

“Mixed” Results

An Introduction to Analyzing Music Production through Eight Commissioned Metal Mixes

Jan-Peter Herbst, Eric Smialek

Dieser Artikel führt Musiktheoretiker:innen und Analytiker:innen in Aspekte der Tontechnik und Produktion ein, die innerhalb der populären Musik weitgehend übertragbar sind. Wir reflektieren, warum die Analyse von Klang und seiner Produktion nützlich ist, gefolgt von einer Fallstudie zu »In Solitude«, einem original zu diesem Zweck geschriebenen und produzierten fünfminütigen Lied, das von acht der weltweit führenden Metal-Produzenten verschiedener Generationen abgemischt wurde. In dieser Fallstudie zeigen wir, wie Analytiker:innen Aspekte der Lautstärke grob auf einer virtuellen Klangbühne abbilden können, die Hörer:innen für subtile, aber folgenreiche Produktionsnuancen sensibilisiert. Um zu erklären, wie dieser virtuelle Raum funktioniert, führen wir die Leser:innen durch psychoakustische Beziehungen von Breite, Höhe und Tiefe. Anhand dieser Konzepte und ihrer Beziehungen zur (Psycho-)Akustik zeigen wir, wie die teilnehmenden Produzenten Hörereindrücke einer virtuellen Klanglandschaft unterschiedlich manipulieren. Diese Produktionsentscheidungen spiegeln nicht nur Umgebungseindrücke wider, sondern reflektieren breitere ästhetische Trends im Zusammenhang mit Subgenre-Erwartungen im Metal. Wir hoffen, dass diese Fallstudie als Grundlage dient, um Analytiker:innen populärer Musik zu helfen, Tontechnik in ihre Analysen einzubeziehen. Solche Arbeit muss das Rad nicht radikal neu erfinden; sie kann Produktionsqualitäten in traditionelle Beobachtungen zu Harmonie, Melodie, Rhythmus und Form integrieren. Wir argumentieren, dass die klanglichen Nuancen der Produzenten die Normen und Erwartungen einzelner Subgenres, wie sie auf der Aufnahme – dem maßgeblichen Text innerhalb der populären Musik – existieren, widerspiegeln und prägen. Dementsprechend könnten Analytiker:innen fragen, wie diese Normen durch Produktionsentscheidungen ebenso wie durch Elemente des Songwritings bedient oder durchbrochen werden.

This article introduces music theorists and analysts to aspects of audio engineering and production that are widely transferrable within popular music. We reflect on why analyzing sound and its production is useful, followed by a case study of “In Solitude,” an originally written and produced five-minute song that was mixed by eight of the world’s leading metal music producers spanning several generations. In this case study, we demonstrate how analysts can roughly map aspects of loudness onto a virtual soundstage that helps sensitize listeners to subtle, but consequential, production nuances. To explain how this virtual space works, we guide readers through psychoacoustic relationships of width, height, and depth. Using these concepts and their relationships to (psycho-)acoustics, we show how the participating producers variously manipulate listener impressions of a virtual soundscape. Beyond environmental impressions, these production decisions reflect broader aesthetic trends related to subgenre expectations in metal. We hope this case study serves as a primer in helping popular music analysts include audio engineering in their analyses. Such work need not radically reinvent the wheel; it can integrate production qualities into traditional observations such as harmony, melody, rhythm, and form. We argue that the producers’ timbral nuances reflect and inform the norms and expectations of individual subgenres as they exist on the record, the definitive text within popular music. Accordingly, analysts might ask how those norms are served or thwarted by production choices as much as songwriting elements.

SCHLAGWORTE/KEYWORDS: Analyse; analysis; Metal; music production; Musikproduktion; sound engineering; Tontechnik

Any discussion of the role of technology in popular music should begin with a simple premise: without electronic technology, popular music in the twenty-first century is unthinkable.

—Paul Théberge¹

1. INTRODUCTION

As popular music analysis coalesced as a field, particularly during the 1990s, authors such as Allan F. Moore and Paul Théberge increasingly emphasized the significance of technological mediation and produced sound – developments that have been recognized as consequential since Walter Benjamin’s famous essay on “The Work of Art in the Age of Mechanical Reproduction.”² As Théberge’s quote in our epigraph implies, technology has shaped the very artistry and aesthetics of the past 100 years to the point where an awareness of what happens behind the curtain of the audio engineer’s workshop can fundamentally alter one’s music analysis.

In 2006, the *Journal on the Art of Record Production*³ was founded, which created an innovative space for analyzing popular music recordings with practical insights and a deep understanding of recording, mixing, and production. Foundational works, such as *The Art of Record Production* volumes 1 and 2,⁴ as well as Routledge’s book series “Perspectives on Music Production,” address how practices and decisions in music production shape listening experience. In parallel but independent developments, works by Moore⁵ and William Moylan⁶ and colleagues⁷ introduced inventive methods for visualizing spatial relationships of sound. Albin Zak’s *The Poetics of Rock*⁸ offered a pivotal glimpse into the ethnography of record production, showing us “from the inside,” as Robert Walser’s review put it, “how complex and multifarious this kind of composition is [...] on every page.”⁹ Here, the kind of composition to which Walser refers is the combined efforts of not just songwriting and arrangement but also their refinement by producers to become the “tracks” we hear.¹⁰

Metal music’s distinct aesthetics of heaviness, transgression, and extremity provide an ideal context for the study of music production. While the exigencies of metal music production are well documented,¹¹ the genre’s exaggerated sound qualities invite an exploration of hyperreal mixing practices and their effects on interpretation and meaning. Metal’s aesthetic of intensification makes production choices particularly salient, inviting

1 Théberge 2001, 3.

2 Moore 1993; Théberge 1997; Benjamin [1935] 1969.

3 <https://www.arpjournal.com/asarpwp> (4 Jul 2024).

4 Frith/Zagorski-Thomas 2012; Zagorski-Thomas et al. 2020.

5 Moore 1993, 2012.

6 Moylan 2002; see also Moylan 2020.

7 Moylan et al. 2023.

8 Zak 2001.

9 Walser 2007, 256–257.

10 Zak 2001, 24.

11 Herbst/Mynett 2022a, 2022b.

analysis of how producers achieve sounds and effects that exceed or transcend the everyday sound environment through their mixing strategies, a phenomenon relevant to contemporary music production more broadly. While heaviness is defined here as a central aesthetic criterion in metal – resulting from a complex interplay of sonic attributes, arrangement choices, and performance qualities¹² – this article uses it as a case study to explore the broader impact of audio engineering on musical meaning, demonstrating how production decisions reshape the listener’s interpretive experience.

As a combined study of record production and popular music analysis intended to introduce music theorists and analysts to aspects of audio engineering and production that are widely transferrable within popular music, our article offers a case study of “In Solitude,” a five-minute metal song originally written and produced by Mark Mynett and Jan-Peter Herbst for the “Heaviness in Metal Music Production (HiMMP)” research project.¹³ The research consisted of a three-component design: 1) writing, arranging, and recording the song; 2) creating multiple mixes by professional record producers; and 3) documenting the process and the underlying conceptions ethnographically through video interviews. The first goal was to offer an intuitive working environment with a track that would sound convincing as a “real” song. During the mixing stage, each of the eight participating producers received the reference mix (see Audio Example 1) but was instructed to actively make creative decisions. For instance, producers selected different guitar and bass source tones by “re-amping” with other hardware or digital amplifiers using direct injection (DI) recordings (i.e., the raw, non-amplified signals from the instruments’ outputs). They could also enhance or replace drum sounds with samples, i.e., layering the naturally recorded drum shells (e.g., kick or snare) with sounds from other drum kits or replacing the recorded material altogether to significantly alter the drum sound. Their choices could then be compared to their interview statements about production during the documentation stage, particularly regarding aesthetic preferences and guiding philosophies. Following Zak’s lead, we use producer interviews to gain direct insights into their working process, at times connecting the statements of eight participating producers with examples that highlight each producer’s individual style. Our analytical approach is two-pronged: we conducted an in-depth qualitative analysis of semi-structured interviews using both emic and etic frameworks, and we performed a comparative technical and auditory analysis of the recorded mixes to map the interplay between production decisions and listener perception. This dual analysis not only examines the technical intricacies of mixing but also lays the foundation for exploring how these decisions can shape listener interpretation.

