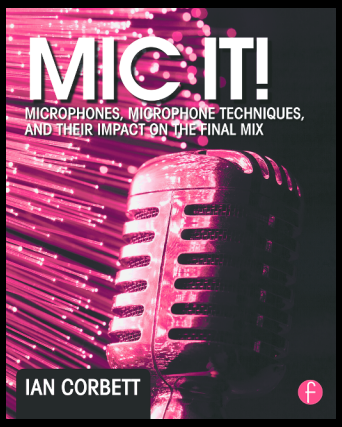
A **Routledge/Focal** FREEBOOK

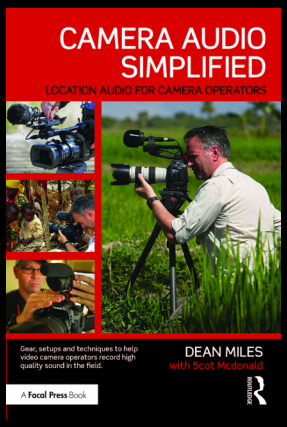
Sound Production Today

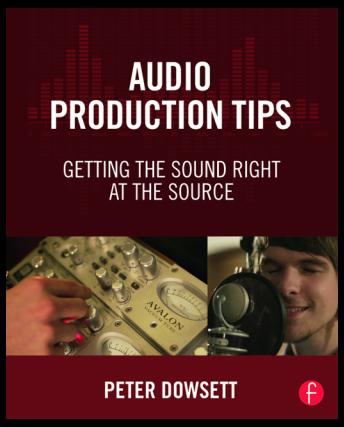
Focal Press
Taylor & Francis Group

TABLE OF CONTENTS

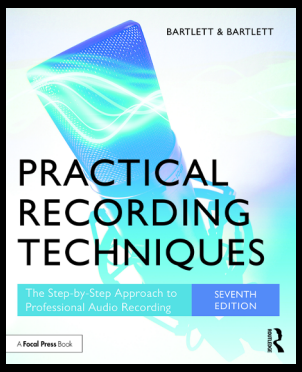
Learn more about Audio Productions with the full books excerpted in this FreeBook:

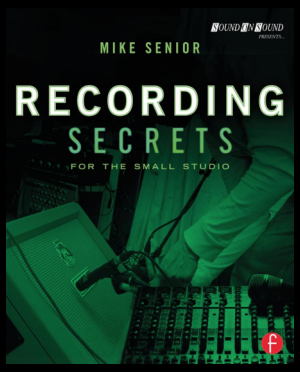


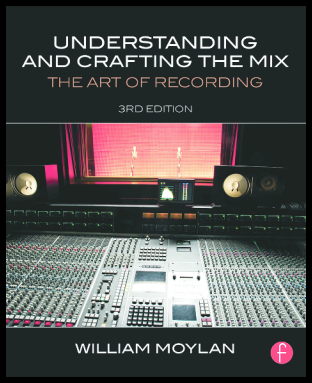




Related titles:







Visit Routledge.com/audio to browse our full range of Audio titles

Introduction

This FreeBook is a one-stop shop for great sound engineering tips from some of today's leading audio experts. Whether you are recording music or shooting live audio, you are bound to find some useful information within these pages.

CHAPTER ONE of this FreeBook comes from **Ian Corbett**'s book ***Mic It!*** The sample is from chapter three, "Good sound" where Corbett teaches reader what "good sound" is and how to achieve it in a home recording studio.

> **About *Mic It!:*** Capture great sound in the first place, and spend less time "fixing it in the mix" with **Ian Corbett's *Mic It! Microphones, Microphone Techniques, and Their Impact on the Final Mix*.** With his expert guidance, you'll quickly understand essential audio concepts as they relate to microphones and mic techniques.
>
> **Dr. Ian Corbett** is the coordinator of the Audio Engineering Program, and Professor of Music Technology and Audio Engineering at Kansas City Kansas Community College. He also owns and operates "off-beat-open-hats – recording and sound reinforcement" which specializes in servicing the needs of jazz and classical ensembles in the Kansas City area.

CHAPTER TWO features a chapter on Sound Elements from ***Camera Audio Simplified***. If you are making a film using a video camera and need a little help with the audio, this section is for you.

> **About *Camera Audio Simplified*:** You've spent a lot of time learning how to use the features of your video camera to take amazing pictures, but chances are the audio always gives you trouble! To achieve great sound to go along with great video, you need to go beyond "automatic" mode and take control – know as much about microphones as you do about lenses with the help of ***Camera Audio Simplified*.**
>
> Author **Dean Miles** gives you the skills you need to capture quality location audio with your camera mic, wireless system, lavaliers, and handheld microphones. You'll get step-by-step guidance to help elevate your sound from amateur to professional – and build your career and video production business.

CHAPTER THREE, Demystifying Recording Levels from ***Audio Production Tips,*** breaks down the science of sound with easy-to understand examples.

Audio Production Tips: Getting the Sound Right at the Source provides practical and accessible information detailing the production processes for recording today's bands. By demonstrating how to "get the sound right at the source," author Peter Dowsett lays the appropriate framework to discuss the technical requirements of optimizing the sound of a source.

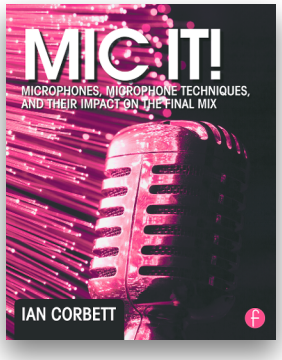
Peter Dowsett is a British audio engineer known for his experience with many facets of the music industry including studio engineering, mixing, mastering and live sound. In the studio he has worked at Metropolis studios with clients including Pharrell Williams, Snoop Dogg, Rick Ross, *Downton Abbey*, Nick Jonas, and Dappy. His work has been synced on ABC, Channel 5 and featured on cover CDs *for Metal Hammer* magazine.

Happy reading and recording!

CHAPTER 1

Good Sound

Chapter 1: Good Sound



The following is excerpted from Mic It!: Microphones, Microphone Techniques, and Their Impact by Ian Corbett.

Learn more:

RECOGNIZING GOOD SOUND:

Before even thinking of plugging in and setting up a microphone, it is important to understand what the desirable characteristics (and undesirable characteristics) in sound are. That's why there are some chapters in this book like this one – that don't directly discuss microphones, but discuss things you need to know in order to use microphones to capture the best, most suitable sound possible.

There's no simple answer to the question "what is good sound?" The best answer might be along the lines of "whatever is stylistically and artistically appropriate." Good sound is subjective. One person's ideal guitar sound may be another person's worst nightmare – however that is often related to whether a sound is appropriate to the context it's in, rather than the sound being simply "good" or "bad." But bad sound certainly does exist! Poor quality sound sources, poor quality equipment, bad recording techniques, and poor mixing skills can all result in inappropriate, questionable, or just plain "wrong" sound!

How do you learn to record and mix well? There are basic concepts and skills that should be mastered before developing your own style. Musicians develop their skill sets and individual musical style by listening to other musicians, emulating them, and eventually synthesizing many influences into their own unique characteristics. As a sound engineer or producer, you should similarly find good quality recordings, listen to them, analyze them, and try to emulate them – building up your skills and techniques before eventually developing your own style.

Listening to, and becoming intimately familiar with, a wide variety of musical and production styles will make you more marketable in the industry – if you only listen to hip-hop, good luck when an acoustic folk band shows up for a show or session you're working! Hopefully, like many audio professionals, you can combine your love of great sounding recordings with artists you enjoy musically. Don't just listen to music because you like the style or the artists – when starting to listen critically, it's often easier to concentrate on the sound, and not be distracted by the music, when you're listening to artists and styles you're not a fan of. It's very easy for your love of an artist or band to persuade you that the recording is better than it actually is!

What do you need to be aware of, how should you be listening, and what should you be listening for, in order to identify the desirable characteristics of a recording?

SOUND REPRODUCTION FORMATS

MONO

A mono, or monophonic playback system has only one loudspeaker, as in Figure 2.1.

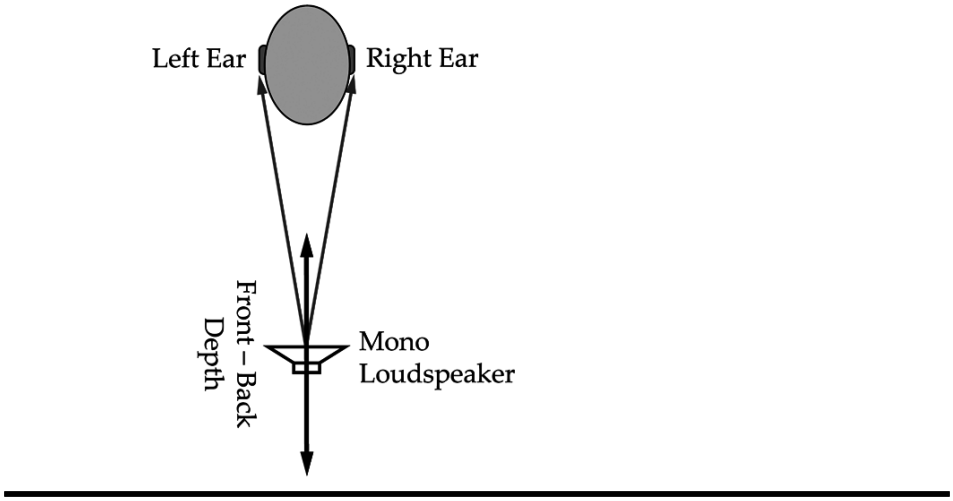


Figure 2.1 *A mono reproduction system with a single loudspeaker. Sound from the loudspeaker arrives at both ears at the same time (if the loudspeaker is centered). The mono format only allows sounds to be positioned along the loudspeaker's limited front/back axis. The result is a spatially congested image, making mix clarity a challenge.*

Mono playback systems include some TVs, bedside alarm clock radios, ceiling type loudspeaker systems in retail outlets (which have many distributed loudspeakers, that are all sent the same signal), and speakers on many tablets and mobile phones. Mono systems are one-dimensional. They can offer a sense of front/back depth through the use of creative recording and mixing techniques, but there is no sense of wide space, and all the sounds come from the same location. It is potentially difficult for a listener to clearly hear everything that might be going on in a mono mix, which makes achieving a good mono mix very challenging.

A small nasty sounding mono Auratone type loudspeaker is a common feature in many professional recording studios – enabling engineers to anticipate the effects of poor quality, mono sound systems, and make sure their product translates acceptably to them.

STEREO

Stereo, or stereophonic systems feature two playback channels – left and right. Different signals are sent to each, as shown in Figure 2.2. In a stereo loudspeaker

receive the same wavefront at the same time, followed by identical slightly delayed inter-aural crosstalk. The listener perceives this sound as coming from a central location, directly between the loudspeakers – where there is no physical loudspeaker. The illusion of a sound located where there is no loudspeaker is called a phantom image.

The stereo format is a two-dimensional format when creative recording and mixing techniques are used: • There is clearly a sense of left/right directionality, because it is possible to position sounds anywhere between the loudspeakers, and to create the illusion that sound is coming from just beyond the physical loudspeakers.

• It is also possible to create the illusion of front/back depth.

• To a lesser degree, it is possible to create the limited illusion of a third dimension – height.

CHECK YOUR MONO COMPATIBILITY!

Checking your stereo project sounds good in mono – its mono compatibility – is very important. You never know what type of system your project will end up being heard on. Cheap TVs, cell phone speakers, elevator music systems, bedside clock radios, and most ceiling speaker systems in shops and stores are all mono systems. The last thing you want is for your mix to sound bad, or for sounds or certain frequency ranges to partially, or in extreme cases, completely disappear when a stereo mix is summed to mono.

Most mixing consoles and DAWs have a mono button in their monitor section. This sums the stereo mix to mono to check this compatibility. The feature is there for a reason – use it! Significant tone, timbre, or frequency balance (equalization) changes, when a mix is folded down to mono, usually indicate a mono compatibility problem that needs fixing.

