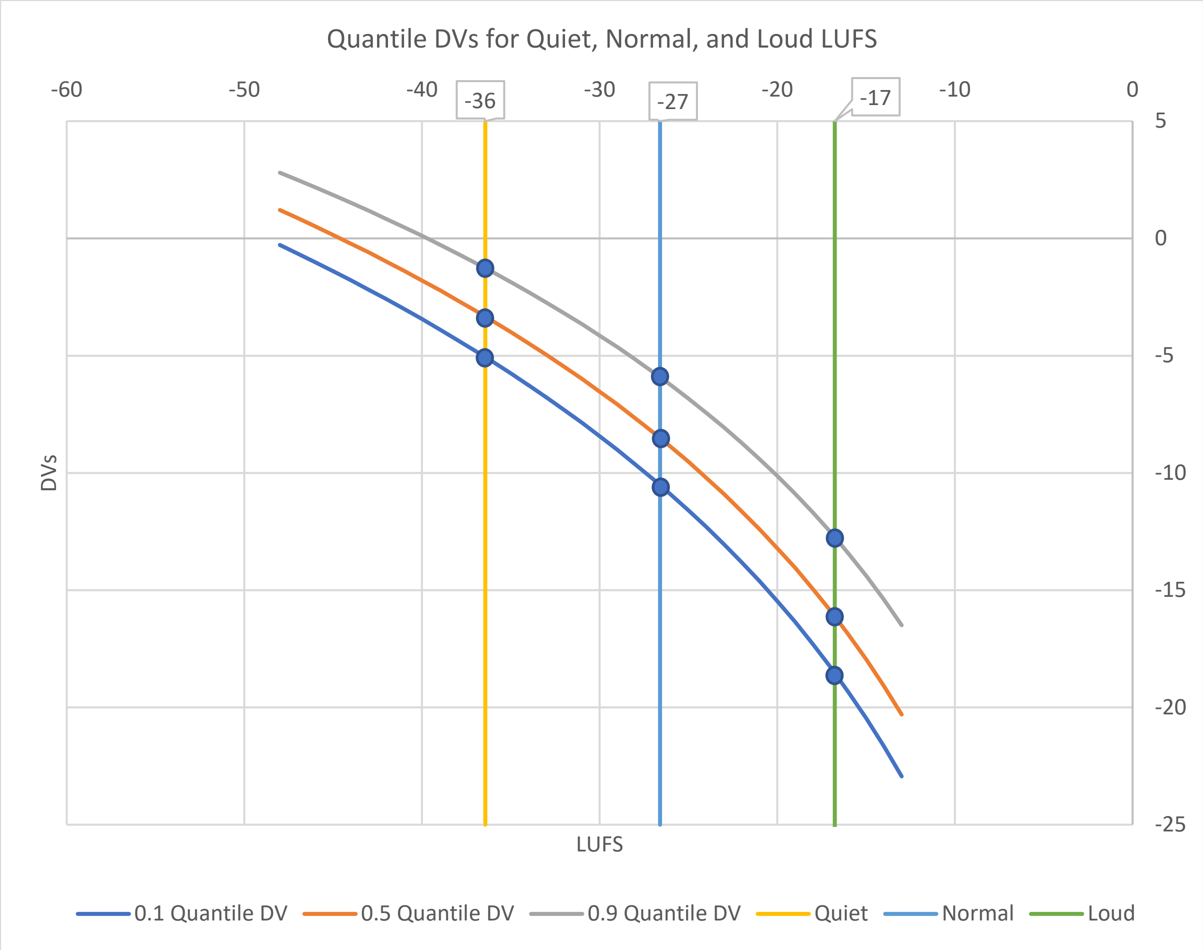


Figure 34: DVs at the intersection points between the Quiet, Normal, and Loud LUFS and the 0.1, 0.5, and 0.9 Quantile Regression slopes.

	Quiet	Normal	Loud
LUFS Range	-47.9 To -33 dB	-33 To -21 dB	-21 To 0 dB
0.1 Quantile DV (dB)	-5	-11	-19
0.5 Quantile DV (dB)	-3	-9	-16
0.9 Quantile DV (dB)	-1	-6	-13

Table 14: The selected DVs for the Quiet, Normal, and Loud programme loudness bands.

Conclusion

AD Dip Values are an important part of the AD production process and the lack of a standardised approach to setting DVs has led to inconsistencies observed in commercial AD services. By analyzing the common practice of professional mix engineers, we have been able to show that the setting of AD Dip Values is highly correlated to the both the programme material's mean integrated loudness (LUFS) at the AD cue point, and that there is also clear correlation between the Sound Engineers' chosen playback loudness (LUFS) of an AD narration track and the programme material's LRA value.

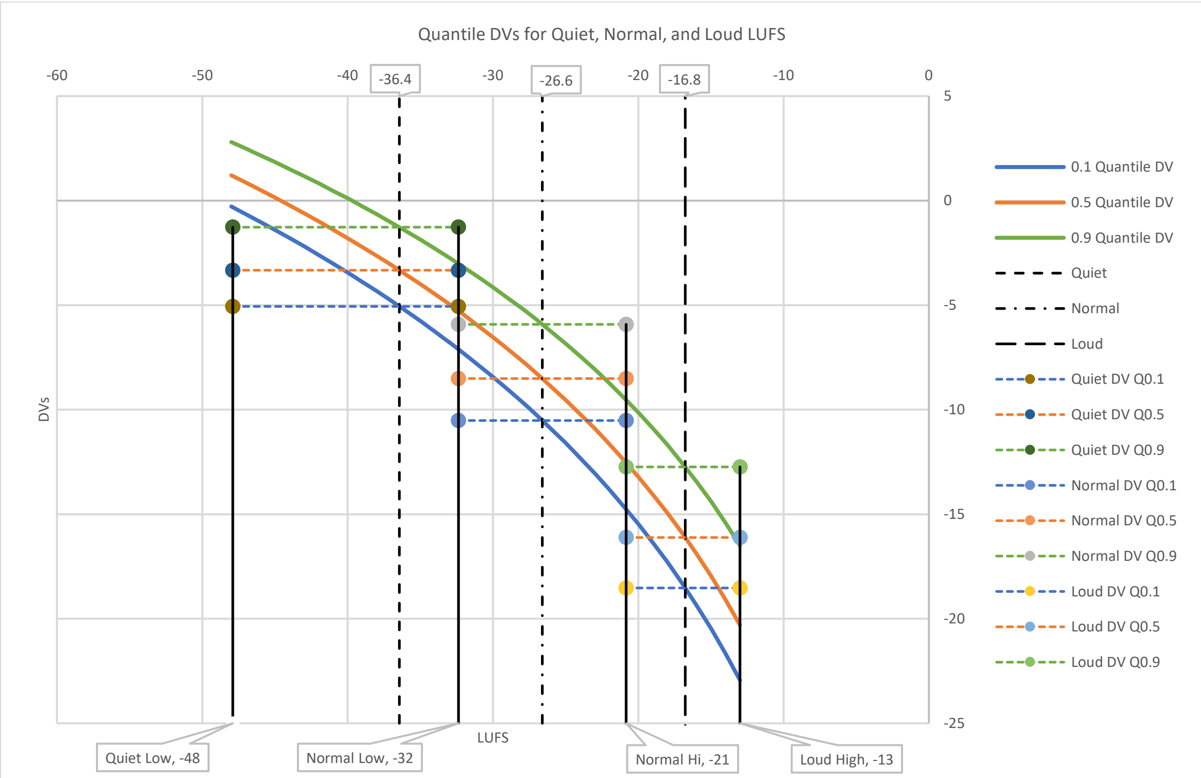


Figure 35: The three DV settings for Quiet, Normal, and Loud programme loudness ranges.

In Table 14 we have presented a set of three DVs for each of the *Quiet*, *Normal*, and *Loud* programme types, and Figure 35 above shows these values as derived by the 3 quantile regressor slopes for the participants' mean DV dataset of Figure 29. These DV sets are consistent with values proposed in [9] for a similar programme dip variable known as Loudness Difference (LD) and provide a range of Dip Values for assessment on an audience. This assessment was carried out using listening tests designed to elicit an audiences DV preference for *Loud*, *Normal*, and *Quiet* programme types and is presented in the next section of this report.

Stage 3: Listening Tests

Listening tests were carried out over two weekends in February 2023 across a number of subject groups. The aim of the tests was to determine the preferred quantile DVs of those presented in Table 14 for three programme types congruent with the *Loud*, *Normal*, and *Quiet* loudness bands as determined in Figure 32.

Listening Test Setup

Participants: The listening tests were primarily focused on AD service users but also included candidates from a wider end-user profile and therefore involved 11 non-expert listeners who identified as blind (average age 44), 8 vision impaired non-expert listeners (average age 44), and 8 sighted non-expert listeners (average age 34). 8 expert listeners were also invited to participate in the listening tests. The expert listener group consisted of the same 8 mix engineers that participated in the Stage 2: Benchmarking experiment. All participants had English as their first language and were verified not to have any known hearing impairments. It is suggested that in future research a similar listening test be carried involving a subject group with known hearing impairments as an AD user with hearing impairment may have DV preferences that differ to those users who do not. Moreover, the service should ideally be as inclusive of, and function well for, as wide an audience as possible. All non-expert listeners received a gratuity for participating in the listening tests.

Method: Three listening tests were carried out and each was a multiple stimuli test in which three different DV conditions were presented. Because of issues around accessibility for the blind and vision impaired participants, it was decided that a more suitable rating scheme be applied. Therefore, rather than use the typical rating-based scales or sensory profiling found in [20], participants were asked to choose the condition containing their preferred DV for the given programme type, or to indicate if they had no preference. This yielded a more accessible process for all. All further questions were then asked in relation to the subject's preferred DV condition only. Data relating to the participants' responses was gathered using an online Microsoft Forms™ questionnaire.

Conditions: For each listening test, three DV conditions were presented in a randomized sequence across participants in an attempt to minimise order bias[6]. It is possible that some contrast bias may still exist in sequences moving from the largest DV condition to the smallest or vice versa, affecting how participants perceive the differences between items.

For the listening test correlating to each of the *Loud*, *Normal*, and *Quiet* programme types, three DV conditions were presented as determined by the 0.1, 0.5 and 0.9 quantiles in Table 14. In the first listening test, DV conditions of -19, -16, and -13 dB were presented for a *Loud* programme type containing integrated loudness values between 0 and -21 LUFS at the AD cue points. The second listening test presented the three DV conditions of -11, -9, and -6 dB for a *Normal* programme type with integrated loudness values between -21 and -33 LUFS at the AD cue points. And the third listening test presented the DV conditions of -5, -3, and -1 dB for a *Quiet* programme type containing integrated loudness values between -33 and -48 LUFS at the AD cue points.