This research employs a novel, quasi-experimental ethnographic design where eight professional producers mixed the same song independently in their own studios, thereby faithfully capturing their engineering practices within ecologically valid professional environments. The HiMMP team (Herbst and Mynett) conducted video-recorded interviews in the producers’ studios shortly afterward to capture their fresh reflections. From 2020 to 2024, fifteen hours of semi-structured interviews documented technical processes and producer conceptions of “heaviness.”¹⁴ These interviews consistently addressed key top-

12 Ibid.

13 Further information and a list of the project’s publications are available on the project website <https://www.himmp.net>.


14 All videos are available at <https://www.youtube.com/@HiMMP-Research>.

ics, such as dynamic range, stereo imaging, and aesthetic intent, while allowing for unexpected insights.¹⁵ While the quasi-experimental setting involves some differences from commercial environments, our combined emic and etic approach provides a robust basis for comparing mixing practices.

We selected producers based on professional standing, coverage of historically important regions for metal (UK, USA, and Scandinavia), and diverse experience across metal subgenres. Participating producers included Jens Bogren (b. 1979, SWE), Mike Exeter (b. 1967, UK), Adam “Nolly” Getgood (b. 1987, UK), Josh Middleton (b. 1985, UK), Fredrik Nordström (b. 1967, SWE), Buster Odeholm (b. 1992, SWE), Dave Otero (b. 1981, US), and Andrew Scheps (b. 1969, US).

Regrettably, a gender balance among the participants could not be achieved. This limitation arises not only from logistical constraints but also from the broader underrepresentation of women and non-binary producers, especially at the highest international level.¹⁶ The sample reflects the current state of hegemonically dominant metal music production, centered around the Global North. Consequently, the findings should be understood as an invitation to learn from widely known representatives of metal while acknowledging that metal music production is a global phenomenon, with many producers of diverse genders, ethnicities, ages, and regions equally deserving of recognition.

While previous studies¹⁷ have used retrospective interviews or single-case analyses of producer methodologies, to our knowledge, no study has directly commissioned multiple producers to mix the same recorded material for a controlled, comparative analysis. This innovative approach fills a significant gap in the literature.¹⁸ By directly comparing the mixes of “In Solitude” produced independently by eight renowned metal producers, our study offers unprecedented insights into the diverse strategies and tacit knowledge that underlie contemporary metal music production. This controlled comparison isolates production variables and offers robust empirical evidence on how mixing decisions shape heaviness in metal music.

 https://storage.gmth.de/zgmth/media/1222/Herbst_MixedResults_Audio01.mp3

Audio Example 1: Reference mix of “In Solitude”

One of the innovation goals of the data¹⁹ is that this situation would not ordinarily arise in the music industry – few labels would hire multiple engineers to create different versions of the same song. The resulting resource is thus unparalleled even among dedicated learning offerings such as “Mix with the Masters” or “Nail the Mix.” Accordingly, YouTube footage²⁰ comparing the mixes has been met with curiosity and amazement. Some comments express relief that there is no single correct way to produce a song: different respected producers arrive at distinct results with individual qualities, and each sounds professional. At times, the comments contain puzzling contradictions – making declarations about authenticity yet favoring mixes that counter them – reflecting how aesthetic

15 See Herbst/Mynett 2025c.

16 See also Herbst/Mynett 2021.

17 For example, Turner 2009, 2012; Mynett 2017.

18 See Thomas 2015, 259.

19 All data is freely accessible at <https://huddersfield.app.box.com/s/8gren2ma4kesvf5vwip2axzz1v8sawur>.

20 https://www.youtube.com/watch?v=U_TtJo2_bh8 (8 Aug 2024).

beliefs and assumptions sometimes clash with empirical insights into the engineering process.²¹ Overall, the dataset is valuable for studying metal music, heaviness, and for broader reflections on popular music analysis. Its value is especially evident from the following highlights:

1. nine different versions of the same track are uncommonly available for comparison;
2. audio stems for drums, bass, guitars, and vocals are available from all versions, allowing for detailed analysis and comparison;
3. unlike traditional analysis of released music, the produced result can be compared with the underlying audio material ('multi-track' recording; see Figure 1);
4. composing an original song allows audio examples to be freely distributed without copyright restrictions.

Building from ethnographic and analytical evidence, our study makes arguments about the emerging differences between how groups of producers construct a virtual space within their music production practices.

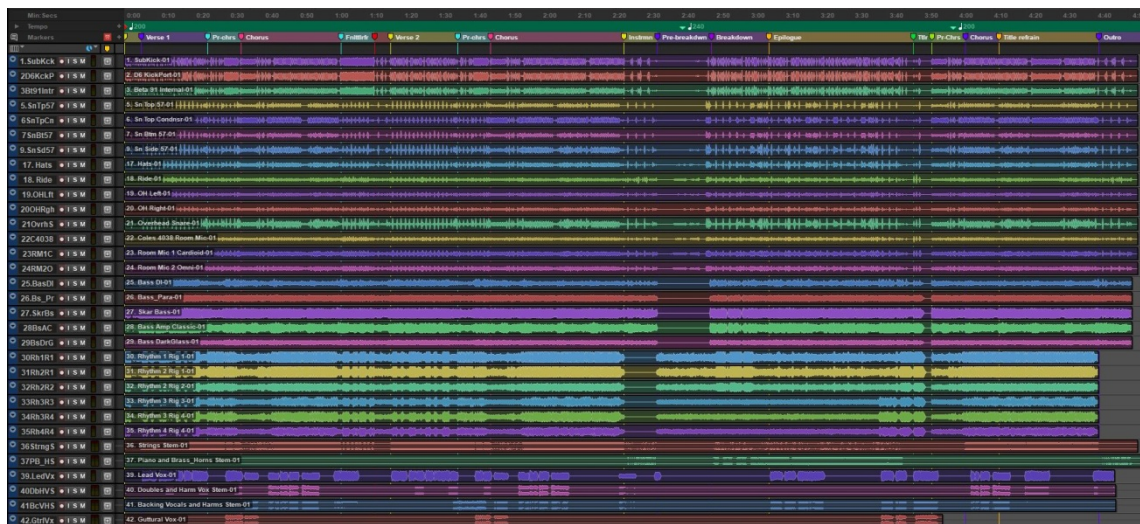


Figure 1: Arrangement view in Pro Tools; all tracks except toms and production effects

Music production allows artists to use sonic space as an artistic parameter in ways unavailable in the non-digital world. Overlapping with recent trends in electronic styles of popular music, such as like EDM (Electronic Dance Music) and hyperpop, metal producers have increasingly attempted to transcend the familiarity of acoustic soundscapes. Whereas acoustic instrumentation – unamplified folk ensembles, classical orchestras, choirs, and pianos – mostly depends on the natural acoustics of a performance space, heavy metal has long employed techniques that deliberately depart from such realism. For instance, early heavy metal did not necessarily aim to reproduce a live venue's acoustic environment. Instead, it often embraced deliberately contrived sonic effects, such as isolating individual drum parts, applying noise gates to distorted guitars, and employing pronounced stereo-width expansion. These techniques produce a hyperreal or even surreal aesthetic. By contrast, some producers strive to foster impressions of live fidelity by mimicking the natural acoustics of a performance space. Such contrasting

21 See Herbst/Mynett 2025a.

strategies and techniques suggest that “live fidelity” in heavy metal is not an inherent property of the sound but rather a specific artistic choice associated with heaviness, transgression, extremity, virtuosity, and authenticity.