SURROUND SOUND

Multichannel surround sound formats, such as 5.1 and 7.1, utilize additional loudspeakers which surround the listener. They offer the advantage of being able to envelop the listener with sound coming from behind and to the side of the listening position. Additionally, a dedicated center channel loudspeaker can take the place of center phantom images, and provide alternate sonic imaging qualities. New formats with height channels are being explored and developed – these truly result in a three-dimensional listening experience. Surround sound formats are beyond the scope of this book, and stereo production should be mastered before attempting any type of surround sound production.

MONITORING OPTIONS – LOUDSPEAKERS, HEADPHONES, AND EARBUDS

Stereo is the most common consumer listening format, and a pair of studio loudspeakers, called monitors, are the preferred listening system when recording, mixing, and mastering in stereo. Boom boxes, hi-fi systems, stereo televisions, car stereos, and the plethora of portable media players (PMPs) such as MP3 players, iPods, and multimedia mobile phones all offer two-channel "stereo" playback. But are all of these devices actually stereo?

Figure 2.2 showed that a stereo system creates inter-aural crosstalk that our hearing system uses to determine the directionality of sound. Figure 2.3 shows a pair of headphones (or earbuds) on a listener. There is no inter-aural crosstalk in this system – the sound from the left driver goes only into the left ear, and the sound from the right

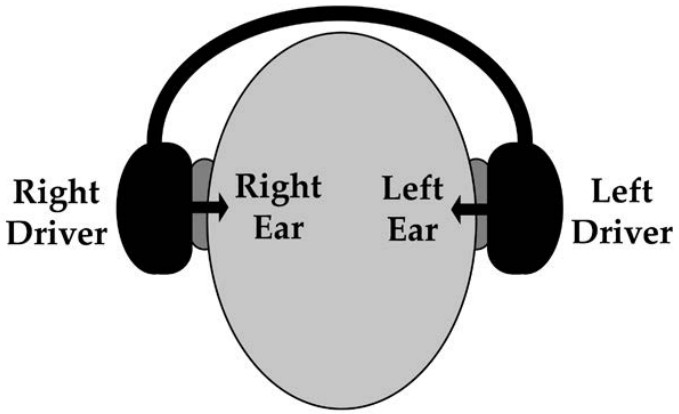


Figure 2.3 *A binaural headphone or earbud system lacks the inter-aural crosstalk necessary to make it a true stereo system.*

driver goes only into the right ear. The only directional information presented to and processed by our brain is the amplitude difference of the sound between each ear. For this reason, headphones and earbuds should not really be called stereo, despite the labeling to the contrary on the retail packaging! They are in fact binaural – featuring two channels, minus the inter-aural crosstalk necessary to make them a true stereo system.

There are several reasons that mixing using high quality, professional loudspeakers is preferable to mixing using headphones:

- Mixes created on loudspeakers translate very well to headphones. Mixes created on headphones do not translate as well to loudspeaker systems.
- "But most of the listeners will be using earbuds anyway." It is perfectly OK, and

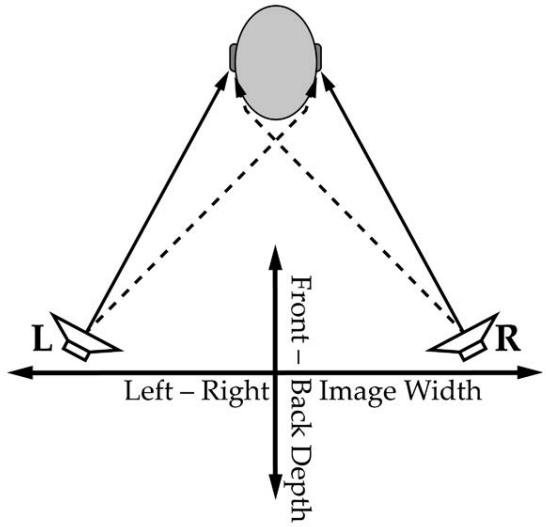


Figure 2.2 *A stereo reproduction system uses two channels, each sending its own unique signal.*

system, sound from the left loudspeaker travels to the left ear, and the right ear – where it arrives slightly later, changed in timbre due to the extra distance, and filtering/equalization effects of wrapping around the face and head. These delay and EQ effects are known as HRTFs, or head related transfer functions. Similarly, sound from the right loudspeaker travels to the right ear, and also the left ear – again, slightly delayed and changed in timbre. The sound transfer to each "opposite" ear is known as inter-aural crosstalk, and is an essential and desirable component of a stereo playback or monitoring system.

The inter-aural crosstalk between the speakers and each opposite ear (the dashed lines) is an essential component of the format. The path from the left loudspeaker to the right ear is slightly longer than to the left ear, therefore the left loudspeaker's sound arrives slightly later at the right ear, and its tonality is changed as it wraps around the face and head.

The reason a sound coming from just the left loudspeaker sounds like it's coming from just the left loudspeaker (even though the sound travels to both of our ears) is because our brain uses the HRTF sub-millisecond-level time-delay and the EQ differences of the wavefront's arrival at each ear to determine directionality. We perceive the sound as coming from the direction of the ear at which the wavefront first arrived. This is known as the Law of First Wavefront, or Haas Effect, or Precedence Effect.

If both loudspeakers reproduce an identical sound wave simultaneously, both ears

good, to check a mix on headphones or earbuds. But to create a stereo mix that will best translate to the greatest variety of playback systems, the mix should be crafted on a stereo loudspeaker system. A mix that doesn't translate well to many reproduction systems, or future systems, will have a more limited audience, and a more limited lifespan.

• Headphones do not present the same stereo image or front/back depth that stereo loudspeaker systems are capable of. While a headphone image may be correctly described as more precise and surgical, it is also smaller and more compact. If you are mixing for headphones exclusively, then knowledge of some of the miking techniques discussed later in this book can allow you to incorporate delay-based image steering into your recordings, to add some delay and HRTF-like artifacts into the binaural reproduction system.

• Professional monitor loudspeakers are built to tell the truth, and not to hype specific frequency ranges, or to sound instantly "pleasing" (like most lower cost headphones, and consumer loudspeakers). Professional monitors reveal more problems in your mix or recording than most headphones or consumer loudspeakers. This makes it more challenging to get the sound right on good loudspeakers – but when you do, the results should translate well to a larger variety of playback systems. Most headphones and consumer loudspeakers don't reveal so many details or potential problems – which can lead to unfortunate surprises when a mix is played on a better sound system.

• Many headphones and consumer loudspeakers hype the bass for "instant impact." Professional monitors have a flatter, more linear frequency response in order to more correctly present the sound, and not mask any potential mix issues. This means that a first impression of a professional monitor speaker might be that it lacks bass, and is not as full sounding. If you are new to professional monitors, you need to learn how your monitors sound – intimately. Listen extensively to industry-respected, good sounding reference recordings on your new monitors. Recognize the characteristics of good sound on those monitors, and get used to the balance of the low, mid, and high frequencies. If you do not know and trust how your speakers sound, you cannot reliably use them to craft a mix that will sound good elsewhere.

There are some great sounding (and usually expensive) headphones on the market, and these can be used for monitoring, subject to the caveats above. Earbuds however, particularly the cheap stock items that the majority of the public happily use, sound horrible. They either lack real bass, have a highly inaccurate "bass" boost that isn't in a really low frequency range, and their high frequency performance is usually not very good.

EARBUDS – A HEALTH HAZARD?

Earbuds connected to portable media players can produce peak SPLs in the ear canal of up to 127 dBA, depending on the device! The average sustained level is a little below the peak figure, but still well in excess of recommended 85 dBA maximum sound exposure level without hearing protection.

Stock earbuds, and most cheap earbuds, are non-sealing "open" designs that let in extraneous sound so the listener is aware of their surroundings. In noisy environments listeners usually listen at least 6 dB louder than the environmental noise. City traffic is about 85 dB SPL, the noise on a subway train can be over 90 dB SPL, and restaurants and bars can be over 90 dB SPL!

Listening to music at least 6 dB louder than this background noise immediately puts the listener at risk of permanent hearing damage in a short time. Sealing, or "closed" earbuds should promote listening at quieter levels because they block some of the background noise, but many people still end up turning them up excessively loud.

KNOW HOW YOUR MONITOR SPEAKERS SOUND

• If cheap headphones or earbuds are what you are used to listening to, then you will need to retrain your ears by listening extensively on good stereo monitors.

• If you go straight from earbuds to good studio monitors, without learning how the monitors sound, you'll end up recreating the earbud sound on the monitors – and the result will be exponentially worse when your mix is played back on earbuds!

• If you're transitioning to new monitors, it's a good idea to have some respected good sounding commercial mixes spinning continuously, so you can A/B compare the characteristics of your mix to those commercial mixes as you are working.

COMPRESSED AUDIO FORMATS

Dissemination formats such as MP3, AAC, and Ogg-Vorbis, are *lossy* compression codecs. They reduce (or compress) the audio file size (compared to PCM) by removing audio that the encoder doesn't think will be perceived. As a result, songs can be downloaded faster and more songs fit onto a portable media player (PMP). The process is not transparent though, and more and more detail is removed as the bit-rate is reduced. Lossy compression does negatively affect the quality of the sound.

LOSSY COMPRESSION ARTIFACTS

Some of the many undesirable byproducts of lossy audio compression codecs include:

- Frequency content discarded throughout the spectrum.
- Drastic high frequency loss at low bit-rates.
- MP3's reduction of bass frequencies.
- Time and phase smearing.
- Roughness.
- Ringing frequencies.
- Swirling and unstable higher frequencies.
- Loss of transient detail.
- Flattening of dynamics.
- Loss of reverberation tails, and reduction of other time-based effects.
- Stereo image narrowing and blurring.
- The addition of low-level noise.

Lossy compressed audio formats are not formats an audio professional, or aspiring audio professional, should be listening to when learning what good sound might be, and how to achieve it.

Lossless compression codecs, such as FLAC and Apple Lossless (ALAC), do not suffer from the negative artifacts of lossy codecs, however, they do not offer such large file size reductions.

While iPods and most other modern PMPs can store lossless files of one type or another, or even uncompressed PCM files, the device's cheap convertors and amplifiers, and cheap earbuds mean they're still not an acceptable solution. So, when listening, try to use good quality playback systems, and/or better external conversion of the digital bit-stream from a PMP.

Lossy formats on a PMP or computer, and/or the cheap convertors and amps found on PMPs should not be used for any system evaluation, or to analyze mixes – they do not accurately portray the mix as the engineer intended.

DYNAMIC RANGE

If you compare a pop or rock recording from the 1970s to one from the late 2000s one thing will probably be really obvious – the latter one is louder! The compact disc becoming the most popular dissemination format in the late 1980s, and the availability

of the digital limiter in the mid-1990s, prompted competition to have the "loudest" most in-your-face mix – the *Loudness War*. By the late 2000s, the average levels on CDs had become so hot that distortion was apparent on many releases – Metallica's *Death Magnetic* notably attracting negative media attention in 2008.

Some dynamic range compression and limiting are a desirable and essential part of the mastering process – particularly of commercial music. This dynamic control gives recordings extra punch, and can help them play back well on a wider variety of reproduction systems. However, hyper-compression – when the peak levels are reduced too much, and end up very close to the more continuous average levels – results in a squashed, gritty, and often distorted sound, that lacks true punch and detail. The transients, the initial attack portions of loud punchy sounds such as the kick and snare drums, are particularly prone to this type of sonic damage. Hyper-compressed mixes are a product of extreme mastering, recording, and mixing techniques. That level of loudness simply cannot be created using more traditional tools and techniques.

Have you ever wondered why your favorite record sounds so different on the radio than on CD? One of the reasons for the loudness war was competition to have the loudest song on the radio. We are easily convinced that something that is louder sounds better. Industry wisdom was that when scanning similar radio stations, a listener was most likely to settle on the one that was louder. Most radio stations employ equalizers, multiband compressors, stereo "enhancers," and aggressive limiters and automatic gain control prior to their transmitter. So in addition to potential hyper-compression during mixing and mastering, the mix is squashed and processed even more by these usually poor-sounding devices and processes.

TURN OFF THE RADIO!

The radio is not a source of good sound. In addition to pre-transmission processing described in the main text, most broadcasters now use automation software that stores music as lossy compressed files, even for analog broadcast. All digital radio is based on lossy compression codecs, and although some sound better than others, current satellite radio formats exhibit audible artifacts because of the lossy data compression used. Streaming internet radio can be even worse! Many of the younger generation stream radio stations or music on their mobile phones – at truly awful bitrates, much lower than computer-based streaming.