The integrated loudness of the AD narration track was unique for each of the *Loud*, *Normal*, and *Quiet* programme types. It was determined, as outlined in Stage 2: Benchmarking, by the full programme's Loudness Range (LRA) using the look-up table in Table 17, Appendix 4. To note, even though the listening tests were carried out using 1-minute excerpts, it is the full programme's LRA value that is used in Table 17 to determine the playback LUFS value

[6] Order bias refers to a type of cognitive bias that can occur when the order in which items are presented to an individual has an influence on their judgment or decision-making.

for the AD narration track, as ultimately it is an AD service appropriate to the full duration of the programme title that an audience will access. The integrated loudness of the AD narration track was set to -27.1 LUFS for the 3 conditions associated with the *Loud* programme type, -24 LUFS for the 3 conditions associated with the *Normal* programme type, and -25 LUFS for the 3 DV conditions associated with the *Quiet* programme types (see Table 15 below).

	Quiet	Normal	Loud
LUFS Range	-47.9 To -33 dB	-33 To -21 dB	-21 To 0 dB
0.1 Quantile DV (dB)	-5	-11	-19
0.5 Quantile DV (dB)	-3	-9	-16
0.9 Quantile DV (dB)	-1	-6	-13
AD Narration LUFS	-27.1	-24	-25

Table 15: Updated Table 14 to include the integrated loudness values used for each of the AD narration tracks used with the 'Loud', 'Normal', and 'Quiet' programme types, as determined by Table 17.

Programme Types: The three listening tests used 1-minute excerpts from three of the same programme materials used during the benchmarking experiment carried out with the mix engineers in Stage 2: Benchmarking of this research project.

For the *Loud* programme type, the three conditions were presented using an excerpt from a car chase-scene from the motion picture ‘The Bourne Identity’[14]. This scene was chosen as it has a high integrated loudness value at the AD cue points and is reflective of the loudest media content broadcast to include an AD service. The excerpt contained one 2-second dialogue clip, which was useful for participants to calibrate voice levels, a blend of loud impactful sound effects, loud background and foreground music in the form of score, and was deemed to be a good example of loud sound effects and music driven programme material containing little or no dialogue. The full movie has a wide overall LRA value measured at 22.4 LU.

The *Normal* programme material was presented using a similar length excerpt from the television series 'Reeling In The Years'[16]. The chosen excerpt offered consistent integrated loudness values at the AD cue points and the full programme title has a narrow measured LRA of 6 LU. The excerpt contained news footage, a 4 second dialogue clip at the start, a commercial pop ballad with lyrics as both background and foreground music throughout, and was deemed a good example of normal programme-loudness material with a strong music focus.

The *Quiet* programme material used a 1-minute excerpt from the television series 'Fair City'[17]. This title is from the soap opera genre and contains good examples of quiet integrated loudness at AD cue points. The full programme has a measured LRA value of 10.7 LU and consisted of mainly quiet interior scenes with low level background audio at the AD cue points that included room tone ambience, production sound effects, and distant traffic. The excerpt was mainly dialogue driven and did not contain any music.

A fourth example was created for illustration purposes only. It was presented to the participants at the beginning of the test experiment but was not used to gather data during the listening tests. This illustration example was of a *Loud* programme type and consisted of a 1-minute excerpt from the motion picture 'The Greatest Showman'[15]. This excerpt was considered to be an example that illustrated well the DV relationship between the AD narration and the background programme audio at the AD cue points during a loud musical passage.

The experiment started with the three conditions corresponding to the *Loud* programme type, followed by the *Normal*, and then the *Quiet* type, as it was considered easiest for the participants to attune to the DVs applied during loud programme passages, and therefore focusing their listening as each subsequent listening test presented smaller DVs for consideration against quieter background programme audio.

Instructions: The following instructions were given to the participants before the test experiment began:

> "*The aim of the study is to determine your preference for a volume-relationship between the programme material and the audio descriptions. That is, the balance that must be struck between reducing the volume of the programme material so that the audio descriptions are clear, while also maintaining the programme's narrative and structure.*
>
> *You will listen to three versions of the same Loud programme material and then will be asked to pick your preference, only in terms of the volume-relationship between the programme material and the audio descriptions. All questions will then be in relation to your preferred choice. This experiment will be repeated for examples of Normal, and Quiet programme material.*"

Locations: The listening tests involving the non-expert listeners took place at the Training Centre at the National Council for the Blind of Ireland (NCBI). A room was chosen that most reflected the average home TV listening environment (see images in Appendix 5). The expert-listener group carried out the test at the same locations used during the benchmarking experiment and outlined in Stage 2: Benchmarking.

Test Procedure: The listening tests involved 9 items: 3 DV conditions (0.1, 0.5, and 0.9 quantiles Dip Values) for each of the 3 programme types (*Loud*, *Normal*, and *Quiet*). The items were stereo WAV files with a 48Khz sample rate, 24 Bit depth, and were all 1 minute in duration. Each test item featured a mono, centre panned, AD track narrated in English and pre-mixed using the conditional DVs of Table 15 against stereo programme material. The *Loud* and *Quiet* programme types featured male narrators and the *Normal* programme type featured a female narrator. No accompanying video was shown.

Subjects participated in the listening tests individually or in groups of two or three. Participants were seated 2.5m from the stereo playback speakers (see images in Appendix 5). The playback speakers were calibrated using the procedure outlined in [21] but the suggested reference listening level of 73 dBC SPL was deem excessively loud and not representative of the typical listening levels used in a home setting, and therefore what was considered a more appropriate reference listening level of 61 dBC SPL was used. The listening level remained fixed during the listening tests.

At the start of the test, a set of explanatory notes and instructions were read to the participants and the illustration example was played. Participants then had the opportunity to ask questions or have anything clarified before the listening test started proper. Each condition was played in a randomized sequence for the first (*Loud*) listening test, participants were then asked to complete the section of the online questionnaire relating to that programme type. This procedure was repeated for the second (*Normal*) and third (*Quiet*) listening tests. Participants were given the choice to complete the online questionnaire themselves or have a research assistant complete it on their behalf during the listening tests. The listening test questionnaire is available in Appendix 6. No formal debriefing was necessary when the listening tests were completed.

Listening Test Findings

Participant Profile

The aim of the listening tests was to determine if a common preference exists among AD service users for the programme loudness Dip Values applied to *Loud*, *Normal*, and *Quiet* programme types. The non-expert listener group were also invited to complete a questionnaire on the day, and from this survey the following data was gathered.

Vision Profile: As in Figure 36, 71% of non-expert listeners that participated in the listening tests reported to have a form of vision impairment. 41% of participants classified themselves as *Blind*, 30% as *Vision-impaired*, and 29% as *Sighted*.

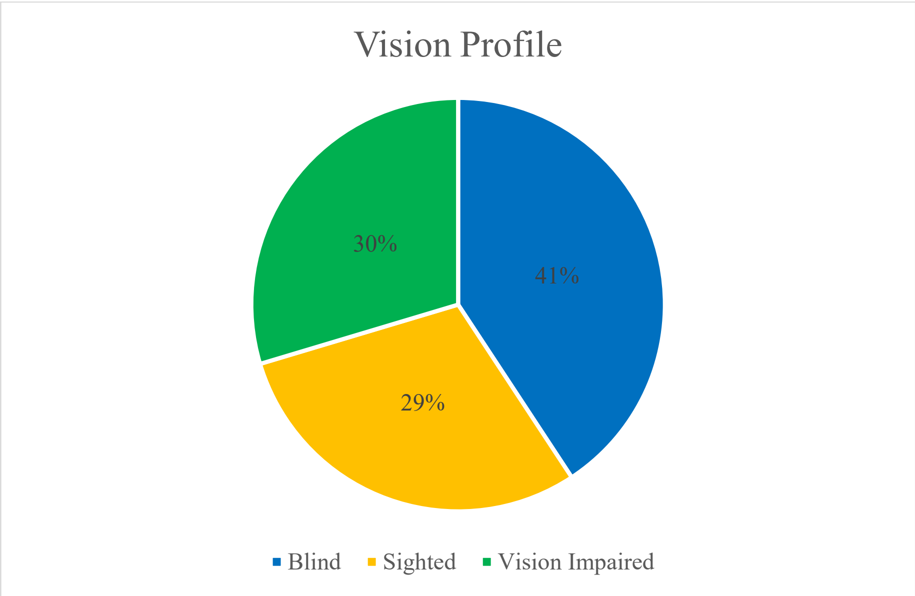


Figure 36: Vision profile of the non-expert listener participants

Viewing Habits: Focusing on the principal AD service user group of participants classifying themselves as either *Blind* or *Vision Impaired*, Figure 37 shows that when asked '*How often do you use Audio Description services?*', 27% of respondents within this user group chose *Daily*, 21% said *Weekly*, 37% chose *Occasionally*, and 5% responded as *Monthly*, *Rarely*, and *Never*.

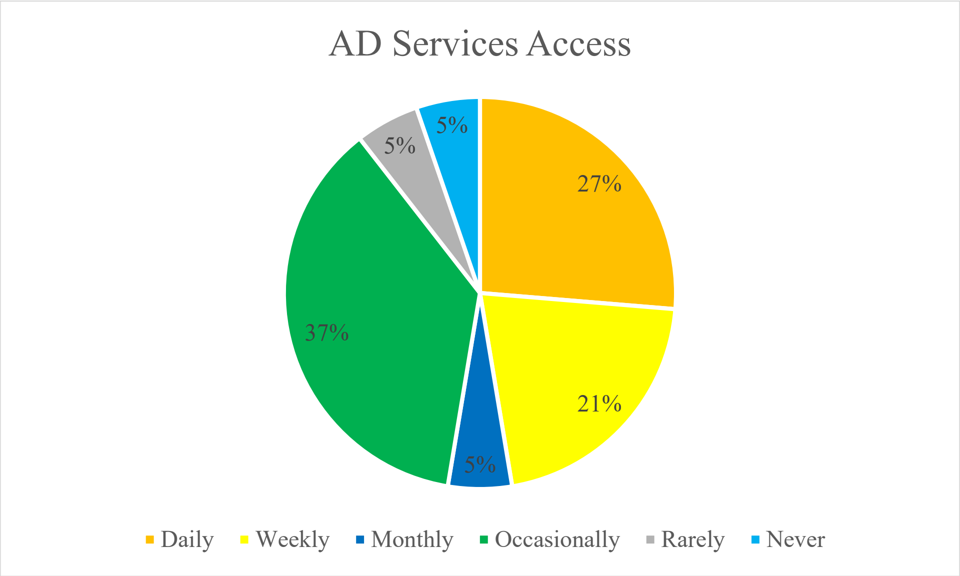


Figure 37: Frequency of AD services access among participants who classified themselves as either 'Blind' or 'Vision Impaired'.

Further analysis shows that if an AD service access frequency of *Monthly* or less is classified as *Infrequent* and an access frequency of *Weekly* or more is classified as *Frequently*, it can be seen from Figure 38 that the majority (53%) of participants who classify themselves as *Blind* or *Vision Impaired* access AD services *Infrequently*.