As we will argue, these contrasting aesthetics loosely map onto two schools of metal production – old-school live fidelity and new-school hyperreality (Figure 2) – with intriguing complications and overlaps along the way. These contrasting approaches to mixing serve as operational concepts for our analysis and represent endpoints of a continuum rather than discrete categories. Although these distinctions are widely discussed informally within fan and musician communities, to our knowledge, they have not been systematically examined in scholarly research, and our study offers preliminary empirical evidence to substantiate and clarify these concepts. While previous writings do not explicitly frame this dualism, Niall Thomas and Andrew King provide comparative perspectives on older and newer metal production practices, and Mynett highlights that metal often departs from a live-fidelity aesthetic.²² Together, these texts support our emerging framework and highlight the need for a systematic investigation of these contrasting approaches in metal production.

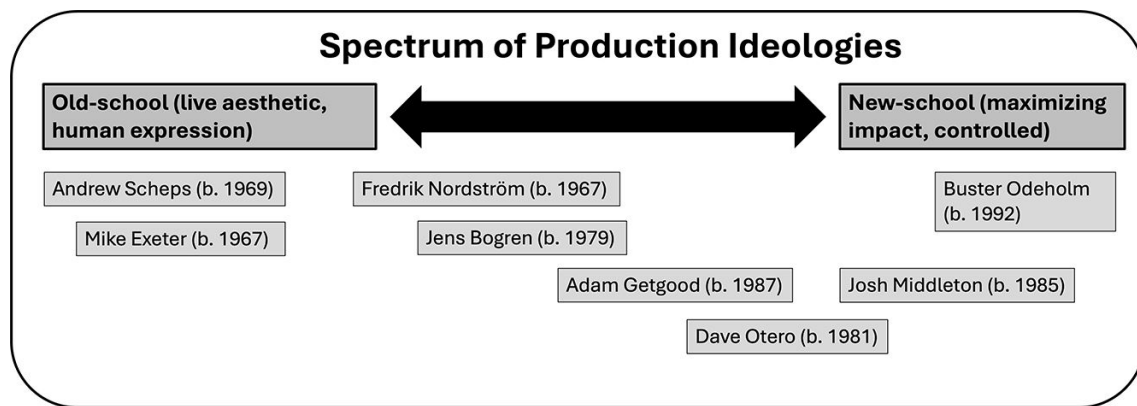


Figure 2: A conceptual spectrum of metal production ideologies, from live fidelity to hyperreality²³

Among the novel contributions of this study is our demonstration of how producers reshape the recorded sound through deliberate mixing decisions. In particular, we introduce and clarify several concepts that have not been systematically examined or connected to musical examples: the meta-instrument, hyperreality in production, cohesion among sound sources, and the V-shape mix. These operational concepts not only illuminate new dimensions of production in metal music but also challenge traditional notions of authenticity and live fidelity by revealing the creative strategies that producers employ to transform the recorded performance into a fully produced product. Using the aforementioned methodological and analytical frameworks, we further demonstrate how analysts can map aspects of loudness onto a virtual soundstage that helps sensitize listeners to subtle, but consequential, production nuances. To explain how this virtual space works, we guide readers through psychoacoustic relationships of width, height, and depth. Through these concepts and their relationships to acoustics, we show how the participating producers variously manipulate listener impressions of a virtual soundscape.

²² Thomas/King 2019; Mynett 2020.

²³ Reproduced from Herbst/Mynett 2025b, chapter 6.

Not merely limited to environmental impressions, these production decisions reflect broader aesthetic trends related to subgenre expectations in metal.

We proceed in three main parts. Our next section directly addresses the methodological and disciplinary question of why analyzing production is important for music theory. We then move to an examination of loudness that establishes the broad division of new- and old-school approaches to metal music production. Finally, our last section explores a virtual space through the three-dimensional metaphors of depth, width, and height. Throughout, we combine and apply a broad array of concepts from music analysis, psychoacoustics, and production techniques to examine how our eight producers shape these dimensions in their mixes.

2. WHY ANALYZE SOUND AND ITS PRODUCTION?

The growing impact of production techniques on the aesthetics of popular music, detailed in historiographies of popular music recording and production,²⁴ makes sound production increasingly unavoidable when analyzing popular music. We concur with Walser that “any cultural analysis of popular music that leaves out musical sound, that doesn’t explain why people are drawn to certain sounds specifically and not others, is at least fundamentally incomplete.”²⁵ Several points are important in this regard.

Firstly, as Mads Walther-Hansen²⁶ has argued, sound perception is not arbitrary. Cognitive metaphors and embodied experiences are widely shared within and sometimes across cultures, reducing the number and breadth of likely interpretations.

Secondly, the analysis of music production is rarely confined to objective, quantifiable observations of sound properties. The “track,” as Zak²⁷ defines it, combines song (representable on a lead sheet), arrangement (visible in a full score), and produced recording (playable through speakers). But rather than a one-way production line, the relationship of song-arrangement-production is better thought of as a feedback loop. That is, the requirements of recording and the capabilities of production technology now play an increasingly influential role in determining the arrangement. They can restrict certain arrangement choices or afford hyperreal sounds beyond the capabilities of a band or ensemble. Paying close attention to the requirements and affordances of the recorded medium allows us to better understand specific artistic choices and their likely interpretations. As Moylan summarizes, “[t]he record is a finely crafted performance; it reframes notions of performance, composition, and arranging around a production process of added sonic dimensions and qualities.”²⁸

Thirdly, and closely related to the previous point, the recording is popular music’s “autographic”²⁹ (i.e., definitive) text, so most analytical contexts will be impacted by music production. Sound engineering impacts some of the most debated issues in popular mu-

24 Chanan 1995; Cunningham 1996; Moorefield 2010.

25 Walser 2003, 21–22.

26 Walther-Hansen 2020.

27 Zak 2001, 24.

28 Moylan 2020, 9.

29 Gracyk 1996, 31–36.

sic. It influences the artistic expression of identity and authenticity during production as well as an audience's mediated experience of them during listening.

Fourthly, timbral nuances – performed with the assistance of music technology and refined during production – shape and expand musical expression. Take guitar tone, for instance. Among the many facets of its appeal, distortion facilitates playing techniques uncommon on the unamplified guitar, including feedback, artificial harmonics, and two-handed tapping.³⁰ Beyond practical advantages, guitar distortion demonstrates how electronic mediation can transform the cultural significance of instruments and sounds – in this case, into a symbol of power and desire.³¹

In studio-oriented music, where sounds can be processed more effectively than in a live performance context, the transformation and extension of acoustic sounds goes one step further. Auto-Tune and Melodyne are pitch modulation tools that have created recognizable vocal sounds important in rap, pop, and electronic dance music. In these genres, and others, velocities (i.e., programmed dynamics), modulation filters of the frequency spectrum, dynamic effects such as pumping (through side-chain compression), and immersive spatial and modulation effects provide many of the expressive qualities most recognizable to audiences. With both technologically mediated live performance and pre-programmed music, the sounds are intentionally designed in detail and can even influence artists to recompose their compositions during the recording and production stages. Quite passionately, Moylan encapsulates both the pervasiveness of record production in the artistry of the final product as well as its innovative abilities to create once inconceivable sounds:

[T]he record is not only a performance and a composition, but something else as well. It is also a set of sound qualities and sound relationships that [...] do not exist in nature and that are the result of the recording process. These qualities and relationships create a platform for the recorded song, and contribute to its artistry and voice. [...] The result is a reshaped sonic landscape, where the relationships of instruments defy physics, where unnatural qualities, proportions, locations and expression[s] become accepted as part of the recorded song, of its performance, of its invented space.³²

Produced sound transforms the arrangement, performances, and aesthetic treatments into a cohesive work with unique qualities that exceed the acoustical limitations of natural space.³³ Zak expands:

The performances that we hear on records present a complex collection of elements. Musical syntactical elements such as pitches and rhythms are augmented by specific inflections and articulations, which include particularities of timbre, phrasing, intonation, and so forth. Furthermore, the inscription process captures the traces of emotion, psyche, and life experience expressed by performers. That is, the *passion* of the musical utterance is yet another element of a record's identity.³⁴

30 Walser 1993, 70.

31 Waksman 1999.

32 Moylan 2020, 10–11.

33 Zagorski-Thomas 2014, 8.

34 Zak 2001, 51, emphasis in original.

All these elements exist within a recording, but how intentional or coincidental they are, as Moylan³⁵ suggests, is not necessarily of concern; the produced record is a unique musical artifact that is unlikely to exist in the same way in the “real” acoustical world, nor is it typically possible to perform identically.