Luckily there is a growing trend against extreme hyper-compression, and more and more artists are bucking the trend. Normalization (level matching) technologies found in playback software and hardware are often automatically employed by cloud and

streaming services, and media playing computer software such as iTunes. Legally mandated, average level standards are being increasingly imposed on broadcasters. All of this makes hyper-compression less beneficial – the systems turn down tracks with higher average levels, so they won't be any louder than better sounding non-hyper-compressed tracks with lower average levels!

Why does this matter? Products that were a casualty of the loudness war are not desirable to emulate, and in the future will certainly not be considered desirable goals.

To learn, evaluate, and analyze good sound, you need to listen to industry-respected recordings, from uncompressed or lossless source formats.

WHAT ABOUT DISTORTION?

Is distortion bad? Well…yes. And no.

Distortion as a result of digital recording levels being too high, digital mix bus levels being over-hot, clipping processors or plug-ins – yes, that's bad. But as discussed earlier, gentle overloading of analog circuits and magnetic tape can produce pleasing effects. What would the sound of an electric guitar be without gentle, moderate, and severe forms of analog distortion, and digital pedal or processor emulations of those characteristics? What would bands such as Nine Inch Nails sound like without deliberate bit-crushing digital distortion?

Distortion is an essential part of some instrument sounds! A clean DI'd guitar track is usually out of place, inappropriate, and just does not work in a heavy rock or blues context – a stylistically appropriate guitar sound has either gentle distortion on the attacks of the notes, or more aggressive, continuous, self-compressing distortion throughout the notes. The loudspeaker cone of a guitar or bass amp distorts the electrical audio signal fed into it, even if all distortion knobs are set to "off." Non-linearities in the cone's behavior are technically distortion, and produce the sound's unique character. The rotating horn in a Leslie speaker distorts an organ's output waveform, and again, this distortion is the instrument's sound and character – without it, a B3 just wouldn't be the same!

Gentle distortions are why one piece of gear sounds different to another, and can be used to give sounds a little extra character. Moderate or severe forms of distortion can be used for creative effect, giving elements of a mix hyper-character, grit and grunt when needed.

Once you have learned how to create clean, technically perfect mixes, carefully applied distortion can be a very powerful and effective tool. As part of the recording chain, it is fair to say that any accidental digital distortion, or much more than gentle analog distortion, is undesirable. You cannot undo distortion, so even if you want to use

distortion effects creatively in your mix, it's better to track and record "clean," and add distortion in a controlled way as part of the mix – in context, when you know how much is needed.

WHAT IS A GOOD RECORDING?

Musicians practice scales as a prerequisite to finding their own style. A similar prerequisite for an aspiring audio professional would be learning to make simple, technically and artistically correct recordings prior to exploiting more creative stylistic techniques. "Technically correct" means free of any undesirable technical artifacts. So what makes a technically good recording?

- Good sounding sources and musicians are a prerequisite.
- Appropriate microphone choice.
- Good mic placement and mic techniques.
- No noise or distortion problems created by incorrect or inappropriate use of any of the equipment in the recording and mixing chain.
- Good balances and use of the stereo soundstage.

"Artistically correct" means that the recording and mixing styles are appropriate to the project, and musical style.

By understanding how to capture, process, and mix sound using equipment technically correctly, you learn to:

- Use the equipment to control the sound.
- Really hear the effect of more creative, artistic use of the equipment.
- Anticipate how creative processing may benefit a project you're working on.

BUT IT'S ABOUT THE MUSIC, NOT THE RECORDING!

Let's not forget one important thing – the music creates a hit song, not the recording! Many hit records are not technically perfect – there may be minor engineering mistakes and errors because a great musical performance trumps a little distortion on a killer vocal take!

ACCURACY

When learning to record and mix, one characteristic to aim for is accuracy. Is the recorded sound a faithful reproduction of the instrument or singer? If acoustic musicians are professional and used to playing together, they know how to blend themselves. You, the recording engineer simply need to capture their performance appropriately. The recording room, mic choice, and mic placement are huge factors that

impact the characteristics of a recording. Accuracy is unachievable if you place the wrong mic in the wrong position on an instrument in a bad room – all you can do is wrestle the sound somewhat into shape as part of the mixing process.

Great sound sources, a good sounding room, the right microphones, and good mic technique will capture sounds that mix themselves more. Getting the initial recording right results in a better, quicker, and easier mix.

NON-NATURAL SOUNDS AND BALANCES

Having discussed the importance of accuracy in a recording, it has to be admitted that most pop and rock music is not about overall accuracy – it's about achieving a sound that is stylistically and artistically appropriate. A good example of this is the sound of a modern rock drum set – the recorded kick, snare, and tom tom sounds are quite different to how they sound naturally from a normal listening position. Microphone and production techniques are exploited to make those sounds larger than life, phat, and in-your-face. They are not necessarily accurate, but they are certainly stylistically and artistically desirable.

In real life, most singers would not be heard above the naturally much louder drums, guitar, and bass amps of a typical rock band. Recording equipment allows the engineer to create non-natural balances so the singer is heard clearly. Effects such as compression and reverb are used creatively during mixing, to improve the way the sounds work as a recording. The balance and mix of sounds and effects we're used to hearing in pop music has little to do with natural balance, but everything to do with style.

"Accuracy" in pop music involves capturing the actual source material in a way that allows construction of stylistically appropriate sounds. If you are not familiar with either the natural sound of the instruments being recorded, or the stylistic goals of the type of project you're working on, how can you record appropriate tracks and create a good mix from them?

It is important to listen, listen, and listen to industry-respected recordings of styles you might one day be called to work on.

WHAT ARE THE ELEMENTS OF A GOOD MIX?

Regardless of musical or production style, there are some fundamental characteristics that are essential and common to any good mix. They include:

- Appropriate frequency balance.
- Clarity and intelligibility.
- Effective use of the stereo image, and stereo imaging concepts.

- Effective use of soundstage depth, and front/back imaging concepts.
- Appropriate focus and amplitude balance.
- Good use of processing and effects.

FREQUENCY BALANCE

A mix should have appropriate amounts of low, mid, and high frequency content. A *real time analyzer* (RTA) is a tool that can be used to visually show frequency content. Figure 2.4 shows an RTA plot of a great sounding mix. The RTA is set to average out the frequency content over fairly long intervals of time, and to respond fairly slowly. If the RTA shows instantaneous readings, it jumps around too much, becomes difficult to interpret, and is dissimilar to human hearing (which has a tendency to determine frequency balance based on longer-term averages).

An "ideal" frequency curve, based on current mixing trends, is for a slight hype in the low end, relatively "flat" content into the mid range, and a gentle roll off up into the highest frequencies.

- If a mix doesn't have enough bass it will sound thin and lacking in power.
- Not enough mid range, and it will sound distant, thumpy, sizzly, and lack diction and clarity.
- Not enough high frequency content, and it will sound dull.

Visual tools are just tools to help train or confirm what the ear is hearing – your trained ear should always be the final judge.

A good way to confirm that your mix is on the right track is to have a respected, good-sounding mix available for playback, and to frequently and quickly switch to, and A/B it with your mix in progress. The commercial mix will probably be louder and more "in-your-face" because it has been mastered, but you should listen to frequency content, not overall loudness.

Your choice of microphone, the position you put it in, and the room you record in are just a few of many variables that impact the frequency balance of the sounds you record. It's essential that you record a sound that will give you the frequency content you need of that sound source in the mix.

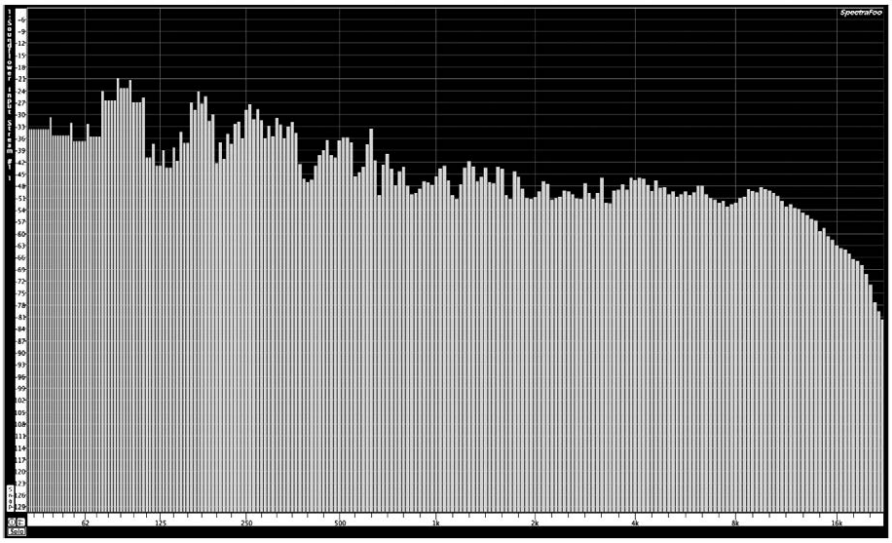


Figure 2 .4 *An RTA plot of a great mix. Note the relative "flatness" of the middle of the plot, with a gentle rise in the bass (on the left), and a rolling off of the high frequencies above about 6 KHz.*

SOME ESSENTIAL FREQUENCY CHARACTERISTICS TO LEARN TO IDENTIFY

• 150 Hz and below: Is the low frequency content of the kick and bass guitar/bass line boomy and undefined, tight and full, or small and compact sounding?

• 150 to 400 Hz: Does too much of the low-mid frequency range cloud the mix and make it too thick and muffled? Or does the mix lack fullness and body because there is too little of this frequency range?

• 400 to 500 Hz: Too much of this frequency range can give the mix a confused, boxy sound. Try banging a cardboard box with a stick, or talking into a cardboard box to get an idea of this quality!

• 600 to 800Hz: Not quite a nasal sound, but like talking into a toilet roll tube. Too little of this frequency range, and many sounds lack definition. Too much of this range results in a honky sound.

• 1 to 1.6 KHz: A lot of nasal, diction, and definition characteristics are contained in this frequency range. Too much of this range can result in a thin, nasal, AM radio, megaphone-like sound. Not enough, and a "smile curve" type of EQ can result, where the mix lacks power and sounds distant.

• 2 to 3 KHz: Probably the ugliest sounding frequency range! Too little of it and the smile curve effects described above can result. Too much, and your mix could have a very cheap

sounding, thin, tinny timbre.

• Around 4 Khz: “Not quite high frequencies” but “not quite mid frequencies”! Too much can give a harsh edge to instruments such as horns, guitars, drums, and cymbals. Too little of this frequency range can reduce a mix’s clarity, particularly of vocals.

• Around 8 KHz: “Proper” brightness and high frequencies. Too little content in this range makes mixes dull and lifeless. Too much, and the mix can become over splashy, too bright, or too sibilant.

• Above 12 KHz: This range contains the “sprinkles” – the magic dust that can give your mix “air,” and a sparkly sheen. Too little of this frequency range and the mix can sound flat and unexciting. Too much, and the mix will be too sizzly.

Figure 2.5 shows an RTA plot of a mix with an inappropriate frequency balance. The low and high frequencies are too loud, and the mid-range too quiet. This mix will have a boomy, bright, and sizzly sound that lacks definition. This is what is commonly referred to as the smile curve or loudness curve sound, and is commonly dialed in by many consumers on their equalizers, or by pressing the “loudness” button on some playback devices. It can be instantly pleasing because it hypes the extreme frequency ranges our ears are less sensitive to, particularly when listening at lower volumes. But at higher volumes, it creates an inappropriate frequency balance.