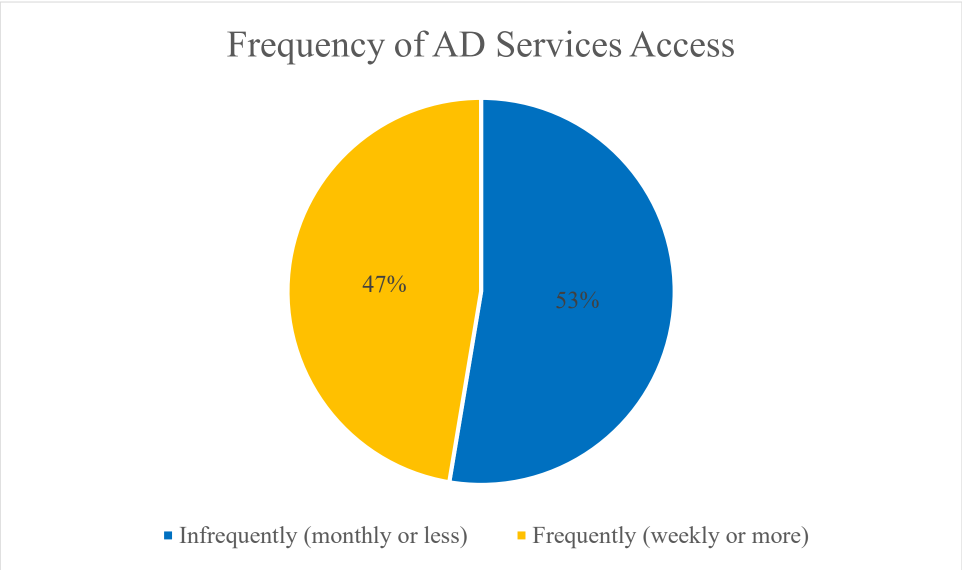


Figure 38:A more generalised outline of how frequently 'Blind' and 'Vision Impaired' participants access AD services.

Qualitative data on the reasons why participants might access AD services in broadcast *Infrequently* was gathered through conversational responses, with respondents stating "*I don't know how to turn it on*", "*I never know if it's available*", or "*I'm just not in the habit*". This would suggest that perhaps a dedicated campaign to raise awareness of AD services and how to access them could benefit the listening experience of the subject audience.

Loud Programme Types

Figure 39 below shows the preferred *Loud* programme type DV distributions for all participants (expert and non-expert listeners) taking part in the first listening test. This listening test presented the three DV conditions of -13, -16, and -19 dB for an excerpt from the film 'The Bourne Identity' with an integrated loudness of -21 LUFS or higher at the AD cues. When asked to choose their preferred DV or indicate if they have no preference, 49% of participants opted for the larger Dip Value of -19 dB corresponding to the 0.1 percentile curve of Figure 35. 31% chose the median value of -16 dB, and 20% of participants preferred the upper DV limit of -13 dB. No participant indicated *No Preference* for the conditions presented in the first listening test. 59% of all participants said that the AD narration in their preferred DV choice was *Clear* and 41% indicated that the AD was *Very Clear*. 100% of participants agreed with the statement "*The Audio Descriptions enhanced my listening experience*".

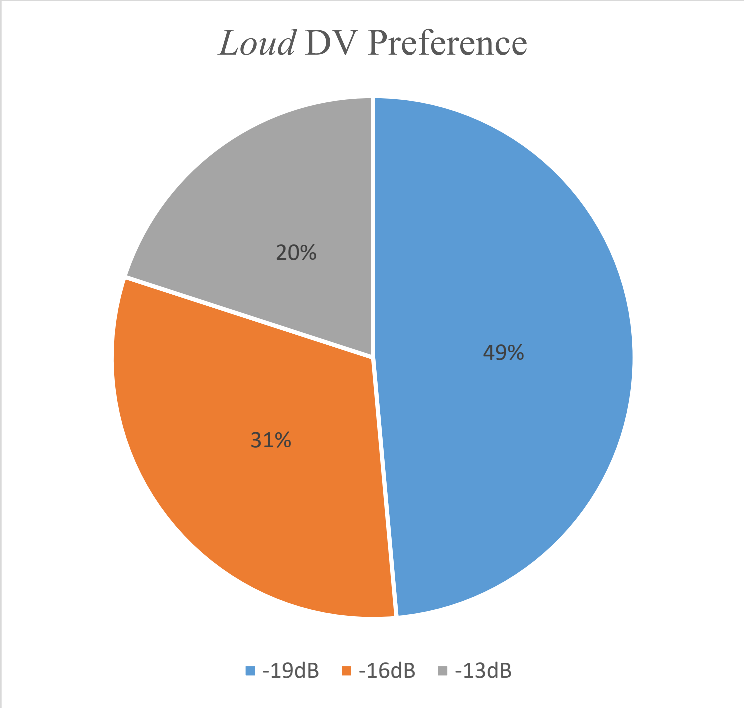


Figure 39: Distribution of preferred 'Loud' DV condition for all listening test participants.

From these results it can be seen that the majority of participants prefer a DV that is 3 dB greater than the median value calculated during the benchmarking experiment for *Loud* programme material, and Figure 40 shows that 67% of non-expert listeners felt that their preferred DV choice was "Just right" for this programme type.

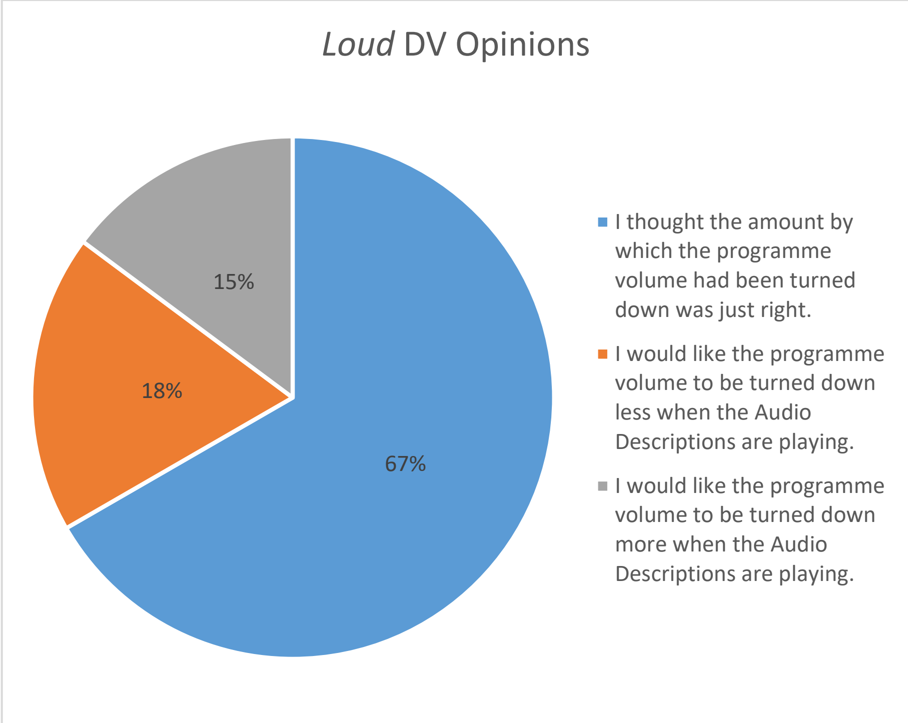


Figure 40: Opinions given by non-expert listener participants on their preferred DV choice for a 'Loud' programme type.

When comparison is made between the expert and non-expert listening groups, Figure 41 shows that the DV preference distribution for the expert listener group, consisting of the 8 mix engineers is centred around the 0.5 quantile value of -16 dB. This is as expected given that the 0.5 quantile Dip Value is derived from the median Dip Values applied by the same expert listeners during the benchmarking experiment of Stage 2: Benchmarking. The DV preference distribution for the non-expert listener group is weighted towards the 0.1 quantile value of -19 dB. The 3 dB difference observed between the two groups correlates well with the findings presented in [9], [22], and [23].

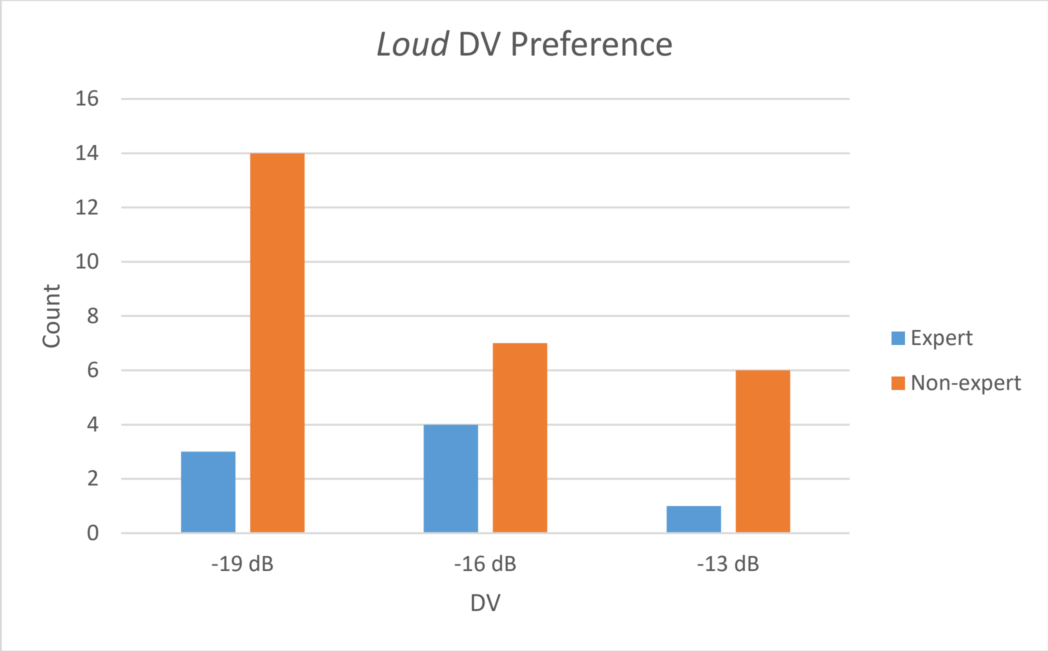


Figure 41: Distribution of preferred 'Loud' DV condition among the 'Expert' and 'Non-expert' listener groups.

It can be seen from Figure 42 that the distribution of preferred DVs for sighted participants is weighted more towards the 0.5 median quantile of -16 dB, whereas vision-impaired participants have a stronger preference towards the larger DV of -19 dB corresponding to the 0.1 quantile, or lower DV limit.