Whether deliberate or not, produced music inevitably contains spatial staging, the three-dimensional placement of instruments, voices, and sounds within a listener’s perceptual imagination. It can be used to digitally simulate spatial relationships between musical objects and influence interpretations of song narratives and performer personae.³⁶ This alternative reality of staged sound simultaneously complicates a recording – adding hyperreal simulacra – and simplifies it, subtracting unwanted features. Unnecessary or detrimental aspects of sound (e.g., noise or excessive frequency energy) get removed, while those more useful or meaningful are emphasized in production. These decisions are what Simon Zagorski-Thomas calls “sonic cartoons,”³⁷ simplified, schematic versions of reality, which occur in most, if not all, popular music recordings. Their purpose is to foreground specific sonic characteristics such as heightened clarity, dynamic impact, and a refined spatial presence. These hyperreal qualities guide the listener toward an aesthetic interpretation focused on intended stylistic and emotional cues. Shaping the representation of a performance or programmed gesture thus “encourages a particular type of interpretation,”³⁸ namely one that privileges these targeted sonic features over the full complexity of the live event. Overall, then, the history of recorded popular music is one of improved “high-fidelity,” the detailed reproduction of acoustical or environmental sounds, but done with the aesthetic goals of augmenting or refocusing meaning away from physical sound sources. With so much craftsmanship focused on carefully sculpting precise sounds, neglecting sound when analyzing popular music risks losing a significant part of the artistry behind making music meaningful to listeners.

Beyond affecting overall sound quality, production decisions fundamentally alter the perceived spatial environment of a recording, an important dimension we observed across our eight producers’ mixes. We now turn our focus to how loudness and virtual space, conceptualized through dimensions of width, height, and depth, serve as models for understanding these effects.

3. LOUDNESS: NEW-SCHOOL HYPERREALITY VS. OLD-SCHOOL LIVE FIDELITY

To enable controlled comparisons between mixes, and prevent louder mixes from automatically seeming “better,” the HiMMP researchers used Scheps’s relatively quiet and uncompressed mix as a loudness-normalization reference. Loudness comparison involves measurements in LUFS (Loudness Units Relative to Full Scale), a perceptual unit that accounts for both the intensity and frequency distribution of sound (e.g., brightness levels). Streaming services such as Spotify use these measurements to achieve relatively uniform volume levels and combat the “loudness wars” prevalent in metal production.³⁹

35 Moylan 2020, 37.

36 Zagorski-Thomas 2014, 70–91; Moore/Schmidt/Dockwray 2011.

37 Zagorski-Thomas 2014, 49–69.

38 Ibid., 55.

39 Smialek 2016, 117–120.

The HiMMP producers delivered their mixes, representative of how their work would be presented to a real-world client such as a band assessing a mix engineer. These mixes typically include a degree of mastering to provide a more complete impression of the intended aesthetic. The loudness profiles in the next paragraph were measured on these quasi-mastered versions, ensuring that our comparisons reflect both the mixing and the subsequent processing choices. This mastering, although minimal, offers additional clues about the producers' aesthetic priorities regarding loudness and dynamic range, especially in the context of the loudness war.

Scheps's mix measures -10.4 LUFS, while the "hottest" mixes are closer to the clipping threshold of zero decibels in a digital system: Middleton (-7.6), Otero (-7.5), and Odeholm (-5.6).⁴⁰ Hot mixes like Odeholm's involve hyper-compression, which cannot be undone later, even if the volume is lowered. The more a producer applies dynamic range reduction – compression, limiting, and clipping – the more the short-term peaks (or transients – the brief, initial bursts of sound that give an instrument its attack) are removed from the signal. This means there is potentially less "punch," a descriptive term for a desirable signal envelope that emphasizes the attack and decay phases,⁴¹ and less timbral individuality.⁴² The more one compresses and distorts, the more sounds become uniform. Herein lies a trade-off: producers can sacrifice some punch to gain an impression of proximity, or "in-your-face" closeness, desirable in metal, resulting from the signal's brightness created by distortion.

A glance at the waveform diagrams in Figure 3 reveals differences in dynamic range between the mixes. The light blue portion of the signal represents perceived and average loudness, while the dark blue represents the signal's short-term peaks or brief transients such as percussive sounds. Comparing Scheps's signal with Odeholm's, we note that the light blue portion appears fairly similar. In fact, all the producers show similar amounts of light blue, which indicates that the perceived loudness is generally uniform. All the sound clips have been normalized to -10.4 LUFS. However, when comparing Scheps's and Odeholm's dark blue portions, stark differences in their levels of dynamic variation and clipping (i.e., the effects of high audio compression that lead to distortion) become apparent. Scheps's signal exhibits a series of spikes reflective of widely varying dynamic levels. During interviews, he stated his dislike for using compression altogether. Conversely, Odeholm's signal is the epitome of what producers affectionately call a "sausage": a signal that severely clips those spikes (i.e., turning a round waveform into a squared one), ensuring a constant dynamic level throughout (Audio Example 2.1). In his interviews, Odeholm spoke about his heavy use of compression and especially distortion to achieve sonic weight and density, intentionally removing transients. Consequently, the dynamic range in Odeholm's mix is very limited. It never reaches a dynamic range of 10 dB, whereas Scheps's mix achieves this dynamic range approximately 40% of the time (Audio Example 2.2). This stark contrast in dynamic range treatment exemplifies the fundamental philosophical difference between old-school and new-school approaches: preserving natural dynamics versus creating a hyper-controlled, maximized sound. Odeholm's squashed dynamic range (5 dB) is more typical of EDM (3–6 dB) while Scheps's 9 dB falls within a rock standard (8–10 dB) (Antares 2024). So the mixes differ in dynamic

40 Digital systems operate under the 0 dB (full scale) threshold. Anything above clips the audio signal.

41 Fenton/Lee 2019.

42 Herbst 2016, 120–134.

range but not in loudness following normalization. Yet even when LUFS units are equalized, brightness differences can make a mix seem louder or “hotter.” By increasing brightness via distortion, producers can gain a competitive edge.