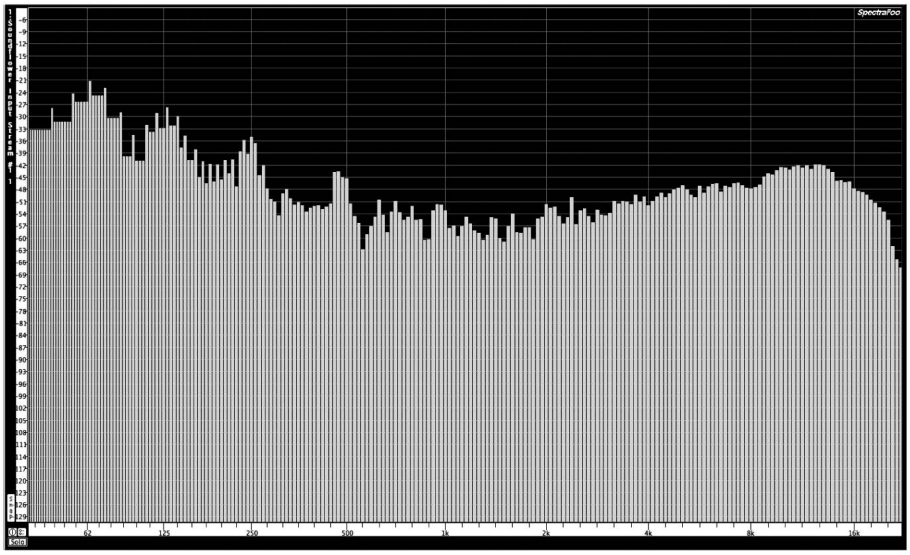


Figure 2 .5 *An RTA plot of a mix with too much low and high frequency content. Note the areas of “smile curve” hype in the low (left) and high (right) extremes.*

Figure 2.6 shows the RTA plot for a mix that is lacking in bass and high frequencies. It will sound thin, AM radio like, and similar to cheap earbuds! If you only listen on earbuds, then this is the type of sound you might think is desirable – because it's the way you're used to hearing things. If you do think cheap earbuds sound good, invest in some good loudspeakers, throw away the earbuds, and re-train your ears to recognize the characteristics of a good mix on good loudspeakers – before you record or mix anything else!

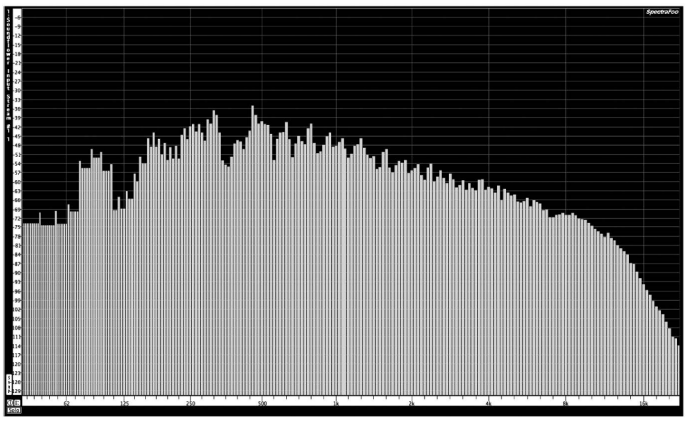


Figure 2 .6 *An RTA plot of a mix that is lacking low and high frequencies. The prominence of mid frequencies produces a thin, AM radio, or cheap earbud sound.*

CLARITY AND INTELLIGIBILITY

Every element in a mix should be able to be heard clearly, or appropriately to its artistic function. Clarity and intelligibility are products of many things, including:

- Overall frequency balance: If the overall frequency balance of the mix is incorrect, then that will negatively impact the listener's ability to hear everything appropriately in the mix.
- Frequency content of the individual elements in the mix: By nature of the format, a stereo mix places many constituent sounds in a relatively small space – the space between the loudspeakers, or between the headphone drivers. When sounds are positioned on top of each other, or in close proximity to each other, they become more difficult to accurately interpret, and their frequency contents sum together to form a new, combined frequency balance.
- Spatial positioning and panning: By panning sounds to different positions we physically separate them, making them clearer and improving the clarity of a mix –

as well as making it spatially more interesting to listen to.

The individual sounds or tracks in a mix might sound great when listened to in isolation, but when put together the sound can become muddled and unclear. Equalization (EQ) can be used to fix this. The unnecessary, or less important frequency components of a sound can be attenuated (de-emphasized) so that more important components of other sounds can occupy that frequency space. This enables both sounds to be heard more easily because they are no longer "fighting" with each other in that same overlapping frequency range.

EQ-ING FOR CLARITY – GUITARS AND VOCALS

PROBLEM: A distorted electric guitar and vocal may both have a lot of frequency content around 2 KHz. They both get in the way of each other in this essential vocal diction and intelligibility frequency range. Neither is clear.

WRONG SOLUTION: Turning one up just obscures the other more, and/or makes the track too loud.

CORRECT SOLUTION: The guitar is less important musically, and also has a greater amount of other beneficially usable frequency content than the vocal – it has its body at lower frequencies, and brightness at around 5 KHz. An EQ attenuation around 2 to 3 KHz on the guitar can set it back in the mix a little, and make that frequency space available to more essential vocal diction and intelligibility frequencies. The problem is solved without turning anything up.

EQ-ING FOR CLARITY – KICK DRUM AND BASS GUITAR

PROBLEM: Kick drums and bass guitars frequently "fight" due to overlapping low frequency content. If a frequency range is congested, then all elements competing in that range lose.

SOLUTION: Attenuating the kick drum in a frequency range above its fundamental pitch (its "boom") can de-clutter frequency space that the bass can then occupy and be more clearly heard. De-emphasizing the bass, around the frequency of the kick drum's fundamental boom will allow the beef and boom of the kick to be heard more clearly.

MUSICAL ARRANGEMENT

Like many things discussed in these early chapters, this has nothing to do with microphones or mic techniques, but the musical arrangement of a song can be the difference between a mix and song that works and sounds great, or one that doesn't

reflect either the band or you as an engineer favorably! The clarity of a mix and the transparency of the sounds in a mix can often be improved by reducing the number of instruments playing simultaneously, or by changing an instrument's part so it is not playing in the same range as another instrument. Musical arrangements should ideally be cleaned up in preproduction meetings well before any mics are set up in the studio. The best mics in the world, good-sounding instruments, and amazing musicians won't help you produce a great mix if the musical arrangement is poor.

Depending on your relationship with the musicians, and your role in the sessions, you may be able to suggest instrumentation and arrangement changes. Even if a band insists on tracking everything (the same way they thrash through a song at a local bar gig), the mixer or DAW has *faders* and *mute* buttons! Just because something is recorded, doesn't mean it has to be used. Taking tracks out of the mix, or turning unimportant tracks down, can improve the clarity and focus of an otherwise cluttered, muddled, and over-busy mix.

If the band or producer is adamant that everything has to remain in, try to find time to do an extra mix, as you would prefer it – even if it's on your own (unpaid) time. Both can be presented to the client. You never know, they might like the stripped down version – and if not, you have a better mix for your resume, demo disc, or professional satisfaction! Regardless, good musicians and good people will be impressed that you cared enough about their project to go the extra mile.

THE STEREO IMAGE

A good mix really takes advantage of the space between the loudspeakers. Leaving all the sound sources panned centrally, parked on top of one another, does not effectively use this space – it's boring, and things are difficult to hear and comprehend. Having sounds come from different locations in the stereo image not only makes the mix more interesting to listen to, but it also means that sounds can have their own spaces – we can hear each element of the mix more clearly, and the clarity and intelligibility of the mix is improved.

FOCUS AND AMPLITUDE BALANCE

Poor amplitude balance between the elements in a mix will negatively affect frequency balance and clarity. For example, if the bass is turned up too loud it will obscure the kick drum, guitars, and keyboard sounds. It will also make the mix generally too boomy, and lacking in mid frequency punch and high frequency brightness.

Amplitude balance is dependent on frequency balance – a good amplitude balance

cannot be obtained until a good frequency balance has been worked out between all the tracks. But a good frequency balance is reliant on the amplitude balance of the tracks! Everything is dependent on what you've done to it, what you've done to everything else, and what you haven't yet done to everything else! Every time you change an amplitude or EQ setting, you have to go back and re-evaluate its effect on every other track or channel you've already worked on. You must be prepared to rework every other track or channel every time you change something!

This sounds like an impossible process, but a common approach might be to:

1. EQ a sound so it sounds good by itself. Or better yet, EQ and balance (channel fader) a small group of related tracks so they sound good together. Don't get too hung up on how an individual soloed track sounds – even though something sounds amazing by itself, it probably won't fit in the complete mix, and may cause other elements to disappear!

2. Add and balance another track, or small group of related tracks – not all of the mix elements, but just more. EQ and balance them so they work with the track(s) worked on previously.

3. Listen carefully to ensure the previously worked on tracks still sound good since the addition of the new tracks. Also, if a new track just isn't clear, or "popping" the way it should, the cause could be some of the previously worked on tracks getting in its way. Go back and further refine the EQ and amplitude balance of the previously existing tracks that need it so the new combination of tracks all work together.

But what is a good amplitude balance?

You need to be familiar with the typical amplitude balances of whatever style of music you're working on (or may unexpectedly end up working on). The desired sound and amplitude of the kick drum and bass in a jazz setting are very different to their respective sounds and levels in a rock project. The only way to learn what is desirable in a particular musical or production style is to listen to industry-respected recordings.

Additionally, a mix needs focus – a focal point. What is the most important element of the mix? In vocally driven music it is of course the lead vocal. If the lead vocal cannot be heard and clearly understood, either because of poor frequency balance, clarity, or amplitude balance problems, the mix has failed.

Prior to the 1980s, vocals were generally louder in the mix than they are today. One current trend is to mix vocals more on par with other elements, rather than having

them "on top." In some mixes, the vocals can actually be quite quiet, but due to their frequency content and the arrangement of the other instrumental tracks around them, they are clearly identified as the focal point. It's harder to mix in this contemporary way, due to the collisions of similar frequencies in the vocal and other instruments that are used at similar amplitudes to the vocal. In addition to more creative EQ, effective musical arranging, and the careful selection of sounds in the mix, the source recording needs to be made with this production aesthetic in mind.

In pop and rock music, the drums drive the rhythm of most songs – particularly the kick drum and snare drum. Usually they are mixed at about equal levels, and of equal focus to the lead vocal (or for dance music they are often the loudest elements in the mix). If the drums are too quiet, the mix will lack rhythmic drive. If the bass is too quiet the mix will lack a solid foundation and "bottom."

There are different ways to develop a mix:

1. Some engineers start with the drums, add the bass, add the rhythm section, and then park the focal point, usually the vocals, on top.

2. Other engineers start with the focal point, and then mix the other instrumentation in around it.

No method is right or wrong. *You* have to develop a process that works for you, and the type of product you are producing. For novice engineers, I suggest mastering the first method before moving on to the second. One thing is guaranteed though – you will end up with very different results using each of these methods – and if your usual method is the first, you should definitely give the second a go! Different projects, music, or production styles will benefit from these different approaches.

PROCESSING AND EFFECTS

Compression and reverb are like audio "glue" and "makeup." Compression can be used to tighten individual sounds, or to give them power and punch. It can also be used gently on the master output bus to apply a bit of glue that gels the whole mix together. Reverb, like real makeup, can smooth over slight blemishes on individual tracks, and generally make things more pleasing.

We usually hear acoustic musical sounds in a room or hall – an enclosed space of some kind. That space (unless it is the open air) imposes its reverberant characteristics on the sound. Some recording studio environments are acoustically dead and dry – particularly iso-booths and home studios treated with acoustical absorption products. When a microphone is positioned relatively close to a sound source in an acoustically dry environment, the recording lacks the spatial reverberation that we are used to hearing

– so artificial reverb can be used to put that sound source back in a more characterful space.

The choice of reverb character is very style and tempo dependent – and it's true that pop music mixing styles have become relatively dry sounding these days. But that doesn't necessarily mean that no effects are used – short room characteristics, as opposed to long swishy halls, are often used, and early reflections can be exploited instead of reverb tails.

Drums sound dull and like the life has been sucked out of them without the sound of a bright reflective room around them. This is why the best professional recording studios have dedicated drum rooms that are relatively reflective – so excitement and energy does not have to be artificially added. Vocals and solo instruments usually sound smoother, and more professional and polished after the addition of some reverb. A common reverb applied to an entire mix, either during the mixing or mastering stage of production, acts like "acoustic glue," gelling the entire band together in a similar environment.

Delays, choruses, flanging, distortion, and a multitude of other creative effects are powerful production tools that should also be used when making an artistically creative recording. Mixes without processing and effects are not as interesting as they otherwise could be!