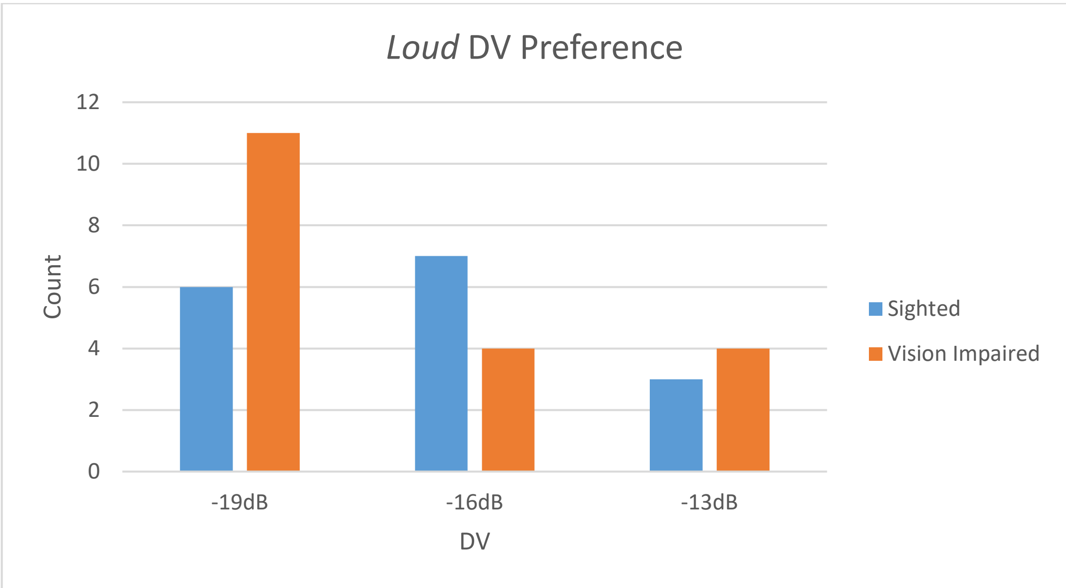


Figure 42: Distribution of preferred 'Loud' DV condition among the 'Sighted' and 'Vision Impaired' subject groups.

The *Sighted* distribution correlated with the distribution shown for the *Expert* listener group. All expert listeners were sighted. The *Vision Impaired* distribution correlates well with the *Non-expert* listener group. All bar one vision impaired participants were non-expert listeners.

Normal Programme Types

Figure 43 shows the preferred *Normal* programme type DV distributions for all participants taking part in the second listening test. This listening test presented the three DV conditions of -6, -9, and -11 dB for programme material with a measured integrated loudness value between -21 and -33 LUFS at the AD cues. When asked to choose their preferred DV or indicated if they have no preference, 31% of participants preferred the larger Dip Value of -11 dB corresponding to the 0.1 percentile curve of Figure 35. 37% chose the median value of -9 dB, and 20% of participants preferred the upper DV limit of -6 dB. 12% of participant indicated *No Preference* for any of the conditions. 78% of all participants said that the AD narration in their preferred DV choice was *Very Clear* while 22% indicated that the AD was *Clear*. 100% of participants agreed with the statement "*The Audio Descriptions enhanced my listening experience*".

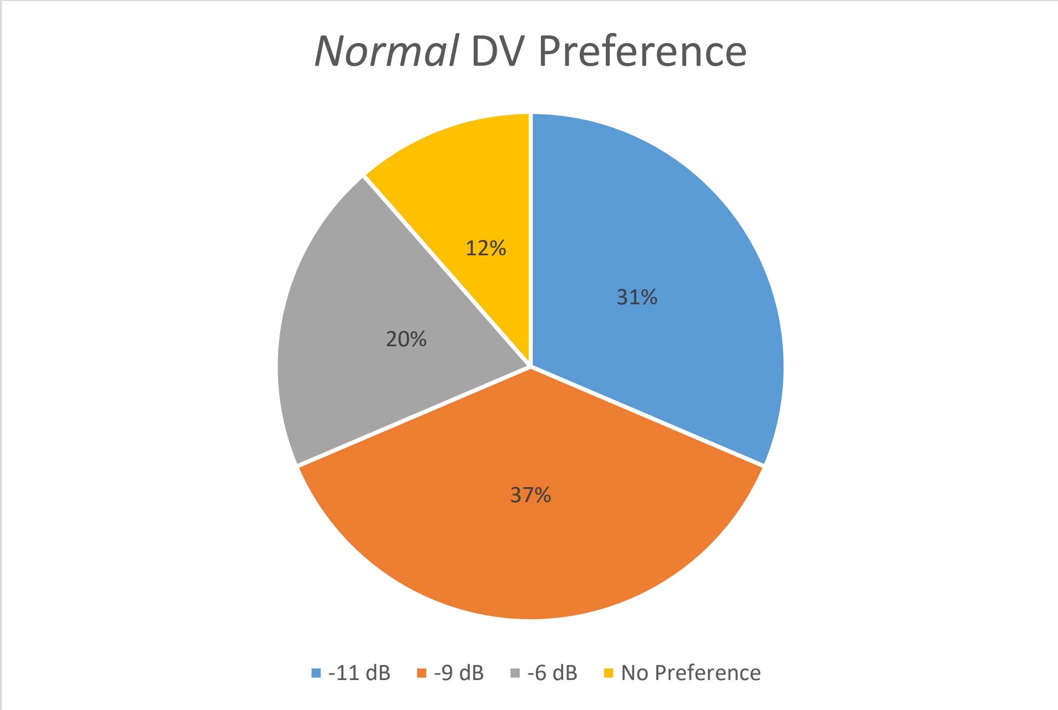


Figure 43: Distribution of preferred 'Normal' DV condition for all listening test participants.

From Figure 43 it can be seen that the DV preference distribution is less defined. The overall preference has moved from the larger DV towards the median or 0.5 quantile value. Here, we also see for the first time that some participants have no preference for any of the presented DV conditions. This would suggest that as the integrated loudness of the programme material reduces at the AD cue point so too does the required Dip Value in order for the AD narration to remain clear, and that as the difference in DV conditions becomes smaller the listeners' DV tolerance increases resulting in a less well defined DV preference. The data presented here correlates with "a so-called 'comfort zone' of about 8 LU around the Target Level" presented in [21].

This is reflected further in Figure 44 and Figure 45 showing the distribution of DV preferences for the *Expert* and *Non-expert* listening groups, and the *Sighted* and *Vision Impaired* groups respectively.

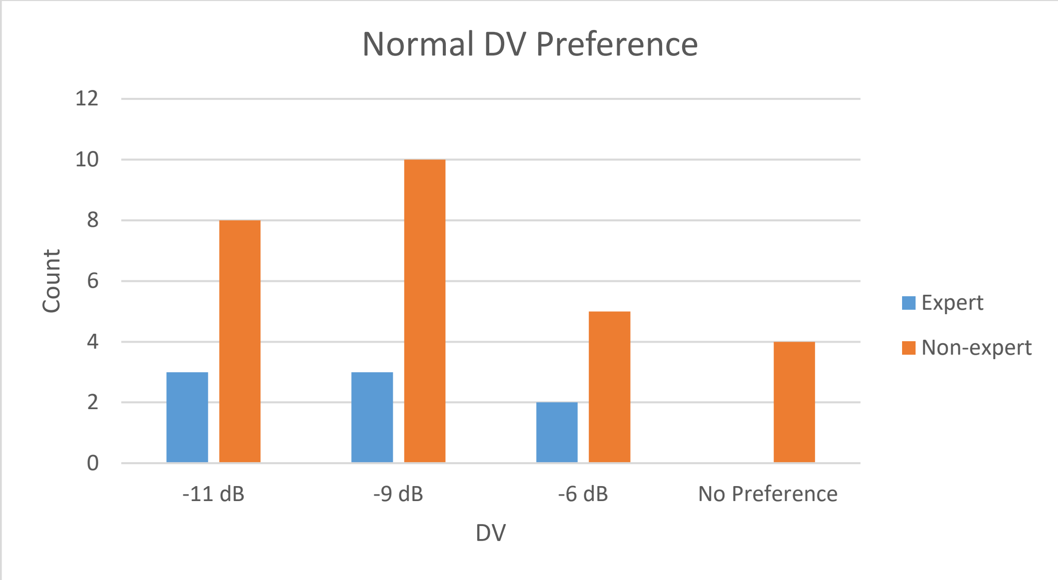


Figure 44: Distribution of preferred 'Normal' DV condition among the 'Expert' and 'Non-expert' listener groups

.

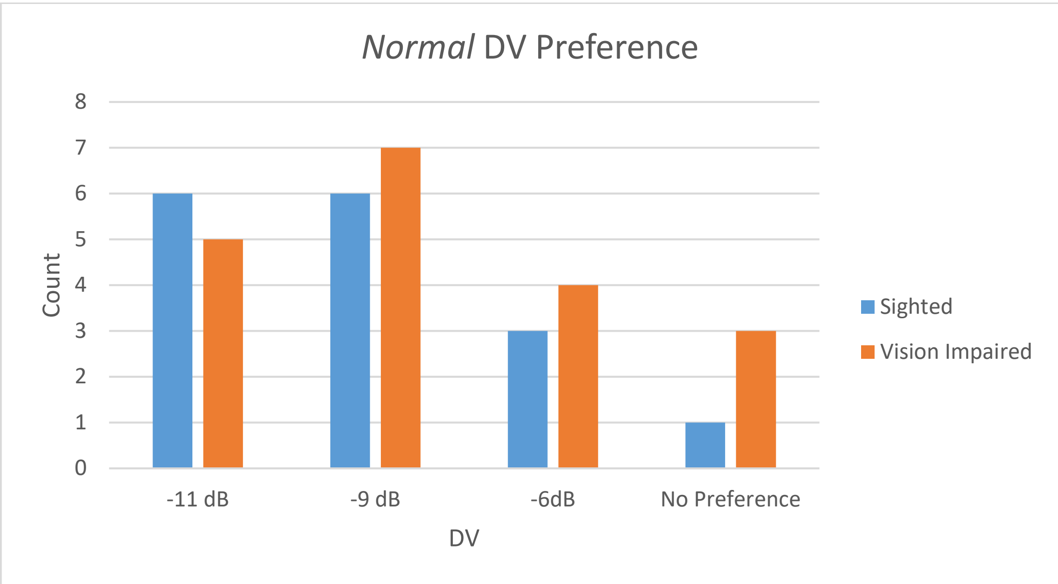


Figure 45: Distribution of preferred 'Normal' DV condition among the 'Sighted' and 'Vision Impaired' subject groups.

Again, the above figures show a less defined DV preference overall, and a migration towards the median or 0.5 quantile value for a *Normal* programme type.

As in Figure 46, the overall satisfaction in DV choice is still high, with 63% of respondents agreeing with the statement that "*I thought the amount by which the programme volume had been turned down was just right*". This would suggest that, although a defined overall DV preference is less clear, confidence in their choice is high among respondents. This again correlates well with the concept of the 'comfort zone' presented in [21]. As with the *Loud* programme type, 15% of respondents said "*I would like the programme volume to be turned down more when the Audio Descriptions are playing*", while only 7% felt "*I would like the programme volume to be turned down less when the Audio Descriptions are playing*".