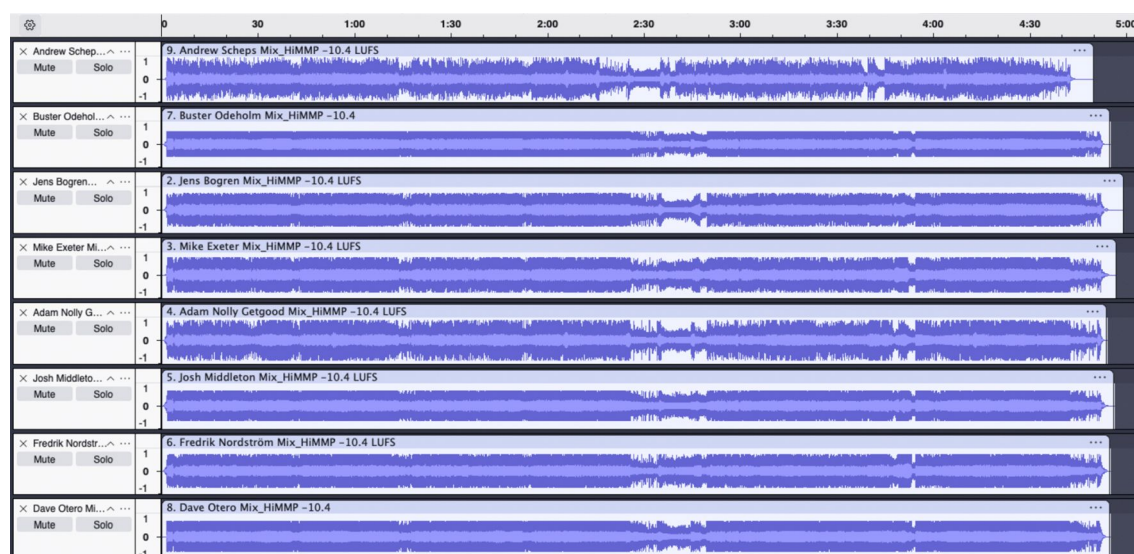


Figure 3: Waveforms of the final mixes of “In Solitude.” Light blue is average loudness, dark blue represents peaks (created with Audacity)

🔊 https://storage.gmth.de/zgmth/media/1222/Herbst_MixedResults_Audio02_1.mp3

Audio Example 2.1: Pre-chorus, Odeholm’s mix of “In Solitude”

🔊 https://storage.gmth.de/zgmth/media/1222/Herbst_MixedResults_Audio02_2.mp3

Audio Example 2.2: Pre-chorus, Scheps’s mix of “In Solitude”

Loudness is not a priority for everyone. One group in the HiMMP sample can be defined by their relative prioritization of dynamic liveness rather than uniform loudness. Exeter and Scheps, while differing in individual ways, both make production choices that retain a sense of fidelity to a live sound. Scheps avoids compression, drum samples, and triggers (devices that electronically capture and enhance drum hits), whereas Exeter uses these techniques to enhance production as long as they result in a believable live performance. By contrast, another group of producers, including Odeholm, Middleton, Getgood, and Otero, shows the opposite tendency. They freely embrace digital alterations to create a hyperreal impression of intensity beyond what is achievable in a live studio performance alone. Odeholm goes as far as stating,

I’m over the sound of a band – a guitar, bass, drums, vocals. That’s boring. I want to maximize everything to make it sound new, or make it sound different, or make it compete with electronic styles that have that full frequency range. Sounds huge, and I want the band sound to be that huge as well.⁴³

43 Information on the respective time and place of the interviews conducted as part of the “Heaviness in Metal Music Production” (HiMMP) research project can be found in the data section after the discography.

A hyperreal approach, such as Odeholm's, relies on "sonic cartoons,"⁴⁴ an approach that enhances and refocuses the musical qualities in a production, going beyond the conventional sense of high-fidelity. It moves away from representing musical performance towards its own ideal, where the carefully crafted production aesthetic becomes the autographic original. Live performance is no longer a primary consideration, as demonstrated by Odeholm and his band, *Humanity's Last Breath*, who recreate the album sound live using elaborate technological means. This includes using instrument tones taken directly from the record, automatically synchronized to the click track, plus an extensive playback of additional layers and effects.

In between these tendencies of live-fidelity and hyperreality is a group of compromises represented by Nordström and Bogren. Nordström (b. 1967) belongs to the same generation as Scheps (b. 1969) and Exeter (b. 1967) but has collaborated with contemporary bands such as Dimmu Borgir, Bring Me the Horizon, and Architects. His peers, Scheps and Exeter, are especially known for their work with older heavy metal bands such as Metallica (Scheps), Judas Priest (Exeter), and Black Sabbath (both). Exeter, in particular, developed an analog, live sound during a period when critics and fans used hard rock and heavy metal interchangeably.⁴⁵ Prior work with bands appears to impact a producer's approach and mindset significantly, particularly with influential bands. Referring to Black Sabbath's vocalist Ozzy Osbourne, Scheps noted:

When you're working with a band like Metallica or Black Sabbath, obviously, the history has to be part of how you mix. You're not going to *not* put a delay on Ozzy's vocal; he's never *not* had a delay. So, you have to do that. Finding the right delay for his vocal was a big part of those mixes, to the point where it actually made me think about slap delay more. And so now, slap delay is a big part of what I do with vocals in general. [...] So those bands have history in the sound.

Therefore, based on the clients of each producer, it makes sense that there are old-school, new-school, and in-between groups. These groups correlate roughly with age, although, as will become evident, the relationship is not clear-cut.

Bogren (b. 1979) approaches the age group of the hyperreal HiMMP producers, Otero (b. 1981), Middleton (b. 1985), Getgood (b. 1987), and Odeholm (b. 1992), but differs in his simultaneous interest in digital production enhancement and his ideal of human performance expression. In interviews, Bogren stated that technology can enhance musical expression, but he requires it to have a clear purpose. He seeks a balance between what he believes is best for the music, what the artist requests, what the record company expects, and what fans want.⁴⁶ While such a balance might sound reasonable for anyone, producers cater to different clients and audiences depending on subgenre. Furthermore, Bogren's mediating position between old-school and new-school approaches serves as a helpful reminder that these oppositions represent endpoints on a flexible continuum rather than discrete, mutually exclusive classifications.

44 Zagorski-Thomas 2014, 49–69.

45 Brackett 2014, 291.

46 Herbst/Bauerfeind 2021.

4. VIRTUAL SPACE

In addition to loudness, these schools differ in how they shape virtual listening space. Moore models such differences with the “sound-box,”⁴⁷ a three-dimensional space that translates subjective aural impressions of spatialization into imagined physical locations of width, depth, and height (Figure 4). Width refers to the lateral spread of sound, depth to the perception of distance, and height to the placement of sounds in terms of high and low pitch. Within this section, we will explore these dimensions in detail sequentially. As an entry point, we first introduce a two-dimensional visualization tool for analysis, the soundstage.

The soundstage follows Moylan’s concept of the “perceived performance environment,” which captures the measurable spatial attributes of the mix.⁴⁸ The resulting spatial representation maps psychoacoustic cues that create spatial impressions, regardless of whether listeners consciously interpret them as instruments on a physical stage. Listeners may also engage with the mix more abstractly, a possibility reinforced by the mixing ideal of the “meta-instrument,” discussed later in our section on high/low-frequency treatment, or height. Whether listeners visualize the traditional stage suggested by Moylan’s model or experience abstract spatial impressions, our framework accommodates diverse ways in which metal listeners construct meaning from the mix.

Figure 5 visualizes sound distribution in Scheps’s mix of the first chorus (Audio Example 3.1) using Moylan’s two-dimensional soundstage in relation to a listener.⁴⁹ When listening to Scheps’s production, one strains to hear the vocal harmonies as if emanating from the very back of a live stage. Accordingly, Figure 5 positions them at the back of the stage, using dotted lines to indicate that one may not detect them at all. By contrast, the bass guitar figures prominently near the front, along with a similarly salient snare drum and lead vocals. As one moves backward, other instruments in Scheps’s mix draw more from the width effects of panning and so appear towards the sides (guitars, synth pads) or across the stage (toms). The kick is unusually subdued in the mix – a striking contrast with Odeholm (Audio Example 3.2), as we will see (Figure 6) – much like the cymbals. For those, we use arrows spanning the width of the stage to indicate that various cymbals are placed far back in the mix.