SONG STRUCTURE AND ARRANGEMENT

Rarely does having all the musicians thrash away on the same riff or rhythm incessantly for an entire song produce an interesting record that a listener wants to listen to repeatedly. A great record is all about drawing the listener in – to do that there must be flow, development, tension, and release:

• BORING: A verse of vocal and full instrumentation followed by chorus of the same full instrumentation. In terms of intensity, the song hasn't gone anywhere and has nowhere to go.

• INTERESTING INTENSITY AND TEXTURAL CHANGE: Having only the drums, bass, and keyboard in the mix for the verse, before adding the guitars for the chorus.

A great band, made up of experienced musicians who are good arrangers, usually orchestrates itself. Different musicians and musical parts will come in and out effectively, and they will naturally pitch their parts so that they aren't in similar octaves or frequency ranges. Mixing is a lot easier when a band does this for you!

Other bands, whose regular performances might be thrashing away at a local bar or

club, might need some production assistance in order to get their songs (which may well be acceptable alongside the other distractions of the live performance environment) to translate and work well as recordings – where the sole focus is the music. If the guitar and keyboards are playing material that is too similar, it may be beneficial to suggest that one of the musicians tries something a little different. This could be as subtle as changing the synth sound, slightly varying the rhythm of a part, changing the octave, or even not playing that section.

How much you, the engineer, can suggest to the band musically depends on your relationship with the performers – and whether there is an actual producer present. If there is a producer, you should keep your mouth closed, and not have an opinion even when asked! If you are hired because of your relationship with the musicians, and you are respected for your opinions, then it would be appropriate to educate the band, and shape their performance into a better recording.

MAKING A GREAT RECORD

A GREAT PERFORMANCE VS THE BEST TECHNOLOGY

Time for a reality check!

The best equipment, and recording and mixing techniques can't turn a poor performance, or bad sounding instruments, into anything other than a "slightly improved" version. And intonation processing software can only do so much before it turns the vocal into a "robot voice"!

It is very frustrating trying to record and mix bad sounding instruments or singers! Corrective EQ can only do so much, and only fix certain problems – for example, a poorly tuned drum set, drum set rattles and buzzes, loudspeaker cone noises, or wrong inflections in the vocal performance *cannot* be fixed in the mix. If you find yourself spending a long time struggling to fix one specific problem, it could be caused by issues not related to recording or mixing. Solutions?

- Locate better sounding instruments.
- Exercise producer skills and coax, guide, and nurture the performer(s) to a confident performance.
- (Or get better performers!)

WHAT DO YOU REMEMBER ABOUT YOUR FAVORITE SONG?

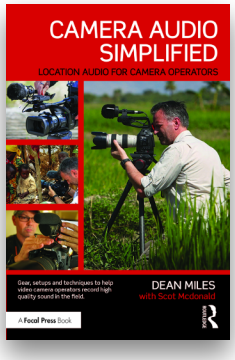
Is it the pristine vocal sound recorded with a very expensive high end condenser microphone? Or is it the catchy hook and the words – which were captured with a cheaper dynamic mic, because that's what was plugged in when the engineer decided to hit the record button "just in case"?

For audio people it's probably (and hopefully!) a combination of both. But for most consumers, it's usually the latter. Of course, better equipment (a great mic, and a great pre-amp) mean better sound, and that has to be desirable – but the fact of the matter is that many top selling records have been recorded on less than the best equipment, and do sometimes exhibit technical problems that are easily overlooked in the context of a great musical performance.

While "we'll fix it in the mix" is certainly a myth, the goal of 100 percent technical perfection does have to be balanced with knowing when a magical and unrepeatable musical performance has been captured.

CHAPTER 2

Sound Elements



Learn more:

Sound Elements

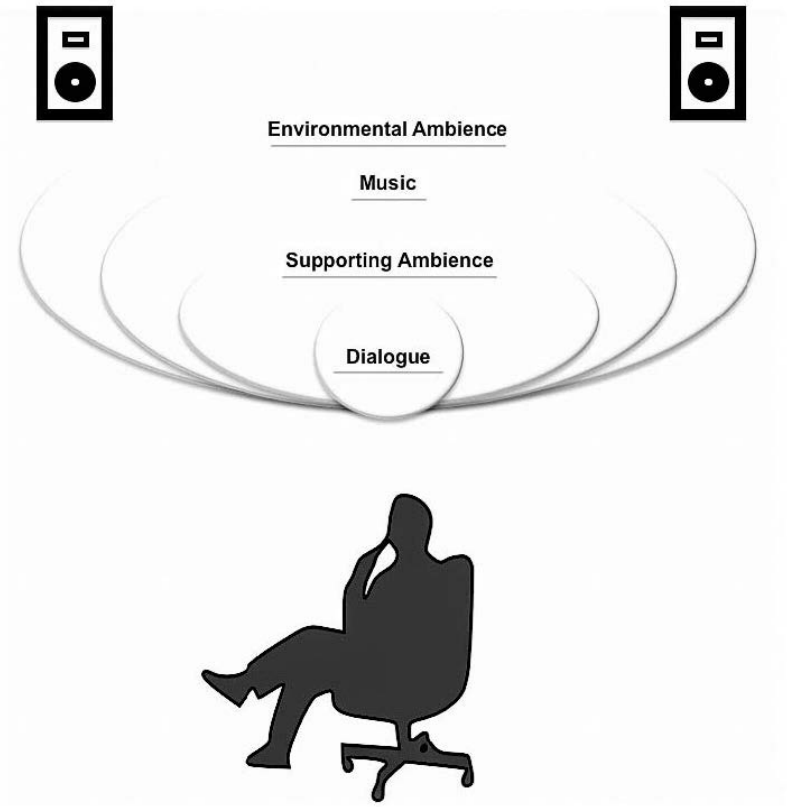
Sound Elements Sound elements are components of sound that are recorded during the production process and then edited together to create the rich stereo mix that we all know and love. When I'm creating a stereo mix for broadcast, I need a variety of things to get the job done right. These include dialogue, supporting ambience such as room tone, environmental ambience such as the din of city traffic or birds and wind in a park, and finally, music or an orchestrated soundscape of some kind.

What sounds would bring this shot to life?

A good way to understand sound mixing is to remember that sound has depth of field just like a visual image. When everything is in perfect focus, there's a lack of separation and all the elements compete equally for your attention. But when there's a shallow depth of field and only the primary subject is in focus, the image now has depth and separation between the various elements. It purposely draws your attention to a specific spot. Sound is very much the same. So, when a beautifully composed image and a perfectly mixed sound track work together, it's magic time! Let's use the shooting of an in-depth, sit-down interview to help explain how to layer sound to create a stereo sound mix. When you shoot an interview, you draw the viewer's eyes to the subject by having them sharply in focus and separated from the background. We do the same with sound. Dialogue is typically the loudest sound element, and by panning it to the middle (so the sound is in the centre of the mix) it's clear, sharp and draws your attention. Just

like your picture, the dialogue is in focus. The next layer of your shot usually consists of elements that are slightly behind the subject such as pieces of furniture, artwork, lamps, etc. These are slightly out of focus to provide context as well as visual depth and support to the overall image. Supporting ambience does the same thing. Adding room tone and/or the ambience of your b-reel supports the dialogue and the visuals. It's not as loud as the dialogue and is panned a little wider. You can still make out what the sounds are, but they're not the focus of your listening. The third layer in your shot is the stuff in the deep background or outside edges—often completely out of focus. It doesn't really need to be there but it creates texture and richness—makes your shot more interesting without calling attention to itself. This is true for the environmental ambience of a stereo mix. It's the din of the location and is panned hard left and right in the stereo mix. The overall mix would still work if it wasn't there, but just like your picture, it adds depth and richness. You probably wouldn't notice it unless someone pointed it out. Music also plays an important part in supporting the dialogue as well as the visuals. You can clearly hear it but it's not as loud as the dialogue—it supports the content by creating emotion, tension, sadness, etc.

That's one ugly bird, but the photo's perfect to help demonstrate layers



So when I'm sitting in my mixing chair with all the sound elements I need, I layer them like this:

- I have dialogue panned right up the middle and it's the loudest sound element—that's the sound I want the viewer to focus on
- The supporting ambience is a mix of room tone and b-reel ambience. It's panned a little wider, and not as loud as the dialogue so not to compete for the viewer's attention
- The environmental ambience is panned hard or fully left and right. It's the air of the location that anchors the mix and helps create a continuous bed of sound for the rest of the sound elements to sit in and move in and out of
- Finally, the music is panned a little wider than the supporting ambience and at a volume where you can hear it. But again, it's not competing for the viewer's attention

So why am I going through all the trouble of teaching you about sound mixing? Well, when you're shooting by yourself, you're the one responsible for capturing most—if not all—of these sound elements. If you don't, the post audio editor (which may be you) won't have the elements needed to create the beautiful stereo mix you were hoping for.

Sound Elements You Need to Capture While Shooting

Here are the three sound elements you are responsible for capturing when on location:
Dialogue

Supporting Ambience

Environmental Ambience

Dialogue

As the one-person-camera crew, your first and most critical audio recording will be dialogue. Record intelligible dialogue that meets all sound requirements you learned in Chapter 2. All dialogue recording will be mono! Lavaliers, camera-mounted mics, and handheld mics will all be mono. If you think recording with a stereo camera mic is better, it's actually not.

These are all mono microphones except the built-in camera mic

Stereo dialogue recordings don't play nice with mono dialogue recordings, and they always sound funny in a mix. If you want to record dialogue with a camera-mounted mic (and you will) and you're recording with either a built-in camera mic or a third-party stereo camera mic— you should replace it with a mono version.

> Although **room tone** is part of the supporting ambience category, I want to talk about it now because you'll record it with the same mic you recorded your dialogue. Room tone is a dialogue-free recording of the room's ambience that's crucial for creating smooth audio edits. The best time to do this is immediately after you've finished a dialogue recording—such as after an interview or several shots for a scene.

Ask everyone to stay where they are and stop talking. Everyone was absolutely quiet during the entire interview so it's not a problem to ask them to be quiet for another

minute or so. Widen out your frame to visually indicate you're recording room tone so it's easy to find during editing. It's important to record with the same mic in the exact same position if at all possible. Start recording and identify what you're doing by saying "30 seconds of room tone." Record for a minimum of 30 seconds.

Remember, your room tone should sound relatively the same as it did during the interview or scene. If a lawnmower just started up and it wasn't there during the interview then it's not going to do you much good.

Avoid the temptation to think you'll grab it later or at the end of the shooting day. Chances are you'll forget all about it and finally remember in the car on your way home!

Supporting Ambience

This is audio recorded when shooting b-reel with the camera-mounted mic. It should be usable whenever possible. Here are a few things you can easily do while shooting b-reel to improve the ambience tracks:

• Be quiet! I'm not sure if most camera ops have A.D.D., but as soon as you hit record on a locked-off shot, you're rifling through your camera bag! It's a cacophony of zippers and Velcro!

• No talking. There's nothing worse than people talking when shooting b-reel, it renders the ambience useless. This would include you, Mr. or Mrs. Camera Operator. Camera ops and directors are the biggest offenders when it comes to chatting all over b-reel ambience— zip it!

• Be aware of the ambience around you. Listen for sounds that don't work with what you've got framed. People talking in the background (even quietly) can cause distractions and if an extraneous sound occurs while you're shooting, shoot it again! You'd shoot again if the camera move wasn't perfect

• Make sure the ambience matches the frame. Long lens close ups can be problematic if the audio 15 feet away doesn't match the close-up. After chapter three: sound elements 33 you've finished with the long lens shot, move in and grab 15 seconds of ambience with the camera mic so you can replace the ambience in editing

• Make sure sound-specific shots have usable sound. If you're shooting b-reel that has sounds specific to that shot (a close-up of someone putting a golf ball) the sound needs to be usable. It's very difficult if not impossible to recreate these specific supporting ambience sounds in post

Environmental Ambience

Environmental ambience is a recording that's used to anchor the entire sound mix. As I described earlier, it's the element you wouldn't know is there unless someone pointed it out. Taking the time to record it does take away from shooting pretty pictures, but it's an extremely important sound element.

Now, how far away do I have to walk to gather some ambience?