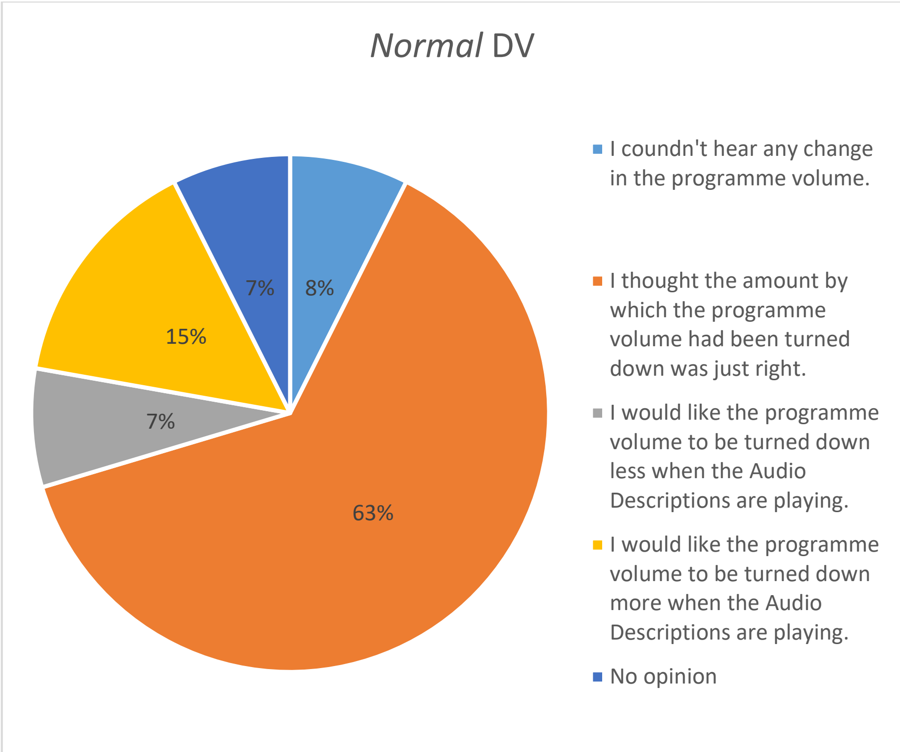


Figure 46: Opinions given by non-expert listener participants on their preferred DV choice for a 'Normal' programme type.

Quiet Programme Types

The distribution of DV preferences for the three conditions presented in the third listening test for a *Quiet* programme type can be seen in Figure 47. Here, the overall preference is again focused around the median DV of -3 dB. The DV conditions presented in this listening test were subtle and required concentrated listening for a difference to be distinguished and therefore it can be seen that a larger percentage of participants opted for the *No Preference* choice (20%).

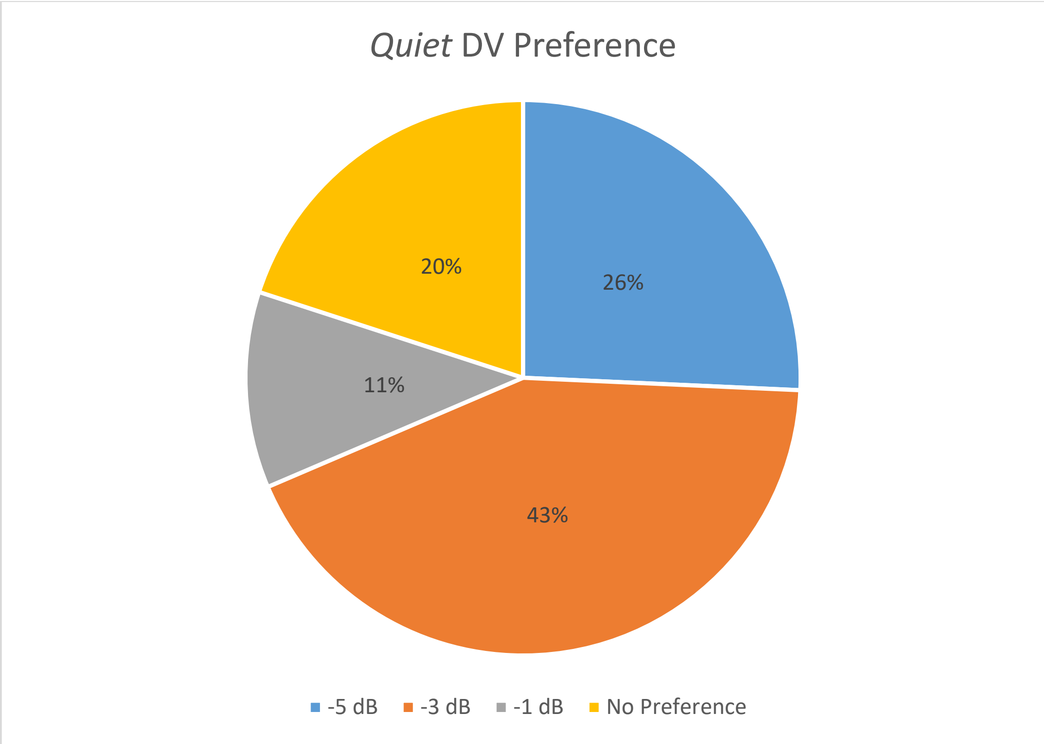


Figure 47: Distribution of preferred 'Quiet' DV condition for all listening test participants.

Further evidence for an overall preference can be seen in the histograms of Figure 48 and Figure 49 showing distributions about a peak at the 0.5 quantile value for the different subject groups.

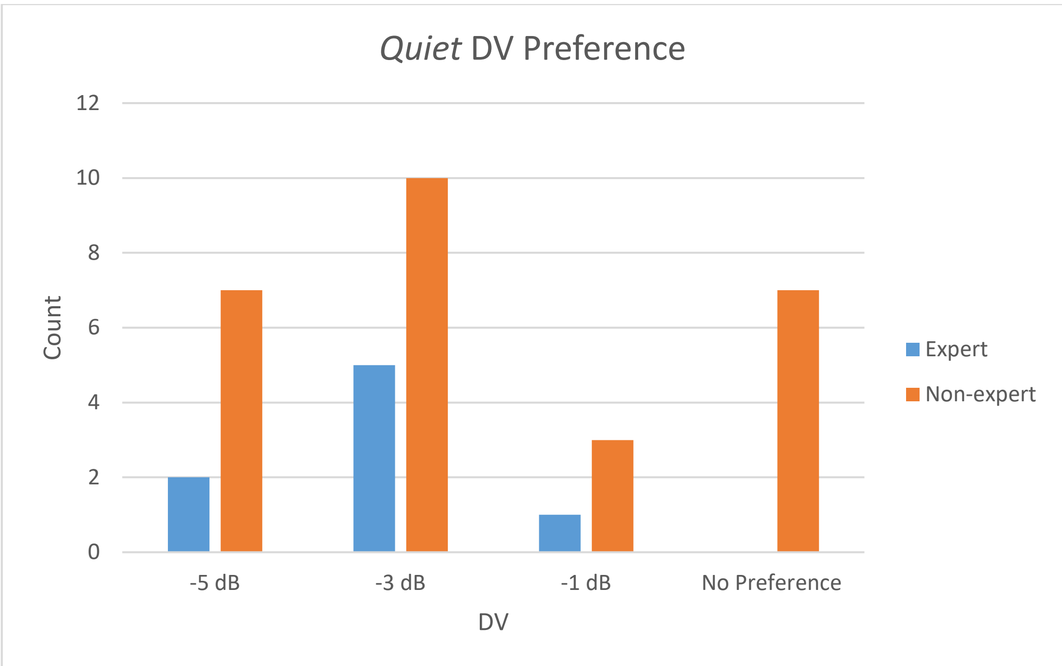


Figure 48: Distribution of preferred 'Quiet' DV condition among the 'Expert' and 'Non-expert' listener groups

.

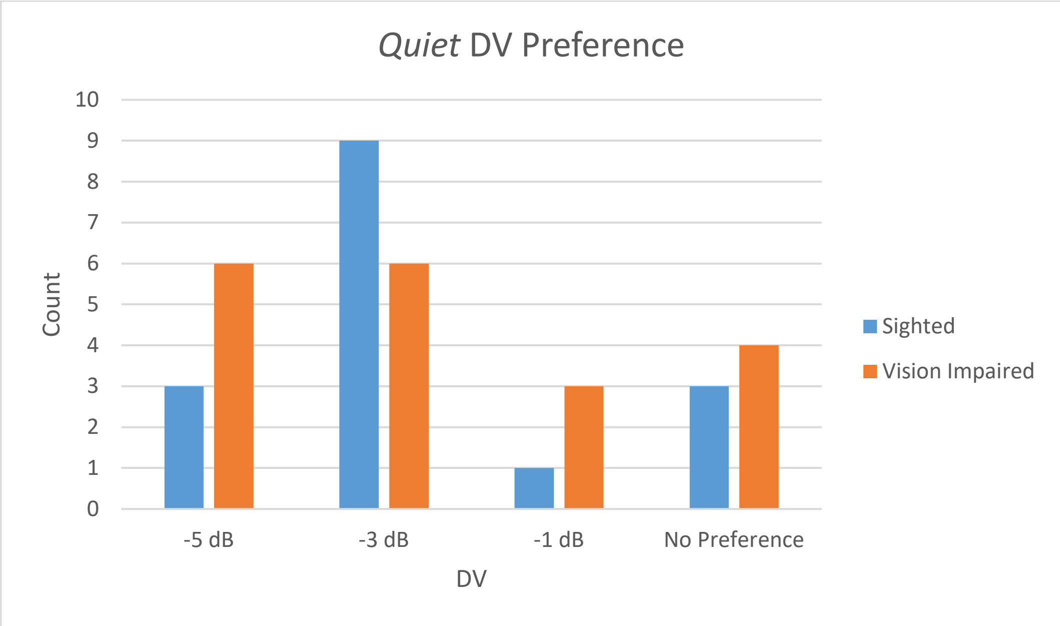


Figure 49: Distribution of preferred 'Quiet' DV condition among the 'Sighted' and 'Vision Impaired' subject groups.

Figure 50 below shows that confidence in preferred choice is high at 63% among participants, 7% of listeners had *No opinion* and only 1 participant claiming "*I couldn't hear any change in the programme volume*".