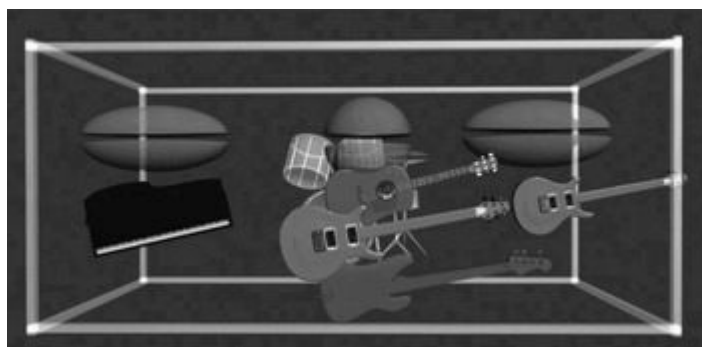


Figure 4: Moore’s sound-box, reproduced from Dockwray and Moore⁵⁰

47 Moore 2001, 120–126.

48 Moylan’s framework, developed into a textbook from his extensive professional practice as an audio engineer, offers a practical, experience-based model for understanding how spatial cues, such as panning, reverberation, and frequency balance, contribute to listeners’ perceptions of a recording’s structure. His approach is particularly valuable for analyzing metal production because it acknowledges both natural acoustic phenomena and the intentional manipulation of these phenomena to create specific aesthetic effects.

49 Moylan 2002, 49–54.

50 Dockwray/Moore 2010, 193.

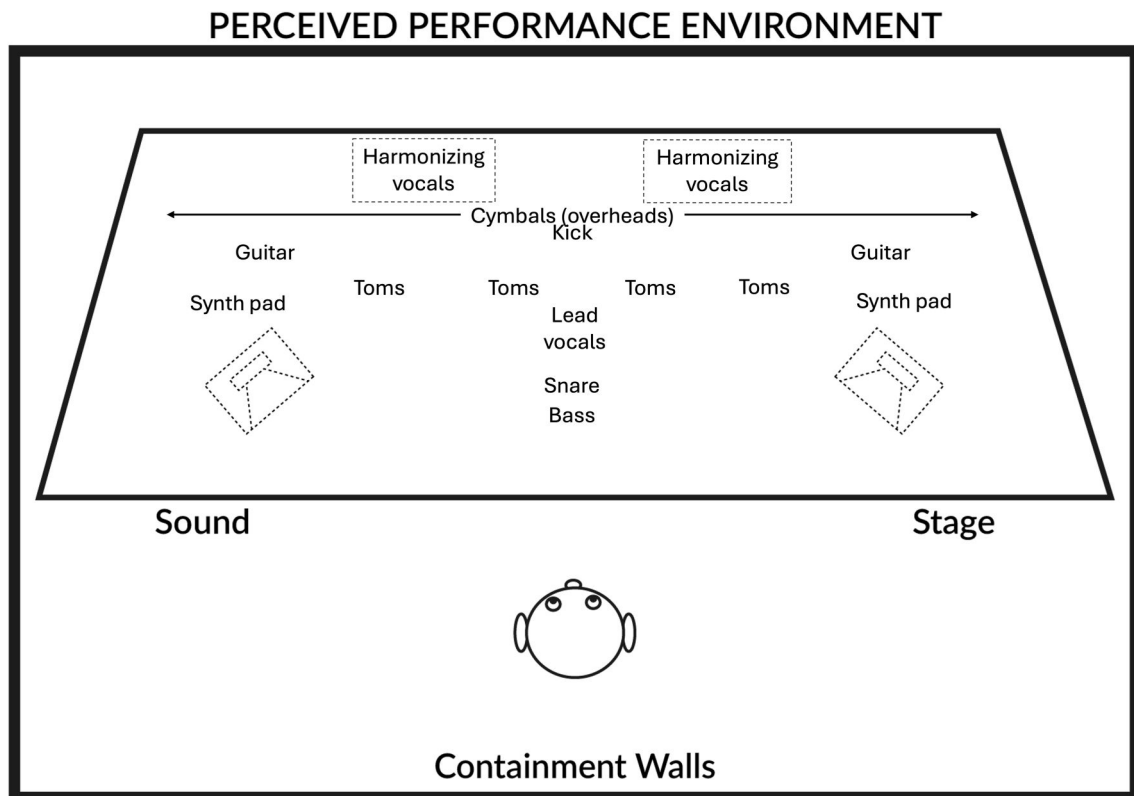


Figure 5: A spatial representation of sound distribution in Scheps's mix of "In Solitude," modeled after Moylan.⁵¹ It shows the width and depth dimensions; height is not displayed.

🔊 https://storage.gmth.de/zgmth/media/1222/Herbst_MixedResults_Audio03_1.mp3

Audio Example 3.1: Chorus, Scheps's mix of "In Solitude"

🔊 https://storage.gmth.de/zgmth/media/1222/Herbst_MixedResults_Audio03_2.mp3

Audio Example 3.2: Chorus, Odeholm's mix of "In Solitude"

Regardless of the extent to which producers prioritize realism and live fidelity, the soundstage is relevant to all recorded music, as it usefully maps listener intuitions around spatial sources and their boundaries. These spatial relationships influence not just aesthetic impressions but also how listeners parse individual elements within a mix, determining which sounds are foregrounded, which blend together, and how these relationships shift throughout a song. Such considerations indicate how the soundstage serves as a model for analytical listening. It is a practical tool for interpreting how spatial cues in a mix correspond to listeners' intuitive experiences of sound placement and separation, insights that are particularly valuable in understanding the aesthetics of metal production.

Depth: Loudness, Brightness, and Spatial Cues

Technically and perceptually, an impression of *depth* is determined by three primary variables: loudness, brightness, and spatial cues. These three dimensions work together to create the listener's perception of distance and proximity in a recording, with each dimension offering producers different tools for spatial manipulation. The most important factor is vo-

⁵¹ Moylan 2002, 50.

lume, as we encountered above with the quiet backing vocals at the back of Scheps's stage. Secondary considerations include brightness and the spatial cues of reverb and delay.

The relationship between physical depth and acoustic brightness is environmentally learned.⁵² Brightness, which music psychologists have shown to correlate with a high spectral centroid (i.e., a high average frequency within the spectrum),⁵³ creates a sense of directness and proximity based on learned ecological perceptions of how sound details change over distance.⁵⁴ Distant sounds exhibit a pattern whereby high-frequency sounds dissipate more than low sounds because their shorter wavelength cycles encounter more friction in the air. Whispers, for instance, cannot carry well over long distances, as one can imagine. Sibilants, the high-frequency sounds in "s" consonants, behave similarly, making them a tool for producers to create an impression of closeness. They enable Odeholm's previously cited goal to "maximize every detail on every level" partially by reinserting hyperreal amounts of high-frequency detail that would ordinarily be lost in an acoustic setting.

The in-your-face proximity that Odeholm seeks can be dramatically enhanced by not just raising brightness through equalization but also through the more drastic route of distortion. An equalizer raises brightness by selectively boosting the volume of individual frequency bands. Conversely, distortion boosts the volume across the entire frequency spectrum, resulting in a more complex timbre due to its enrichment of the signal with harmonic overtones.⁵⁵ By increasing harmonic content uniformly, distortion especially accentuates middle- and higher-frequency overtones, effectively brightening the sound more holistically than the targeted boosts of an equalizer. Rather than the localized brightness achieved through equalization, distortion creates a holistic rise in brightness and an overall saturated sound. This technique exemplifies how producers can exploit psychoacoustic principles to manipulate depth perception; in Odeholm's case, bringing sounds unnaturally close to create a hyperreal sensation of immediacy.