Record this ambience bed separately from any shooting—it's for sound only. Use either the camera-mounted mic or a small, portable field recorder like a Zoom H4n. Now, don't freak out: it doesn't take very long. It's just going to require you to focus on audio for a short period of time.

Record the din or general ambience of the location you're working in. Remember, what you're recording is deep background sound. Think in terms of this sound being "out of focus"—just like in your pictures.

If we look at the golf course scenario for example, you'll need to move away from distinguishable sounds that are too specific (or "in focus") such as the noisy club house, golfers talking loudly, course maintenance workers operating equipment, etc. Walk out into the golf course and find a quiet place away from these sounds. Record a minimum of two minutes.

In remote locations, I often leave my Zoom recorder in an isolated spot to record while we continue our shooting day. I don't bother anyone to be quiet or waste production time sitting out there recording ambience because I'm going to assemble the usable parts after the shooting day. I usually leave it recording for 15 minutes or so, but I've left it for close to an hour on several occasions—I'm getting old!

I know you camera ops are all about visuals, but good audio makes your pictures look better. By taking some time to properly gather these sound elements, the overall quality of your projects will increase dramatically!

Remember, you're going to be judged by the final production and the sum of its parts. If your sound is amateur, your production is amateur!

CHAPTER 3

Demystifying Recording Levels

Learn more:

Demystifying Recording Levels

Although we have outlined exercises to develop your critical listening skills, explored pre-production and discussed compositional and arrangement concepts, we're not quite ready to jump into the recording process just yet. I know what you're thinking: 'I'm bored – just get to the bit with all the knobs and buttons!' While getting creative with production is why you are reading this, taking this time to really explore the foundations of the craft that are so often overlooked will save you a whole heap of pain later on and will allow you to concentrate on the 'good stuff'. So now it is time to recap on some fundamentals, bust some myths and outline some recording practices that will help you record

anything!

ANALOGUE AND DIGITAL AUDIO

Now I hope that all of you reading this will know the differences between analogue and digital audio, but let's just make sure. The process of changing a physical sound wave into an analogous electrical signal using a transducer (such as a microphone) is called analogue audio. In an analogue system, the instantaneous electrical level is directly proportional to the instantaneous air pressure captured by the transducer. The analogue signal is then amplified and can be stored on an analogue medium such as tape (which works by magnetization) or converted further into a series of discrete mathematical numbers. Once it is converted to a numeric form, it is called digital audio. Digital audio works by storing the amplitude of the analogous signal many thousand times a second. These 'snapshots' are called samples and the number of times per second a sample is taken is called the sampling rate. The amount of detail that each of these samples can store are its resolution or bit depth. Once the signal has been converted to a series of numbers, it can easily be stored or processed using a computer.

Both analogue and digital systems have inherent strengths as well as weaknesses, so it would be unfair to call either definitively superior.

Analogue pros

• An analogue signal is an extremely efficient way of representing sound, but its accuracy can be hampered by the recording chain and storage medium. Many producers feel that the subtle imperfections of analogue recording methods sound warmer and more pleasing to the ear.

• By pushing the signal level and overdriving, the analogue signal path can cause saturation, which might, in certain circumstances, deliver even more pleasing results than digital methods.

Analogue cons

- Recordings are susceptible to degradation.
- Copies of the original recording are noisier and more distorted.
- Editing is more difficult and time-consuming.
- Noise from the storage medium (tape hiss) and recording device become a part of the recording.

Digital pros

- Digital editing is more powerful and generally easier.
- Duplicates are exact copies that do not degrade.
- When recording at high enough bit depths, the dynamic range (range between noise floor and clipping point) of digital audio exceeds the human hearing ability.

Digital cons

- Overdriving a digital signal results in unmusical and harsh-sounding distortion.
- With lower bit depths, conversion between analogue and digital must be done carefully to avoid loss of fidelity.
- Digital recordings have been called sonically 'colder' or 'more sterile' than fully analogue recordings. While this is a matter of taste, quality of equipment and processing methods, I have found merit to analogue recording techniques.

Because of the distinct advantages during editing and storage, digital audio has become the prevalent recording medium over the last few decades, so much so that many of today's aspiring engineers know a lot about sampling rates or bit depths but very little of the fundamental electronics terminology that would have been staple knowledge in the 1960s–1980s. Despite the popularity of digital audio, knowledge of basic electronic principles is still relevant in a modern recording environment, not least because the signal will be undoubtedly converted to or from an analogous form at some point (even if it is just to drive your speakers). As well as this, you will often encounter analogue gear and digital models of the same classic gear. Knowing basic electronic principles will also help you to avoid breaking equipment, for example by incorrectly matching amplifiers to speaker cabinets.

Luckily for you, Appendix A covers the rudiments of simple electronic circuitry such as voltage (volts, V), power (watts, W), current (amps, A) and impedance/ resistance (ohms, Ω).

INTRODUCTION TO GAIN STAGING AND METERING

Before we get into gain staging, let's quickly define some terms that will be used throughout this section. First, on a system that captures and reproduces audio, there will always be two operational extremes. The noise floor defines the point at which a signal is so low that hiss and other system noise begin to have a detrimental effect on the quality of capture. The other extreme is the clipping point, which is where a signal level is so high that it cannot be accommodated by the system; therefore, the signal is distorted in such a way that it is detrimental.

Two other terms are derived from these to help describe the quality of audio capture:

1. *Signal to noise ratio* – the level of noise in a recording in relation to usable signal. It can also be written SNR or S/N.

2. *Headroom* – this is the range between your operating level and the system's clipping point.

Gain structure (also known as *gain staging*) is a hot topic in the audiophile community. Bad puns aside, it is basic, key knowledge that is often overlooked entirely. Put simply, gain staging is the process of setting your levels throughout the recording chain. The art of gain staging is to record the signal 'hot' (loud) enough to avoid excess levels of unwanted noise in the recording but also give enough headroom to avoid the audio distorting in unpleasant ways. Over the years, our metering needs have changed with technological advancements, and because of this there are several common metering scales still being used in the audio industry, which has led to some confusion. The period of transition when digital audio was first introduced was particularly troublesome. What is worse is that these scales are completely different and sometimes even incompatible with each other. Demystifying these types of meters is crucial in the modern recording studio, where you are likely to see:

- analogue peak meters;
- VU (volume unit) meters;
- analogue dBu meters;
- digital peak meters; and
- digital RMS ('root mean square')

Before we start to explore the different types of metering, it is important to outline the biggest fundamental difference between analogue and digital metering. On a digital system, zero is the maximum operating level – go any further and undesirable distortion will be introduced. In an analogue scale, zero is the optimum operating level – levels can go beyond this without distortion, and even when the point of distortion is reached it might even be desirable. As we move through this chapter, we will elaborate further on the whys and hows, but understanding this fundamental difference will help

you get to grips with gain staging much more quickly.

Before we outline the common types of audio metering, there is one more caveat. Remember, from Chapter 1: Production Philosophies, Your Ears and Critical Listening, the decibel is a unitless ratio, used to express a change in power. So to calculate absolute values, we need to have a reference value. The

same goes for all signal levels in audio equipment. Regardless of whether the scale is analogue or digital, it is only representing the strength of a signal based around a set reference; what changes between metering scales is one of two factors:

• The reference type and value used to measure the signal strength (e.g. 0.775 volts RMS, 1 milliwatt or 0 dBFS).

• Whether the meter is measuring peak or average (RMS, root mean square) signal level.

ANALOGUE METERS

Some of this next segment might get a bit maths-heavy. I have purposely omitted the working out as what is important is not the calculations themselves but remembering the relationships between scales.

The VU Meter

The VU (volume unit) meter was the first commonly found metering system in analogue pro-audio equipment. Its origins go all the way back to a time when all audio equipment had an input impedance or resistance of 600 Ω. This was due to the fact that when audio was transmitted over large distances, the length of wires could approach the electrical wavelength of the signal, causing reflective noise in the system. Reflection is caused by unmatched impedances, so, to maximize power transfer, a standard impedance of 600 Ω was used.

It was decided that the VU meter should read zero when there was an optimal balance between clean signal (i.e. minimal distortion) and low noise. This equated to a power level of 1 milliwatt (mW).

The VU meter (Figure 7.1) was designed to give an indication of average volume rather than peak levels. It was designed like this because it is similar to how your ears perceive loudness. Due to the slow response time of the needle (approximately 300 ms), they are unable to show peak values. You can find VU meters on much of the classic outboard gear that is still widely in use today.

Because the VU meter is based around an optimal level of performance between headroom and SNR, optimal recording levels were calculated on older large format mixing consoles by setting gain levels to meter roughly 0 VU (volume unit) on the loudest instruments.

Figure 7.1 Example of a VU meter on a UREI 1176 classic compressor (picture courtesy of Metropolis)

Since we know the resistance (600 Ω) and the power level (1 mW), using Ohm's law, the voltage needed to zero the VU meter can be calculated as 0.774596669

The dBu Meter

Since then, the pro-audio industry has moved away from having a standard resistance in equipment, but the same signal strength of 0.775 Vrms remained a reference value for professional audio equipment. A new scale was created to reflect this, known as 'dBu' (decibels unloaded), or, more specifically, '0 dBu'. These days, however, many pieces of equipment employ the dBu scale of reference using LED displays. As the dBu scale's reference is volts RMS, it is also an averaging meter and does not show peak values.

The PPM Meter

To combat this, another type of analogue meter was developed called a PPM (peak programme meter) and these were developed by the BBC to show peak values in the analogue domain.

Hybrids of these two types of meters were also developed and are called 'peakhold meters'. They show both averaged signal level and peak signals by displaying the average value of the signal as a solid bar and the peak value as a floating point at the top of the meter. This peak mark usually has a temporary or permanent 'hold' function to give us time to recognize the peak values.

The dBV Meter

The meters above will cover the majority of the analogue metering in pro-audio applications. However, it is worth noting that since then, a second separate operating level was devised for semi-pro or consumer audio equipment. This is based around an operating level of 1.0 Vrms, which is known as 'dBV'.

Analogue Operating Levels

We can use the power ratios I outlined in Appendix A to work out changes in voltage between signal levels. For example, when we see a figure of +3 dBu, it means the voltage gain has produced a power gain of +3 dB over 0 dBu (which is 0.775 Vrms). We work out the voltage gain ratio from the power gain to find the new voltage level of: 1.095 Vrms.

For +3 dBV, we would use the reference value V, 1.0 Vrms, finding the new voltage to be 1.41 Vrms.

These days, pro-audio equipment operates at +4 dBu, which translates as a nominal voltage of 1.227 Vrms. This level is called 'line level'. Without a 600 load, 0 VU is equal to +4 dBu.

Consumer audio gear operates at a nominal power of −10 dBV, meaning a nominal operating voltage of 0.316 Vrms. Often you will find switches on pro-audio equipment that can switch between +4 dBu and −10 dBV, which makes devices more compatible with each other.

As well as VU, dBu and dBV, there are many other analogue audio signal measurements, such as:

- Vpp (peak to peak voltage);
- dBr (decibels relative to reference level);
- mW (milli-watt); and
- dBm (decibel relative to a milli-watt).

This may be confusing, but the reason for so many scales is the fact that different applications have chosen different types and values as a reference point. In reality, even as an experienced engineer, I only know what about half of these actually mean. In practical terms, you won't come across them often, if ever.

Analogue Scale Headroom

Because analogue meters are centred on optimal recording levels, analogue scale meters have headroom above 0 before they reach their point of distortion. This point is often referred to as clipping. The point at which clipping starts to become audible depends on the quality of the equipment.

Some high-end gear even adds subjectively pleasing artefacts to a signal when it is softly distorting. Because of this, clipping wasn't the primary concern when it came to metering in the analogue domain.

It is a common rule of thumb that any peaks rising up to 8 dB above 0 VU (+4 dBu) are acceptable. Peaks beyond +12 dBu start to audibly distort (soft clipping), and the viability of using such levels is dependent upon the musical context and the quality of equipment used. Undesirable hard clipping can start to occur around +18 dBu or +24 dBu on high-end systems.