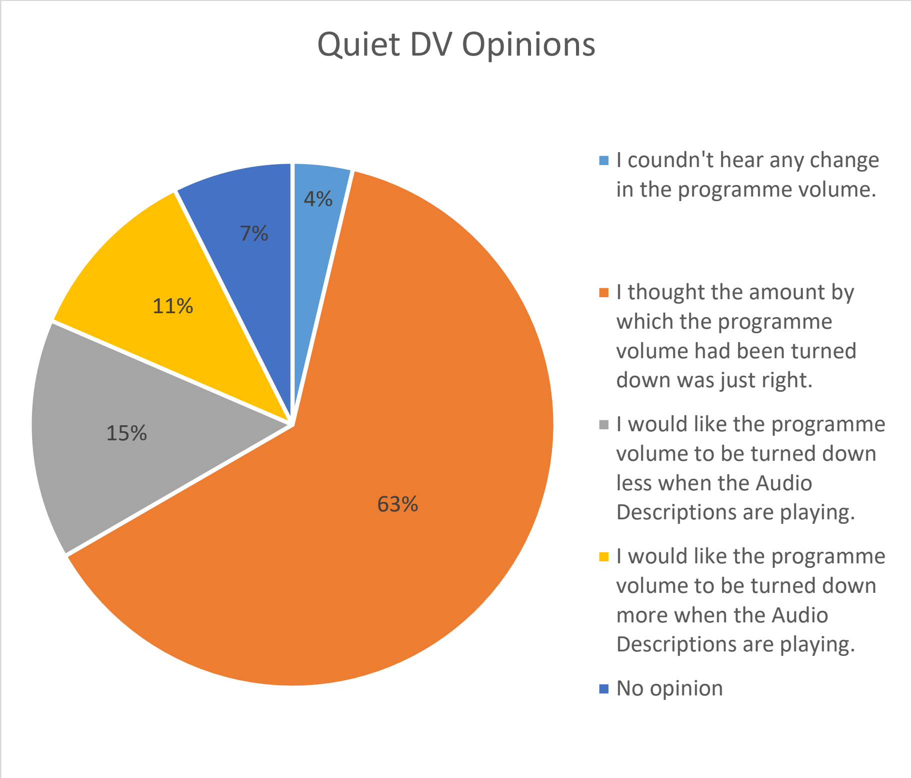


Figure 50: Opinions given by non-expert listener participants on their preferred DV choice for a 'Quiet' programme type.

Analysis

In order to better determine if the DV conditions with the highest frequency of preference are statistically significant, the data needs to be analysed in a way that identifies a pattern or trend. Statistical analysis can help us to test hypotheses that we have formed about the data such as if a relationship exists between the two variables of programme loudness and applied Dip Value.

For this experiment we can set out the following hypothesis:

- H_0: The distribution of preferred DV is even among the presented conditions and there is no DV preference among participants for a defined programme type.
- H_1: The distribution of preferred DV is not even among the presented conditions and there is a DV preference among participants for a defined programme type.

A chi-squared goodness of fit test was chosen as the statistical method deemed most appropriate to determine whether the observed data sample matches a theoretical distribution or the hypothesised distribution. This test was used to determine if a significant difference exists between the observed frequencies of the listening test outcomes and the expected frequencies based on the above null hypothesis. One of the results from this statistical test is the ρ (or significance) value. This value is then compared to the critical value α = 0.05 for a listening test. If ρ is less than α then we typically have good evidence to reject the null hypothesis.

Table 16 below shows the statistical results for the observed data generated by participants in the listening tests carried out for *Loud*, *Normal*, and *Quiet* programme types.

	Programme Type				
	Loud		Normal		Quiet
Non-experts					
Chi-squared statistic	14.62962963		3.37037037		3.666666667
P-value	0.002162136		0.337967876		0.299780589
Experts					
Chi-squared statistic	1.75		0.25		3.25
P-value	0.41686202		0.882496903		0.196911675
All					
Chi-squared statistic	17.45714286		5.571428571		7.4
P-value	0.000569091		0.134428272		0.060184324

Table 16: Results from the chi-squared goodness of fit test as applied to the observed data for the 'Non-expert', 'Expert', and 'All' subject groups in the 'Loud', 'Normal', and 'Quiet' listening tests. The ρ values highlighted in yellow show that a significant difference exists between the observed data and the expected statistical distribution and therefore the null hypothesis can be rejected for this programme type.

The cells in Table 16 highlighted in yellow show where the ρ value is less than the critical value α = 0.05 and that a statistically significant difference exists between the observed data and the expected theoretical distribution. We can therefore reject the null hypothesis for the *Loud* programme type and say that a DV preference does exist among participants. The preferred DV for *Loud* programme types at an AD cue is -19 dB.

Focusing on the data for all participants, the ρ value of 0.134428272 for the *Normal* programme type is significantly greater than the critical value α = 0.05 and therefore the statistical difference between the observed data and the theoretical distribution is such that the null hypothesis cannot be rejected. Looking again at Figure 43, it can be seen that the 0.1 and 0.5 quantile DV conditions have almost equal preference distributions (-31% and -37% respectively) and it could be suggested that an average be taken as the target DV. However, it is the preference of the *Non-expert* and the *Vision Impaired* listeners that needs to be considered and, in review of the distribution presented in the histograms of Figure 44 and Figure 45, it is suggested that the DV for *Normal* programme types at an AD cue be set to the median value of -9 dB.

In the case of Quiet programme types, the ρ value is 0.06 and therefore the null hypothesis cannot be rejected. The calculated ρ value is only 1% greater than the critical value α = 0.05 and, with reference to Figure 47 and the histograms of Figure 48 and Figure 49, the DV condition with the highest preference frequency (43%) is the 0.5 median quantile. It is therefore suggested that the DV for *Quiet* programme types at an AD cue point be set to the median value of -3 dB.

Conclusion

From the above analysis, it is understood that as the integrated loudness of the programme audio increases at the AD cue points, so too does the listener's preferred Dip Value. That is, when the background audio lies within the *Quiet* loudness band, the majority of listeners are satisfied with a 3 dB reduction in the programme audio in order for the AD narration to remain clear. The dip preference is less well-defined when the background programme audio is within the *Normal* loudness band, with almost equal preference shown for the median and lower quantile DVs of -9 and -11 dB. Since the satisfaction rating is high (63% *'Just Right'*) for both conditions, and 78% of listeners report the AD narration to be '*Very Clear*' in their preferred DV, it is expected that either DV would yield a satisfying listening experience but that the median DV of -9 dB be the target value as that setting has the highest preference rating among the target audience for AD services. When the integrated loudness of the programme lies within the *Loud* loudness band, there is a clear preference among listeners for the lower quantile DV condition of -19 dB. Future work should investigate further an audiences DV tolerance about the target value, the influencing factors, and whether the acceptable range is symmetric about the target. Only DVs have been studied here as this is the dominant parameter in AD content creation. The influence of the Fade Value (FV), or rate at which the DV is applied before and after the AD cue, on an audience's listening experience should also be investigated in the future.

Suggested Actions

The application of DVs during the creation of AD content is a technique required to facilitate speech intelligibility within Audio Description services. DVs should be applied in a way that clarifies the AD narration while also maintaining the integrity of the original programme's narrative. Although it is a technique used throughout broadcast and other audio-visual industries offering AD services, optimal values for programme loudness bands such as *Loud*, *Normal*, and *Quiet*, are not documented in the literature. This work fills that gap by presenting a set of recommended DVs correlating to the measured integrated loudness of the programme material at the AD cue point. This was done by documenting the current professional practices used in AD content creation, by benchmarking the mixing practice of 8 professional sound engineers when setting DVs, and by carrying out listening tests to determine an audience's preferred DV for each of the three quantified programme loudness bands of *Loud*, *Normal*, and *Quiet*.

Best practice in the production of AD content is to employ the services of experienced, professional sound engineers when setting DVs. However, this is not always an option and it has been observed that the responsibility of setting DVs often lies with non-sound engineer personnel, such as the narrator. Therefore, the following set of technical actions are proposed as effective guidelines when setting Dip Values during the creation of content for Audio Description services:

- Foremost, in order to maintain a common loudness reference, it is essential that the original programme material be R128-compliant (-23 LUFS ±0.5 LU).
- The undipped integrated LUFS value of the full AD narration track should be set in accordance with the original programme's LRA value. Table 17 offers guide LUFS values (see Appendix 4).

- The original programme material should be dipped -19 dB if its measured integrated loudness is -21 LUFS or higher across the duration of an AD cue.
- The programme material should be dipped -9 dB if its measured integrated loudness lies between -21 and -33 LUFS across the duration of an AD cue.
- The programme material should be dipped -3 dB if its measured integrated loudness lies between -33 and -48 LUFS across the duration of an AD cue.
- No dip should be applied to the original programme material if its measured integrated loudness is less than -48 LUFS across the duration of an AD cue.

In addition, and in response to the qualitative data presented in the Stage 3: Listening Tests, it is suggested that effort be made to further raise awareness of Audio Description services among both the vision impaired community and the wider public. A dedicated advertising campaign outlining where and how to access AD services is one proposed method. Moreover, it is suggested that information on the availability of Audio Description services within amenities such as cinema, theatre, TV, streaming services, museums, galleries, and live events be highlighted and made available through one, easily accessible resource.

From the evidence gathered in the Service-user Survey presented in the Background and Context section of this report, it is proposed that broadcasters make available AD services within their online players and catchup services. It is also suggested that a more robust QC stage be developed within the broadcasting and streaming communities to ensure that AD services do not suffer from the highlighted issues such as out-of-sync AD cues, missing AD cues, inconsistent AD cue levels, inconsistent and inappropriate Dip and Fade Values, and the mono delivery of stereo programme material. In addition, the inclusion of blind and vision impaired service users during the AD production and QC stages is one suggested action that could help mitigate the outlined issues associated with low quality AD.

It is proposed that further investigation be carried out into the impact that Fade Values (FV) have on the perceived quality of AD services and also the development of an algorithm for the automation of the measurement, parameter setting, and application of AD DVs.

References

[1] J. Packer, B. Gutierrez, and C. Kirchner, “ORIGINS, ORGANIZATIONS, AND ISSUES IN VIDEO DESCRIPTION: RESULTS FROM IN-DEPTH INTERVIEWS WITH MAJOR PLAYERS,” 1997.

[2] L. Fryer, *An Introduction to Audio Description*, Illustrate. Routledge, 2016. doi: 10.4324/9781315707228.

[3] “Code on Television Access Services Ofcom’s Code on Television Access Services Ofcom’s Code on Television Access Services Contents.” [Online]. Available: http://stakeholders.ofcom.org.uk/consultations/review-signing-

[4] N. Lodge, “ITC Guidance On Standards for Audio Description,” 2000.

[5] “Global Content Guide For Prime Video Licensors and Vendors v6.0,” 2018.

[6] N. Tanton, T. Ware, and M. Armstrong, “ACCESS SERVICES FOR DIGITAL TELEVISION : MATCHING THE MEANS TO THE REQUIREMENT FOR AUDIO DESCRIPTION AND SIGNING,” 2000.

[7] Sky, “Sky’s Technical Specification For Ultra High Definition Content.”

[8] X. Amatriain, “Netflix Recommendations,” 2009, [Online]. Available: https://partnerhelp.netflixstudios.com/hc/en-us/articles/215510667-Audio-Description-Style-Guide-v2-1

[9] M. Torcoli, A. Freke-Morin, J. Paulus, C. Simon, and B. Shirley, “Background ducking to produce esthetically pleasing audio for TV with clear speech,” *AES 146th International Convention*, vol. 67, no. 12, pp. 1003–1011, 2019, doi: 10.17743/jaes.2019.0052.

[10] EBU, “LOUDNESS NORMALISATION AND PERMITTED MAXIMUM LEVEL OF AUDIO SIGNALS Status: EBU Recommendation,” *Ebu - R 128*, no. June. pp. 1–5, 2014.