Odeholm heavily runs all audio signals through multiple distortion processors (e.g., saturation, overdrive, clipping), including the vocals. His guitars are extremely bright, and the kick and snare provide an impression of hyperreal closeness, as shown in Figure 6 (Audio Example 3.2). The closeness of central instruments and vocals in Odeholm's mix differs noticeably from Scheps's mix, as previously illustrated (Figure 5) (Audio Example 3.1). But however much one might link this use of distortion with new-school production centered around hyperreality, Scheps and Exeter also employ distortion towards different ends. Scheps, in particular, uses distortion to maintain transients and punch qualities (i.e., bursts of dynamic power or weight).⁵⁶ He does not like the sound of compression, which would reduce transients, and instead uses "parallel distortion," a technique of blending a distorted copy of the signal with the unprocessed signal. One could argue that this is a way of foregrounding live authenticity. A live engineer endeavors to preserve the idea of a band and enhance it (e.g., through punch). By contrast, Odeholm views production as autographic art; in his live performance, he strives to replicate the studio sound by incorporating many of the same sonic elements used on the record.

52 See Zagorski-Thomas 2014, 7–14; Walther-Hansen 2020.

53 Schubert/Wolfe 2006.

54 Moylan 2002, 185–188.

55 Herbst 2016, 120–134.

56 See Fenton/Lee 2019 and Herbst/Mynett 2022a, 22.

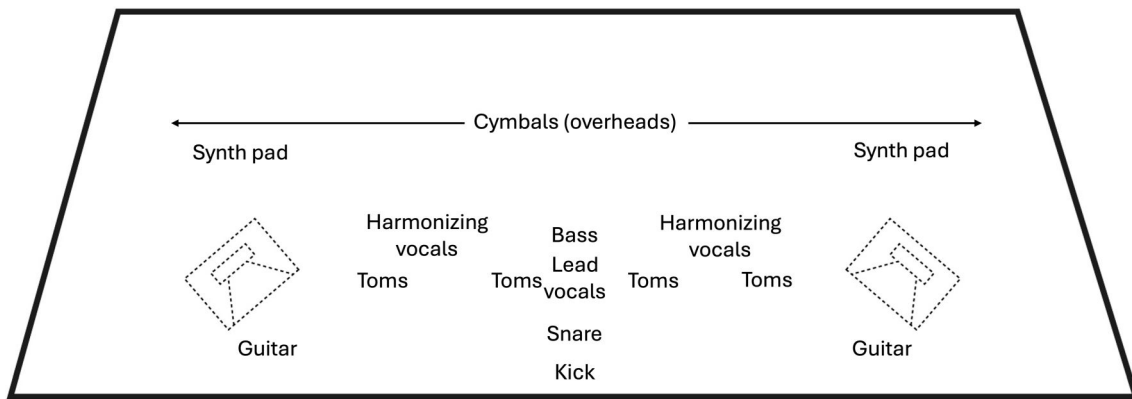


Figure 6: A spatial representation of Odeholm's mix of "In Solitude"

Other important components in controlling depth are time-based effects such as reverb and delay, which create a sense of space and locate instruments within the soundstage relative to each other and to the listener's position. Zagorski-Thomas introduced the concept of "functional staging," which refers to staging instruments and vocals based on ideal performance and listening situations.⁵⁷ Concerning rock music, Zagorski-Thomas argued that higher frequencies use audible reverb and delay to create a sense of being in a stadium or other large venue. By contrast, bass frequencies and instruments are often intentionally mixed "dry" because reverb and delay can detract from their intended function of providing clear rhythmic cues and low-end punch.⁵⁸ It is important to note that this selective treatment of frequencies – boosting reverb on high frequencies while minimizing it on low frequencies – is a basic mixing convention used across genres to avoid phase issues and muddiness. In metal, however, this practice is not merely technical but also contributes to aesthetic goals, exaggerating spatial cues in a larger-than-life manner. Functional staging is another example of "sonic cartoons" where beneficial aspects of sound are exaggerated while others are attenuated for musical and/or social reasons.

Compared to rock music, metal production tends to take this approach one step further. In metal, most instruments are kept completely dry to ensure maximum closeness and impact.⁵⁹ We have already pointed out how Odeholm keeps each instrument at the front of the soundstage or even manipulates sound to appear as if jumping out of the speakers, afforded by maximum loudness, extremely bright (distorted) sounds, and a complete lack of reverb.⁶⁰ There is a certain brutality to this "sledgehammer" aesthetic that forcefully closes the *stage-to-listener distance*, Moylan's term for "the distance between the grouped sources that make up the soundstage and the perceived position of the audience/listener."⁶¹ According to Moylan, the stage-to-listener distance not only "establishes the front edge of the stage with respect to the listener" but also "determines the level of intimacy of the music/recording."⁶² This kind of proximity may be familiar in the context of emotionally sensual performances – crooning, jazz combos, heartfelt ballads, or folk songs. But notions of proximity also apply within this

57 Zagorski-Thomas 2010, 252 and *passim*.

58 Ibid., 257.

59 Mynett 2017, 65.

60 Herbst 2014, 79–88.

61 Moylan 2002, 52.

62 Ibid.

more violent aesthetic where Odeholm is driven to impress upon the listener power and immensity. Hyperreal mixes that seem to leap out at the listener can be found among several new-school producers, albeit to a lesser extent than Odeholm.

Regarding proximity, both generally speaking and between instruments/vocals within the soundstage, different production schools vary in their treatment of a perceptual quality that audio engineers call cohesion. This quality refers to a particular impression of consistency around real or imagined sound sources, best explained through an example. For instance, drum sounds within a mix may or may not seem as though they originate from a single kit as one finds in a live acoustic environment. Indeed, drum sounds are particularly prone to fragmentation when they are recorded with individual spot microphones or when they are enhanced or replaced by samples. A mix engineer might strive for a cohesive mix by treating the drum parts in such a way that they consistently resemble a unified kit. Conversely, they might treat the component sounds independently to arrive at a more abstract or diffuse set of percussive sounds. In this way, the descriptor “cohesive” need not imply a preferential value judgment, and not all mixes strive for this kind of cohesion. While one might expect the groups of new- and old-school producers to fall neatly into two camps regarding cohesion, our examination of their drum treatments reveals that they did not all approach cohesion in predictable ways.

Within the old-school paradigm, one might assume that cohesion would be particularly valued. In a real acoustic setting, which old-school producers generally attempt to reproduce, the cymbals would be relatively balanced or possibly even louder than the other drums (i.e., toms, kick, and snare). Older rock records, such as AC/DC’s “Back in Black,”⁶³ contain relatively loud cymbals because they naturally sound louder in a live environment. Likewise, within the new-school paradigm, one might assume that producers would eschew drum-kit cohesion for the exaggerated, independent treatments that result in hyperreal proximity. One might expect their hyperreal approach to focus on high-impact drums like the snare and kick while attenuating the cymbals, as they easily sound abrasive when heavily processed and distorted.⁶⁴

However, our analysis reveals that the relationship between production approach and cohesion is more complex than a simple old-/new-school dichotomy would suggest. Mixes of some new-school producers, such as Middleton’s (Figure 7) (Audio Example 3.3), do indeed favor separation over cohesion. While his kick drum and, to some extent, his snare are very present and upfront in the mix, the rest of the drum kit – toms and cymbals – sound as if they come from next door. Yet, this is hardly different in Scheps’s mix, where the snare drum dominates, the kick is barely audible and back in the mix (although one could argue that this is how a kick would sound live without amplification), and the toms appear as though they originate from another room. Exeter’s (Figure 8) (Audio Example 3.4) mix similarly emphasizes independence between drum components over cohesion: very loud kick, quiet snare, almost inaudible cymbals, and loud bass, with its low-end emphasized, separated from the mid-frequency focused guitars. In our opinion, the most internally cohesive drum mix is that of Bogren (Figure 9) (Audio Example 3.5). He achieves the balancing act of making the drums sound cohesive as an instrument, with effective shells and audible cymbals within a realistic room. The mix features distinct but not isolated-sounding drums, bass, guitars, and vocals while simultaneously obtaining an effective wall of sound with a balance of sonic weight and

63 AC/DC 1980.

64 Herbst/Mynett 2021, 639.

clarity. It is a contemporary metal mix that showcases and enhances the musical composition and recorded performances without veering into the hyperreal; at the same time, the mix can be perceived as a believable performance within the conventions of rock and metal.