DIGITAL METERS

The invention of digital audio and the CD in the early 1980s caused a revolution in the pro-audio industry. Suddenly, recorded music was much more easily transferred, copied, edited and reproduced. Unlike analogue recording systems, digital signal processing produces no additional noise when changing the amplification of the signal or when going through a string of digital plug-ins (although it would amplify noise that is already there). Despite the practicality of digital audio, there are significant drawbacks: when digital audio clips, it degrades in a far harsher, more undesirable way. Analogue processing and its soft clipping properties create a 'warmer' sound that is subjectively more musical and pleasing to the ear.

The fact that digital clipping is so detrimental to sound quality meant that all current metering solutions were ineffective. It was no longer appropriate to have a meter that showed average volume, but instead a peak meter was needed to show the moment a signal clips. This led to the development of a new digital metering system called dBFS. FS stands for 'full scale', where 0 dBFS is the highest figure possible in the digital domain. Anything that goes beyond that level is a clip. Unlike analogue meters, a digital signal can only be represented by negative figures rising up to its maximum level (0 dBFS). In order to analyse digital average levels in your DAW, you need to use a dBFS RMS meter.

One other point to mention about digital signals is that once a signal reaches peak value, it will simply truncate (cut off) the wave. This is the main reason why digital clipping sounds so intrusive and far more audible. As well as truncation, another form of distortion called aliasing can also occur when a digital signal is clipped, which introduces artefacts. Aliasing will be explained shortly, but for now it is sufficient to say that when a digital system is overloaded, harmonic artefacts caused by aliasing might be introduced. These harmonics are, at worst, unrelated to the fundamental and, at best, less related, and are therefore dissonant/unpleasant to the ear. This is in contrast to an analogue signal: when an analogue signal overloads, it produces distortion that introduces more even-order harmonics, which are more heavily related to the fundamental and are therefore more pleasing to the ear. Therefore, analogue distortion is often referred to as sounding more 'musical'. This section is not meant to scare you away from digital recording; in fact, recent improvements in audio technology and metering, combined with sensible gain staging, means that truncation and aliasing

artefacts shouldn't be a concern as long as you operate the system correctly (i.e. within its limits).

Let's now look at the parameters involved in creating digital audio.

Conversion to and from Digital Audio

To create a digital audio signal, you first must convert a physical waveform into electrical energy (i.e. into an analogue audio signal). You then need an analogueto-digital converter (ADC). This device takes a measurement of the electrical voltage of the analogue signal and represents it as a binary number (made up of zeroes and ones) so that it can be sent to a computer. By taking samples thousands of times per second, you can get a very accurate representation of the original audio signal.

The same process applies in reverse when playing back a digital audio signal. First, you need to convert the digital information back into an analogue signal with a digital-to-analogue converter (DAC), then turn the electrical energy back into sound energy with a speaker.

There are two factors that determine the quality of a digital recording:

Sample rate

The rate at which the samples are captured or played back, measured in hertz (Hz), or samples per second. Common sample rates used are 44.1 KHz (44,100 samples per second), 48 KHz, 88.2 KHz and 96 KHz.

Higher sampling rates allow a digital recording to accurately record higher frequencies of sound. According to the *Shannon-Nyquist theorem*, the sampling rate you choose should be at least twice the highest frequency you want to reproduce.1 The highest frequency a system can accurately reproduce is called the Nyquist frequency, which is half of the sampling rate. If a frequency that is higher than the Nyquist frequency is calculated by a ADC or DAC, it is misrepresented and causes the phenomenon called *aliasing*.

To explain the effects of aliasing, let's imagine we are recording a sine wave (i.e. a single frequency) that is steadily increasing in pitch using a sampling rate set at 48 KHz. When the sine wave is at lower frequencies, the waveform is sampled at many points in its cycle. As the sine wave's frequency increases, less points in its cycle are represented. Once you reach the Nyquist frequency (in our case, 24 KHz), each cycle is only represented by two samples. If the sine wave's frequency rises past this point, the waveform will not be expressed accurately by the system. The result of inadequate sampling is that frequencies above the Nyquist frequency are misrepresented in a way

that makes them the equivalent of a lower frequency. The way that the signal is misrepresented is, in fact, highly predictable: it appears to 'reflect' around the Nyquist frequency. Using a 48 KHz system, this would make a 25 KHz sine wave indistinguishable from a 23 KHz sine wave, a 30 KHz sine wave indistinguishable from the 18 KHz and so on. Once you consider the implications of aliasing on a signal that contains more than one frequency, you can see how detrimental this can be to sound quality. To avoid this, a low-pass filter called an *anti-aliasing* filter is used to block out the frequencies above our audible range, therefore eliminating aliasing artefacts.(2)

The maximum hearing range for a human being is 20 Hz to 20,000 Hz, so any sampling rate above 40,000 Hz (40 KHz) would create an audibly perfect representation of any audible sound. The reason that the lowest standard sampling rate of 44.1 KHz (4,100 samples above 40 KHz) was chosen is due to slope of the anti-aliasing filter and the need to keep the entire audible range unaffected by the filter.

However, despite the anti-aliasing filter, it should be reiterated that aliasing will still occur when the signal has digitally clipped, as the distortion is introduced after the low-pass filtering stage.

Bit depth

This is the number of binary digits used to make the digital representation of each sample. The bigger the bit depth, the more accurate each sample is to its original analogue form.

There are two types of bit depth: fixed or floating-point. These types define how numbers are represented in binary form. A fixed-point bit depth has a set number of digits after the decimal point (or sometimes before), which means that fixed-point bit depths have a rigid minimum and maximum value, but between these ranges it is very accurate. A floating-point bit depth is able to move the decimal point dependent on the number of significant digits, which means that they have a far greater range. When a bit depth is listed, a fixed point system is assumed unless stated. A floating-point system is usually indicated by the word float listed in brackets after the bit depth (i.e. 32-bit (float)). Common bit depths are 16 bit, 24 bit and 32 bit (float).

Higher bit depths give a greater dynamic range and also provide a better SNR ratio, but come at the cost of using far more space on storage media. In a fixed point system, each extra bit gives 6 dB of extra usable dynamic range. Usable dynamic range is the difference in level between the system's clipping point and the level at which you hit the noise-floor. Therefore, theoretically, the dynamic range of fixed-point processing is 96 dB for 16-bit audio, and 144 dB for 24-bit. In reality, the figure for converters is slightly less than the theoretical values. Our own hearing has a dynamic range of approximately 120 dB from the threshold of hearing to the start of physical discomfort;

this means that 24-bit can easily cover the whole dynamic range of our hearing without clipping or hitting the noise floor. A 32-bit float system has a theoretical dynamic range of 1,680 dB, which is rather mismatched for the dynamic range of our hearing, but as we will soon find out it can have an advantage in digital signal processing.

Implementing a Suitable Sampling Rate and Bit Depth

When it comes to choosing a sampling rate and bit depth, there is some contrast in opinion between audio professionals regarding the quality benefits between settings. A rule of thumb is to use the highest sample rate and bit depth your computer can handle comfortably. This rule of thumb is good up to a point; when deciding settings, you should be concerned with the compromise between audio quality and computer performance.

Sampling Rate

We all know that computer speeds are improving all the time and it is no longer the case that you need the latest Pro Tools HDX rig to get the very best audio quality and professional-sounding mixes. You can make very professional sounding mixes with a sampling rate of 44.1 KHz and if that is all your computer can handle, then you are not going to lose out an awful lot. In fact, it is true that any waveform can be completely recreated by using a sampling rate of twice the highest frequency you are wishing to reproduce (Shannon-Nyquist theorem). Therefore, recording at 44.1 KHz or 96 KHz sounds identical. This may seem counter-intuitive but you just have to take it as fact. However, there is a slight advantage to using 96 KHz over 44.1 KHz in the digital processing that comes after you have recorded. To avoid overcomplicating matters, it simply makes the maths involved in non-linear plug-ins slightly more accurate. In fact, when using lower sample rates, many of these plug-ins actually up-sample within themselves, do its processing and then down-sample back down to the session's sample rate. It doesn't make a big difference, but if you've got the power, why not use it? I still currently mix at 48 KHz as I find it to be the most efficient sample rate for the speed of my computer and sizes of the mixes I am completing, and I know many top-level guys with various HD rigs and DSP chips who still do the same.

Please note that I do not advocate the use of 192 KHz sample rates, as any improvements are inaudible at best or actually less accurate at worst over 96 KHz.3 Also, the CPU power and storage facilities needed to process this information start to dramatically drain system resources and slow down workflow in even the best systems.

Bit Depth

Whenever audio is converted between analogue and digital, and vice versa, fixed-point systems are used. This means that any converter, interface or soundcard will be fixed-point. Modern soundcards can handle at least 24-bit recording and computer hard drives are big enough to handle their file sizes, so there are no real benefits to recording at 16-bit any more.

Once the signal is in the digital domain, your DAW can represent and store these 24-bit samples in a floating-point format. Such formatting will not increase the fidelity of the audio in any way (as the source was still a 24-bit resolution). However, there is an advantage; implementing 32-bit (float) within your DAW means that you cannot irreparably damage the audio by clipping a signal within the DAW (as the range of numbers is so huge). There is a catch, though; before the audio signal can be converted back to analogue again (for use on speakers, or to use analogue hardware), it has to formatted back into a fixed-point system. The DAW does this automatically but it means that the signal level must be back

within the range of a fixed-point system to avoid clipping.

If you practise proper gain staging, the difference between 24-bit and 32-bit (float) won't matter! But if you can afford the 50 per cent bigger file sizes of 32-bit float, a safety net can't hurt.

To summarize, you should really be looking to record at a minimum of 44.1 KHz and 24-bit to get professional results.

GAIN STAGING

The biggest problem that inexperienced engineers face when recording to a DAW is that they record everything too hot.4 There are a number of possible explanations for this – it could be confusion/hangover from the analogue days of 'set the levels to zero'; or to avoid the noise floor when recording at 16-bit; or just human instinct to think that 'the little bars should be near the top'. In your DAW, you should leave a sensible amount of headroom to 0 dBFS on every channel.

Here are several issues that will potentially cause problems if you don't leave enough internal headroom even when working all-digitally:

• You run the risk of clipping internally through plug-ins. Imagine that you are peaking at –3 dBFS on a snare drum track and you add a +3 dB high shelf to add more sparkle to the sound. You start to run a real risk of clipping the digital signal within the DAW. You can deal with this by turning down either the input or output of the plug-in down to compensate for the increase in signal. Alternatively, you could just leave more headroom initially.

• If you are using analogue-modelled plug-ins, they have been modelled to work best at a level equivalent of 0 VU, just like the originals. It is not as much that a plug-in will 'expect' a certain input level; it is just that many of these analogue models have the same varying knee behaviours and non-linearity just like the original units, thus making them sound better being run at the same level as in the analogue domain. Each plug-in will be calibrated to a particular level, which you can find in the product's manual. Waves CLA series classic compressors, for example, are modelled with −18 dBFS RMS equalling 0 VU.

• If you are running levels high, it means that your levels are going out to any auxiliary subgroups and master fader equally hot, and you will need to lower the faders significantly to prevent the signal clipping when the channels are summed together.

Because of these factors, I actively try to use similar headroom levels when mixing ITB (in the box) as you would do when using analogue or hybrid systems. It must be stated, however, that working with headroom in your DAW does not mean that quality of audio is adversely affected. The mixer in your DAW is clean right up to the point of clipping (and beyond, with a floating-point bit depth); therefore, the benefits are down to practicality and improved performance from analogue hardware and emulations of such equipment. There are no adverse effects in audio quality in the DAW as long as:

• you aren't clipping on input;

• you aren't clipping in plug-ins;

• you are managing the levels into analogue-modelled gear effectively;

• you leave the master fader at unity (because its insert points are post-fader); and

• you aren't clipping going back into the analogue domain from your DA converter.

So how do you run similar levels in a DAW compared to the analogue domain? Turn it down! Every channel in your DAW should be running at peak values between −6 dBFS and −12 dBFS. Where you set the peak values in this range also depends on the nature of the instrument, which we will get to soon. If in any doubt about how hot to record an instrument, always lean towards recording at the lower end of this spectrum as you can always boost it in the digital domain later. Even more experienced engineers can get confused with recording levels. I often see professional engineers recording all of their tracks at −18 dBFS into their DAW. While this is not going to do any harm, it is also not likely to get the best out of any analogue equipment prior to the AD conversion. This is because the recording level is likely to be lower than what you would set in the analogue domain.