[11] I. Radiocommunication Bureau, “Algorithms to measure audio programme loudness and true-peak audio level BS Series Broadcasting service (sound),” 2015. [Online]. Available: http://www.itu.int/ITU-R/go/patents/en

[12] E. - Miles, “TECH 3342 LOUDNESS RANGE: A MEASURE TO SUPPLEMENT EBU R 128 LOUDNESS NORMALIZATION SUPPLEMENTARY INFORMATION FOR R 128,” 2016.

[13] International Telecommunication Union, “Proposal for the rationalization of nomenclature used in Recommendation ITU-R BS.1770 and ITU-R BS.1771,” 2010.

[14] Doug Liman (Director), United States. *The Bourne Identity [motion picture]*, (2002).

[15] Michael Gracey (Director), United States. *The Greatest Showman [Motion Picture]*, (2017).

[16] RTÉ (Producer), “Reelin’ In The Years [Television series],” RTÉ (Broadcaster), Ireland, 1999.

[17] RTÉ (Producer), “Fair City [TV Series 1989–present],” RTÉ (Broadcaster), Ireland, 2019.

[18] RTÉ (Producer), “What Makes My Day,” RTÉ (Broadcaster), Ireland, 2016.

[19] “Stellar - AD - Yella Umbrella Ltd.” http://www.yellaumbrella.tv/stellar-2-audio-description/ (accessed Oct. 20, 2022).

[20] B. Fazenda, “How to Design and Conduct Listening Tests for Audio and Acoustics,” *University of Salford Manchester*. 2015.

[21] E. - Miles, “TECH 3343 GUIDELINES FOR PRODUCTION OF PROGRAMMES IN ACCORDANCE WITH EBU R 128,” 2016.

[22] M. Armstrong, “From Clean Audio to Object Based Broadcasting,” 2016, Accessed: Apr. 11, 2023. [Online]. Available: http://www.bbc.co.uk/rd/pubs/whp

[23] “334-8565331284877938815-Best_practice_tips___Clear_Sound”.

[24] Avid, “Pro Tools Reference Guide,” 2022. [Online]. Available: www.avid.com/patents.

Appendices

Appendix 1

Audio Description Industry Survey

Overview

This study is being conducted by Dundalk Institute of Technology (www.dkit.ie) as part of an Audio Description (AD) research project. This is a questionnaire survey about current AD production practices in industry. It will help better understand the current roles and practices of those creating AD media content, and if a standardised approach to AD production is welcome. The survey is focused mainly on standards in setting AD Dip and Fade values and is aimed at those responsible for setting these values.

You have been asked to participate in this questionnaire because of your expertise around the creation of audio described content.

Please forward the survey to team members you feel best placed in this area of your AD production workflow. The questionnaire should take between 10 and 12 minutes to complete.

Research Area

Audio Description is a form of verbal commentary used to provide detail in relation to important visual plot and character information in a media production, such as a TV Programme, for the benefit of vision impaired consumers. AD describes context, body language, expressions and movements, helping to make the programme's visual content and narrative clear through sound.

During the production of AD, the narrator records descriptive passages in the gaps between the programme's dialogue. In order for the narration to be clear, the original programme material may necessitate a reduction, or Dip in volume so as to prevent masking of the audio descriptions. The Dip and Fade Values, or amount and rate by which the programme material is reduced, is dependent on the programme's volume at the AD cue point.

This research is funded by the Broadcasting Authority of Ireland.

Data Protection

This questionnaire is intended for research purposes. Your participation is voluntary and confidential, and your data will be anonymised for the purpose of data processing and data analysis. Data will be stored only for the duration of the research in a manner compliant with Data Protection legislation. Any concerns can be communicated to kieran.lynchk@iadt.ie

Location

1.In what country do you produce AD content?

Capacity

2.What role(s) best describe your AD production responsibilities?

Producer

Director

Scripter

Narrator
Sound Engineer

3.How much involvement do you have in the creation of AD content?
None
Little
Some
Lots
All aspects

Audio Technology
4.Please rate your knowledge of the following audio technologies.

	Bad	Poor	Fair	Good	Excellent
Microphone Techniques					
Recording Levels					
Compression					
Equalization					
Metering					
Loudness Standards					
Volume Automation					
The Decibel scale					
Audio Editing					
Mixing					

Dip and Fade Settings
5.Are you aware that the AD Dip and Fade values can be manually adjusted during the production process?
Yes
No

6.Do you set the Dip and Fade values during AD production?
Never
Sometimes
Always

7.What determines the AD Dip and Fade value?
Default settings
Recommendations and guidelines
The programme loudness at the Audio Description cue point
Your own judgment

8.What is your default Dip Value in dB? Please write NA if not applicable.

9.What recommended values or guidelines, if any, do you use for setting Dip Values? Please write NA if not applicable.

10.Do you adjust the Dip Value over the duration of the programme?
Never
Rarely
Sometimes
Often
Always

11.What is your default Fade Down and Fade Up Vales in milliseconds? Please write NA if not applicable.

12.What recommended values or guidelines, if any, do you use for setting Fade Values? Please write NA if not applicable.

13.Do you adjust the Fade Value over the duration of the programme?
Never
Rarely
Sometimes
Often
Always

Production Standards
14.Is there a Quality Control stage to your AD production workflow?
No
I don't know
Yes

15.Do you ever receive feedback from AD service end users or broadcasters?
Yes
No

16.Have you ever received complaints from the end user or broadcaster about AD service quality? If 'Yes', please use the 'other' text box to elaborate on the context of the complaint.
Yes
No

Standardisation
17.How useful would a set of recommendations on setting Dip and Fade values be to you?
Extremely useful
Somewhat useful
Not useful
No opinion

18.Would you like the setting of Dip and Fade values to be an automated process?
Yes
No
Maybe

Software
19.What software do you use to create AD content?
Audacity
Audition
One Dub
Pro Tools
Reaper
Starfish
Yella Unbrella

AD Users
20.Are you an AD service user?
Yes
No

Appendix 2

Instructions to Benchmarking Participants

Audio Description (AD) Dip Value Benchmarking.

Instructions to Sound Engineers:

Please read the Participant Information Leaflet.

Please Sign the Participant Consent Form.

Please read the following instructions on the recommended approach to setting AD Dip Values.

Please contact Kieran Lynch if you have any questions:

+353 86 8299692 kieran.lynch@dkit.ie

You are being asked to set a static audio descriptions level and appropriate volume dip values for programme material of various loudness ranges. Only mix those clips that are active and marked in red.

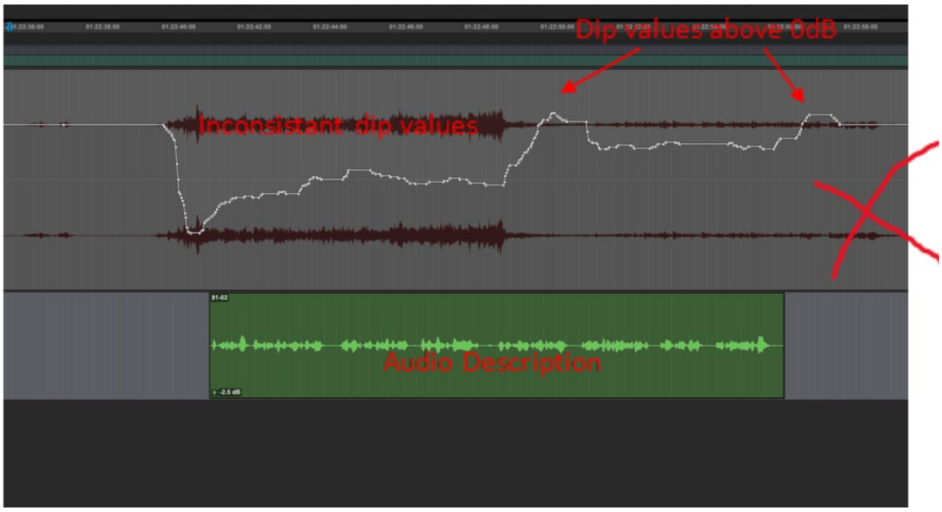
When you open a Pro Tools session, use the 'Save as' function and append the session title with your name.

All the supplied programme material are R128 compliant. Please do not apply any overall gain to the programme material and **do not adjust the clip gain already applied**. Do not apply processing such as compression, EQ, trim, or clip gain to the programme material. The only adjustments to the programme material are to be made using volume automation moves.

1. Set the programme material fader to 0dB.
2. Set the overall AD playback volume so that the descriptions are clear and well defined at the point where the programme material is quietest i.e., a point in the programme where no dip (volume automation trim) is required. The AD playback volume should then remain static and should not be adjusted for individual AD cue points.
3. As the programme material plays, use your discretion to determine if a dip needs to be applied to the programme material at each AD cue point.
4. Apply dips using volume automation trims so that the audio descriptions are clear, present, and well defined but dips should not be distracting or detract from the programme's narrative.
5. Apply the appropriate dip values to the programme material at the AD cue points using volume automation only.

6. Apply only negative trim values (dips) to the programme material's volume automation lane.
7. Volume automation trim values should be consistent in their values when applied and should always dip from 0dB and return to 0dB.

Above is an example of best practice; Dips are applied with consistant values, dipping from 0dB and returning back to 0db using appropriate fade down and fade up values.



Above are two examples of inconsistant Dip Values applied to programme material; Avoid complex fader moves, positive Dip values (above 0dB), and make sure to return the programme volume to the 0dB point between AD cues.

The Pro Tools automation 'Preview' and 'Manual Write' functions may help in applying consistent dip value to the programme material at the AD cue points.

8. Automation dip fade ins and fade outs should be applied so that each dip feels natural and smooth.

The order of preference for mixing is:

1. The Bourne Identity
2. The Greatest Showman
3. Reeling In the Years 1979
4. Fair City (Male AD)
5. What Makes My Day
6. Fair City (Female AD)

When finished, please email me a zipped folder containing your Pro Tools sessions only - no media – and your signed consent form.

Appendix 3

The RTE AD Production Workstation Setup

Microphone: Shure SM7 B

Pre-amplifier: Cloud Microphones Cloudlifter CL-1 Mic Activator

Audio Channel: Grace m103

Audio Interface: SSL 2+

Computer: Windows 10

Software: Yella Umbrella Stellar V5.3

Appendix 4

AD Narration LUFS Look-up Table

Programme LRA (LU)	1	2	3	4	5
AD Narration LUFS	-23.0	-23.2	-23.4	-23.6	-23.8
Programme LRA (LU)	6	7	8	9	10
AD Narration LUFS	-24.0	-24.2	-24.4	-24.6	-24.8
Programme LRA (LU)	11	12	13	14	15
AD Narration LUFS	-25.0	-25.2	-25.4	-25.6	-25.7
Programme LRA (LU)	16	17	18	19	20
AD Narration LUFS	-25.9	-26.1	-26.3	-26.5	-26.7
Programme LRA (LU)	21	22	23	24	25
AD Narration LUFS	-26.9	-27.1	-27.3	-27.5	-27.7
Programme LRA (LU)	26	27	28	29	30
AD Narration LUFS	-27.9	-28.1	-28.3	-28.5	-28.7

Table 17: Look-up table for setting an appropriate AD narration LUFS value for a programme with a particular LRA value.