So why have professional engineers made this 'mistake'? To answer this, we need to take one more look at comparisons between metering scales. When engineers started

to integrate both analogue and digital processes in hybrid setups, manufacturers suggested recommended digital levels for engineers to calibrate their analogue VU meters to. In Britain, the calibration process involves sending an uninterrupted sine wave test tone at 1 KHz out of their system at −18 dBFS and setting the analogue equipment's meter to equal 0 VU (in the United States, it has been recommended at −20 dBFS).

This calibration process caused much confusion with engineers when it came to implementing recording levels. What engineers extrapolated from the calibration process was that they should record everything into the DAW at −18 dBFS too as this would equal 0 VU. This is wrong! Comparing 0 VU and −18 dBFS is like comparing apples to oranges, as VU is an averaging level and dBFS is a peak level. The level calibration process is based on a continuous sine wave and, because a sine wave is a pure tone consisting of only one frequency and the tone never stops, the calibration tone has an identical dBFS peak and RMS level. In practice, the RMS and peak levels of instruments won't be equal – the disparity between them will be based on the transient nature and decay of the instrument. Your decision on where you place the peak level between −6 and −12 dBFS should be based on:

- the transient nature of the instrument; and
- the density of the mix.

For instance, a snappy snare drum transient with a short decay is going to have a higher peak value but a much lower RMS level. On the other hand, a heavily distorted electric guitar is likely to have a much smaller disparity between its peak and RMS level. In practice, this means that you are likely to want to have the snare drum peaking a little higher in level so that it feels more balanced with the guitar without having to adjust your faders. If your track's layering is going to be very dense with keyboards and guitar parts, I often give these instruments a little less level on input so that I am not running into problems overloading plug-ins later when it comes to the mixing process.

When recording in the digital domain, we need to see any clips in real time, so monitoring peak levels is still pivotally important, but for balancing purposes you also need to be aware of the RMS signal in the same way as you would in the analogue domain. You can use an RMS meter if you wish, but over time you will get used to setting recording levels by only looking at a peak meter and by balancing instruments.

HYBRID SYSTEM CALIBRATION

If you are working in a hybrid system, where you have analogue equipment (console, analogue summing mixer, or other hardware) along with your DAW, it is important to use the calibration method mentioned above. However, you can only do it this way if your DACs have an output level trim, which is considered a premium function in

converters, and generally only the most expensive will have that option.

If you don't have this option available to you, the simplest and easiest way is to consult your interface or converter's technical specs. For instance, I own a Prism Sound Orpheus, and after looking at the specifications I can see that the analogue inputs and outputs are +18 dBu at 0 dBFS. Therefore, to output at line level, I need the signal being sent to my analogue hardware to be at −14 dBFS RMS in my DAW.

This means that I have 14 dB of headroom from line level to peak. This headroom is less than on some of the flagship converters, which are usually set between −14 dBFS and −18 dBFS. So on the Orpheus, we need an output signal level of −14 dB RMS to the analogue gear to equal 0 VU. Ideally, we would like to work at the UK or US standard levels, which are −18 dB RMS for the UK and −20 dB RMS for the US.

In order to do this, we would need to adjust the interface's settings to make the system reach line level at −18 dBFS instead of −14 dBFS. Unfortunately, this is a rather premium feature for a DAC and is not offered on the Orpheus. So I just have to remember when mixing that I have to output at −14 dBFS RMS to outboard equipment. In practical terms, this means I have to be slightly more careful with peak values in the DAW and channels may require more compression or limiting to avoid clipping while retaining an RMS of −14 dBFS. This is because peak values will be proportionally closer to 0 dBFS with a system calibrated to be line level at −14 dBFS than one calibrated to be −18 dBFS or −20 dBFS. If you cannot get hold of your interface's tech spec or it doesn't mention the output sensitivity, then you can use a digital multi-meter (DMM) to work it out. All we need to know is that 0 VU = 1.227 AC volts RMS on the output of the DAC:

1. Insert the black patch lead that comes with the DMM into the COM port, and the red patch lead into the VΩ port, then set the multi-meter to read AC voltage (~).

2. Set up an auxiliary channel to send a 1 KHz sine wave tone out of your specified output at −25 dBFS RMS.

3. Insert a lead into the chosen output and on the other side of the cable place the multi-meter's probes on the + and − connectors. These are pins 2 and 3 of an XLR connector or the tip and ring of a TRS connector.

4. Keep gradually increasing the output level of the sine wave until your DMM reads 1.227 AC volts RMS (if it only displays two significant figures, use 1.23).

5. Once it reads 1.227 AC volts RMS, read the level in dBFS that you are sending out of the output. This level will be 0 VU for your audio interface/converter.

Table 7.1
0 dBFS = +18 dBu
−1 dBFS = +17 dBu
−2 dBFS = +16 dBu
−3 dBFS = +15 dBu
−4 dBFS = +14 dBu
−5 dBFS = +13 dBu
−6 dBFS = +12 dBu
−7 dBFS = +11 dBu
−8 dBFS = +10 dBu
−9 dBFS = +9 dBu
−10 dBFS = +8 dBu
−11 dBFS = +7 dBu
−12 dBFS = +6 dBu

Once you know how your system is calibrated, you should have no problem running some of your key instrument faders in the DAW at near unity, and even after summing it should also leave plenty of headroom on the master fader.

CONSIDERATIONS OF WORKING WITH HYBRID/ANALOGUE SYSTEMS

Even those without experience recording with tape in a fully analogue system will probably be aware of the argument that analogue processes sound warmer and subjectively better than digital ones. Such is the reverence of analogue equipment that the world's top plug-in manufacturers are making emulations of the very best tape machines, consoles, EQs, compressors and anything analogue they can get their hands on!

But why is this happening? Digital recording is a definite improvement on analogue systems in terms of data storage, portability, ease of use, editing facilities, flexibility and workflow, but analogue systems had innate faults and tonal colour caused by the altering of the harmonic structure that can subjectively make the end result more pleasing to the ear. These faults, including the distortion introduced by overdriving tubes, tape saturation, the top-end loss in tape machines, tonal colour of classic hardware and countless other small imperfections, can help bring a sense of cohesion and 'glue' to a multitrack recording.

This does not mean that analogue is better than digital. In fact, unless you have the budget for the top-level consoles or hardware and the maintenance budget to match, it usually isn't. I am not saying there aren't bargains out there, but try equipment out before you buy it. In my experience, on a smaller budget you get comparable (if not better) results by opting to buy plug-ins that emulate top-end analogue technology rather than buying cheaper hardware.

Working digitally is not an excuse for why you are not getting warm, punchy and pleasing results while recording and mixing entirely in the digital domain, or 'in the box' (ITB). There are plenty of techniques you can use to make the result of a digital recording sound more 'analogue'. Many top-level mixers such as Tchad Blake, Andrew Scheps and Dave Pensado have mixed commercial grade results entirely ITB. I'd go so far as to say that the difference you hear in commercial recordings is actually 90 per cent the skill of the engineer, the quality of the rooms used to record/mix in, the quality of performers and instrument build, and only 10 per cent is the 'sheen' added by the top-of-the-line pro-audio equipment.

If you do have the budget to buy analogue hardware or even want to set up a classic analogue-only system based on a tape machine, you will need to think a little bit more about signal-to-noise ratio. Fully analogue recording processes were notorious for developing electronic hardware-based SNR issues. A few standard operational procedures help to reduce the causes of a low SNR:

• Use amplifier volumes and mic positioning to maximize level without having to crank the pre-amp's gain.

• Minimize the amount of equipment that the audio passes through. Only make the recording chain as long as it needs to be. This includes tape-bouncing only when necessary.

• Avoid reducing gain in one processor only to increase it again in another.

• Regularly maintain analogue audio gear.

MONITORING LEVELS AND THE K-SYSTEM

Now that we have learnt more about metering, we can start to calibrate monitoring levels in a more

fail-safe way.

In the music industry, it has become standard practice for producers to set their own monitoring levels. In the film and television industry, however, significant standards have been set to help unify the levels of audio. This level is calibrated so that 0 VU equals 83 dB SPL.

Renowned mastering engineer Bob Katz sought to change this with the KSystem, which involves a three-tiered monitoring gain and metering system, which you change dependent on application.(5) While the K-System is not universally implemented by audio professionals, it is a great starting point for the inexperienced to make sure they are monitoring correctly. It is designed to:

1. Allow you to easily return to a calibrated listening level that is a compromise between the point at which your ears have the flattest frequency response (Fletcher-Munson), and also a comfortable listening level that can be used for extended periods of time without ear fatigue.

2. Optimize headroom in your mixes and to make sure that dynamic range is not excessively reduced, which can result in loss of excitement and impact. The three tiers are called K-20, K-14 and K-12. The proposed applications for each of these are as follows:

• K-20 is for music with a high dynamic range, such as the large theatre film mixes, classical music, etc. (20 dB of headroom).

• K-14 is for the mixing of standard pop-rock music, etc. (14 dB of headroom).

• K-12 is for programmes dedicated for TV or radio broadcast (12 dB of headroom).

To properly integrate the K-System into your workflow, you need to be aware of three factors:

1. You must use the standardized monitor level outlined by the K-System. This states that each monitor should be set to a level of 83 dB with an SPL meter, which should be set to C weighting with a slow attack.

2. You must use the K-Meter on the mix bus to ensure the sections of your track are falling within the required levels suggested by the K-System. Luckily, many of the most popular limiters and spectrum analysers come with K-Metering.

3. The K-System was developed to help govern the overall dynamic range of a mix. The K-Meter was only designed to meter an entire mix, and therefore is not relevant when determining correct levels on individual tracks within a mix (proper gain staging determines this).

Monitor Calibration for the K-System

Here are the steps to correctly set up your monitors for the K-System:

1. Turn your sound card's output to 0 dBFS (and any monitoring controller you might have).

2. Output pink noise from your DAW at −20 dBFS RMS.

3. With the pink noise outputting to your monitors, each speaker needs to be turned up to 83 dBSPL using an SPL meter at the listening position. You need to make sure that you do each speaker separately, and that only the speaker you are setting is turned on. You also need to make sure that the SPL meter is set to C weighting (which has a flatter bass response) and that the attack time is set to slow.

Now that you have done this, you can operate at K-20. If we then choose to operate at K-14, our unity point (0) is 6dB higher than when we were operating in K-20. To maintain our ideal monitoring level of 83 dB, we need to trim 6 dB from our monitor output. To operate at K-12, we need to trim 8 dB.

Using the K-Meter

To implement the K-System, you need to insert a plug-in that has a K-Metering feature on the last insert slot of the mix bus. Both PSP's Xenon and FabFilter's Pro-L brickwall limiters have this capability. To work out how to use the meter, let's examine the K-14 scale some more.

The K-meter is split into three operating ranges, the green, amber and red zones. Your aim is to operate in the green zone (less than 0) during the quieter sections of a song, the amber zone (0 to +4) for louder sections such as choruses, and the red zone (+4 and above) for the loudest sections (typically last choruses/solos/outros). Values over +8 should be seldom used, as this starts to leave the mastering engineer too little headroom to work with. These limits are guidelines and should be based on personal taste and the style of song. For instance, a particularly muscular heavy metal track may be in the red zone more often than a pop track. Similarly, a stripped-down acoustic track may always operate in the green or amber zones.

NOTES

1. The reason most ITB mixes don't sound as good as analog mixes, accessed January 2013,

www.gearslutz.com/board/so-much-gear-so-little-time/463010-reason-most-itb-mixesdona-

t-sound-good-analog-mixes-restored.html.

2. Hugh Robjohns, What is 'aliasing' and what's the cause of it?, accessed January 2013,

www.soundonsound.com/sos/jan06/articles/qa0106_2.htm.

3. Shannon-Nyquist sampling theorum, accessed January 2013, http://en.wikipedia.org/

wiki/Nyquist%E2%80%93Shannon_sampling_theorem.

4. Dan Lavry, The optimal sample rate for quality audio, white paper accessed January 2013,

www.lavryengineering.com/pdfs/lavry-white-paper-the_optimal_sample_rate_for_

quality_audio.pdf.

5. Bob Katz, Mastering Audio: The Art and the Science (Focal Press, 2007).