Appendix 5

Listening test setup:

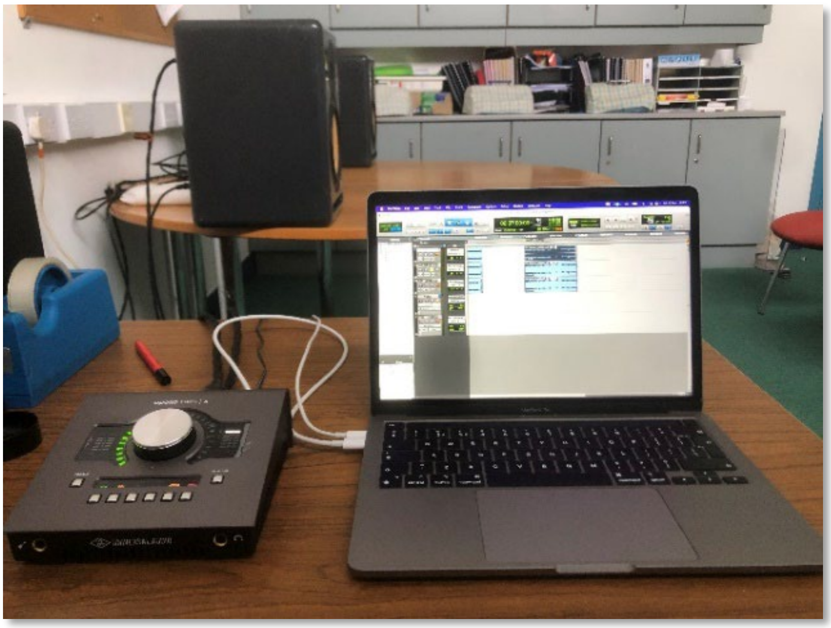


Image 5: Playback system - Pro Tools, MacBook Pro, UA Apollo Twin audio interface.

Image 6: Playback speakers - KRK V8 Series 2.

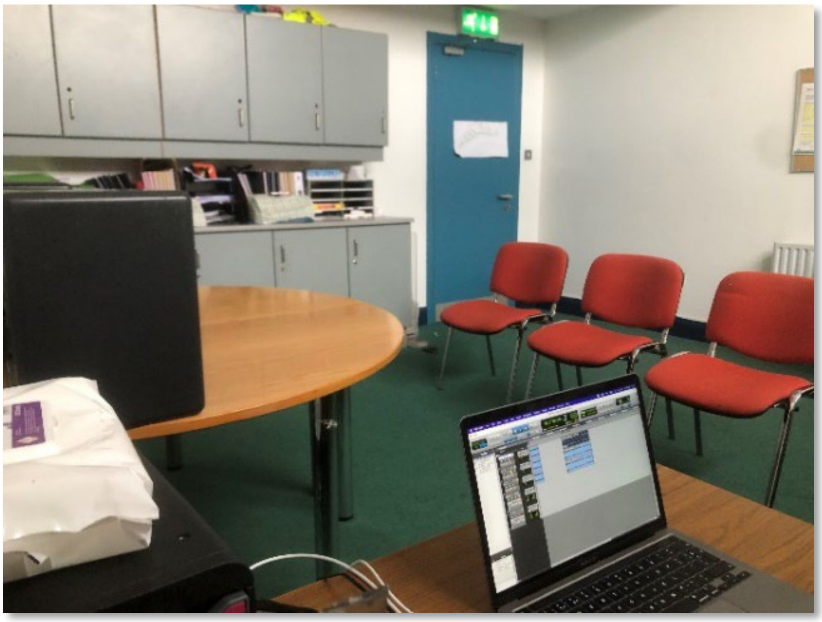


Image 7: Listening test room setup.

Image 8: Listening test setup - listener proximity to the speakers.

Image 9: Listening test setup - listener perspective.

Appendix 6

Listening test questionnaire:

Audio Descriptions Listening Test Questionnaire

Thank you for agreeing to participate in this research. These listening tests are the final stage of a research project that aims to develop a set of standards and recommendations for setting Audio Description dip values.

Please listen to the relationship between the programme material and the audio descriptions and complete the questionnaire to the best of your ability.

Please note that the questionnaire aims to collect opinions and that there are no wrong or right answers.

Required

1.Please enter your assigned ID below.

2.Which of the following best describes your vision?

- Blind
- Vision Impaired
- Partially Sighted
- Sighted

3.Do you suffer from any known hearing impairment?

- Yes
- No

4.Which age group are you in?

- 18 - 24
- 25 - 34
- 35 - 44

- 45 - 54
- 55 - 64
- 65 or older

5.How often do you use Audio Description services?

- Daily
- Weekly
- Monthly
- Occasionally
- Rarely
- Never

6.When this research project is complete, would you like to receive a copy of the final report?

- Yes please
- No thank you

Listening Test 1 - Loud Programme Material

Please wait for the researcher to play the first set of audio examples before proceeding.

7.Which of the three audio examples did you prefer.

- Example 1
- Example 2
- Example 3
- No preference

8.In your preferred choice, did you feel that the Audio Descriptions were:

- Very clear
- Clear
- Unclear
- Very unclear

- No opinion

9.Which of the following statements is true for your preferred choice?

- I would like the programme volume to be turned down more when the Audio Descriptions are playing.
- I would like the programme volume to be turned down less when the Audio Descriptions are playing.
- I thought the amount by which the programme volume had been turned down was just right.
- I couldn't hear any change in the programme volume.
- No opinion

10.Which of the following statements is true for your preferred choice?

- The Audio Descriptions enhanced my listening experience.
- The Audio Descriptions were distracting.
- No opinion

Listening Test 2 - Normal Programme Material
Please wait for the researcher to play the second set of audio examples before proceeding.

11.Which of the three audio examples did you prefer.

- Example 1
- Example 2
- Example 3
- No preference

12.In your preferred choice, did you feel that the Audio Descriptions were:

- Very clear
- Clear
- Unclear

- Very unclear
- No opinion

13.Which of the following statements is true for your preferred choice?

- I would like the programme volume to be turned down more when the Audio Descriptions are playing.
- I would like the programme volume to be turned down less when the Audio Descriptions are playing.
- I thought the amount by which the programme volume had been turned down was just right.
- I couldn't hear any change in the programme volume.
- No opinion

14.Which of the following statements is true for your preferred choice?

- The Audio Descriptions enhanced my listening experience.
- The Audio Descriptions were distracting.
- No opinion

Listening Test 3 - Quiet Programme Material

Please wait for the researcher to play the third set of audio examples before proceeding.

15.Which of the three audio examples did you prefer.

- Example 1
- Example 2
- Example 3
- No preference

16.In your preferred choice, did you feel that the Audio Descriptions were:

- Very clear
- Clear

- Unclear
- Very unclear
- No opinion

17.Which of the following statements is true for your preferred choice?

- I would like the programme volume to be turned down more when the Audio Descriptions are playing.
- I would like the programme volume to be turned down less when the Audio Descriptions are playing.
- I thought the amount by which the programme volume had been turned down was just right.
- I couldn't hear any change in the programme volume.
- No opinion

18.Which of the following statements is true for your preferred choice?

- The Audio Descriptions enhanced my listening experience.
- The Audio Descriptions were distracting.
- No opinion

Appendix 7

Audio Description Service User Survey

Overview

This study is being conducted by Dundalk Institute of Technology (www.dkit.ie) as part of an Audio Description (AD) research project. This is a questionnaire survey designed to gather information in relation to AD services currently available in broadcast and on streaming platforms. It will help better understand the current standards in AD services, and if a standardised approach to AD production is welcome. The survey is focused mainly on the AD service end user experience.

You have been asked to participate in this questionnaire because of your expertise in the access and use of AD services in broadcast and on streaming platforms.

Please forward the survey to anyone you feel is qualified to participate. The questionnaire should take between 10 and 12 minutes to complete.

Research Area

Audio Description is a form of verbal commentary used to provide detail in relation to important visual plot and character information in a media production, such as a TV Programme, for the benefit of vision impaired consumers. AD describes context, body language, expressions and movements, helping to make the programme's visual content and narrative clear through sound.

During the production of AD, the narrator records descriptive passages in the gaps between the programme's dialogue. In order for the narration to be clear, the original programme material may necessitate a reduction, or Dip in volume so as to prevent masking of the audio descriptions. The Dip and Fade Values, or amount and rate by which the programme material is reduced, is dependent on the programme's volume at the AD cue point.

This research is funded by the Broadcasting Authority of Ireland.

Data Protection

This questionnaire is intended for research purposes. Your participation is voluntary and confidential, and your data will be anonymised for the purpose of data processing and data analysis. Data will be stored only for the duration of the research in a manner compliant with Data Protection legislation. Any concerns can be communicated to kieran.lynch@dkit.ie

1.In what country do you live?

Enter your answer

2.What is your age range?

18 to 24

25 to 34

35 to 44

45 to 54

55 to 64

65 or over

3.Are you vision impaired?
Yes
No

4.How would you rate the quality of your hearing?

Excellent

Very good

Good

Average

Fair

Poor

Very poor

5.Are you an AD services user?
Yes
No

6.Which AD services do you use?

TV services
Streaming services

Other online services

Live events

Cinema

7.What AD service providers do you use?

RTE

Virgin Media

TG4

BBC

ITV

Channel 4

Netflix

Apple TV

Amazon Prime

Other

8.If you answered Other to the previous question, please use the space below to list the other AD services providers you use.

Enter your answer

9.How happy are you with the AD services you use?

Very happy

Somewhat happy

Neither happy nor unhappy

Somewhat unhappy

Very unhappy

10.Have you notices any inconsistencies between the AD services you use?
Yes
No

11.Please give details of inconsistencies observed between the AD services you use?

Enter your answer

12.How would you rate the quality of AD services that you use?

Excellent

Very good

Good

Fair

Poor

13.What aspects of AD services have you experienced to be of a low quality?

Quality of the AD recordings

Quality of the descriptions

Quality of the Narrator's delivery

The volume of the AD relative to the programme dialogue

The volume of the programme during AD passages

Other

14.If you answered Other to the previous question, please use the space below to tell us about the other aspect of AD services you have experienced to be of a low standard.

Enter your answer

15.Is the programme material ever too quiet during the descriptive passages?

Never

Sometimes

Often

Always

16.Is the programme material ever too loud during the descriptive passages?

Never

Sometimes

Often

Always

17.Are the descriptive passages ever too quiet in comparison to the programme material?

Never

Sometimes

Often

Always

18.Are the descriptive passages ever too loud in comparison to the programme material?

Never

Sometimes

Often

Always

19.Have you ever made a complaint to a service provider in relation to the quality of their AD?

Yes
No

20.If you answered Yes to the previous question, please let us know about the nature of your complaint?

Enter your answer