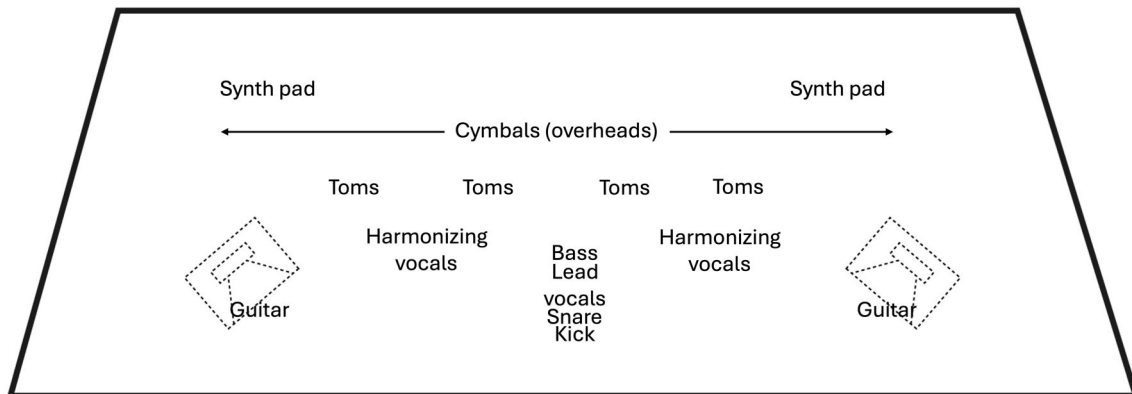


Figure 7: A spatial representation of Middleton's mix of "In Solitude"

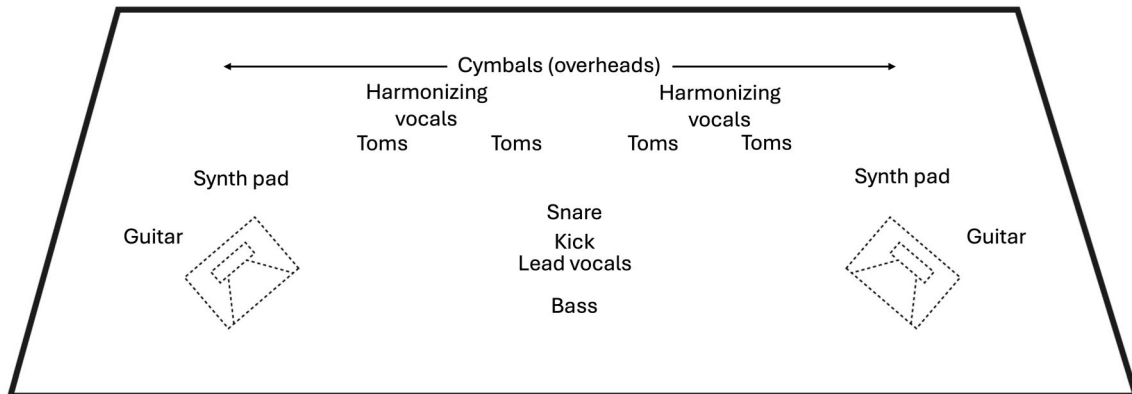


Figure 8: A spatial representation of Exeter's mix of "In Solitude"

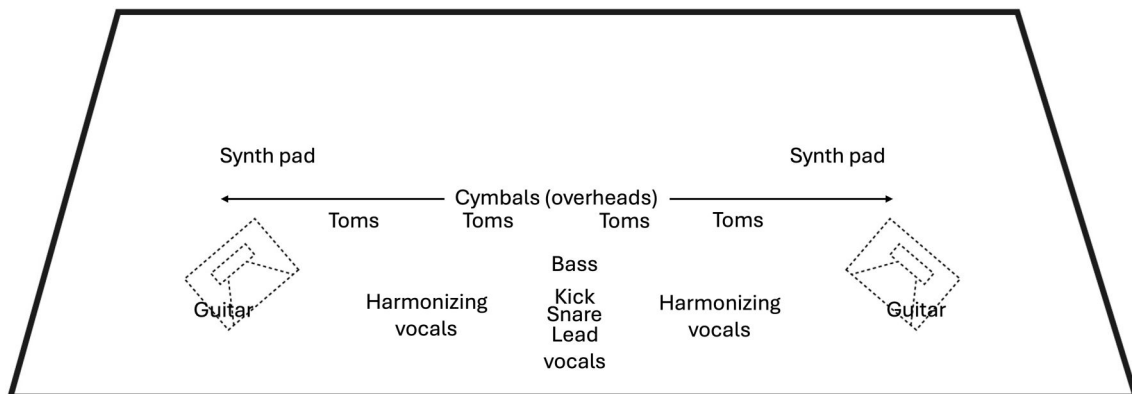


Figure 9: A spatial representation of Bogren's mix of "In Solitude"

🔊 https://storage.gmth.de/zgmth/media/1222/Herbst_MixedResults_Audio03_3.mp3

Audio Example 3.3: Chorus, Middleton's mix of "In Solitude"

🔊 https://storage.gmth.de/zgmth/media/1222/Herbst_MixedResults_Audio03_4.mp3

Audio Example 3.4: Chorus, Exeter's mix of "In Solitude"

🔊 https://storage.gmth.de/zgmth/media/1222/Herbst_MixedResults_Audio03_5.mp3

Audio Example 3.5: Chorus, Bogren's mix of "In Solitude"

Width: Panning and Stereo Impressions

Width is equally as important as depth for two reasons.⁶⁵ Firstly, a wide stereo impression facilitates the larger-than-life wall of sound desired in most metal genres. One way to achieve a wide impression is by using two slightly different versions of the same instrument or part – one panned to the left and the other to the right – whereas a narrow impression results when both versions are centered equally. However, panning is not the only way to achieve an impression of width. In addition to the horizontal separation of sound sources, producers can generate width by distributing frequency differently across channels and across the stereo image generated by time-based effects. Collectively, these factors combine to enhance perceptual wideness. And the wider the sound, the more it enhances impressions of size,⁶⁶ thus contributing to “heaviness,” one of metal music’s defining ideals.⁶⁷

Secondly, width can supplement depth, augmenting impressions of proximity. For example, imagine listening to a band play live while standing close to the stage: the speakers are spread left and right, conveying great width (Audio Example 4.1).⁶⁸ If you stood further back, that impression would narrow, as the speakers would seem closer together (Audio Example 4.2).⁶⁹ A similar effect can be achieved digitally. Panning or psychoacoustic processors can alter width, modifying the entire stereo field as if a listener were moving across a fixed arrangement of sound sources.⁷⁰ Figure 10 illustrates this effect by comparing two opposite width configurations of a Waves S1 Stereo Imager plugin. As Figure 10 shows, narrowing the depth makes a listener experience a more distanced clustering of sound sources, while increasing the depth has the effect of spreading them as though they were heard up close. In this way, width is closely linked to perceptual proximity, which producers can manipulate through panning.

Unsurprisingly, most HiMMP producers panned all guitars fully wide, maximizing the overall width. Although drum-overhead microphones, capturing cymbals and the entire drum kit, are sometimes panned slightly more narrowly, the producers panned them fully wide as well, possibly to reinforce the boundaries of the soundstage and maximize width.

65 While headphones, stereo speakers, and other spatial audio configurations afford different spatial perceptions, our discussion applies equally to each mode of listening. For a comparison of headphones and stereo speakers, see Herbst 2014, 79–88.

66 Herbst 2014, 86–87.

67 Herbst/Mynett 2022a, 19–21.

68 Herbst 2014, 86–88.

69 Of course, not all listeners imagine this kind of live staging while listening. Nevertheless, this thought experiment well illustrates the relationship between width and perception of distance.

70 While the perceptual effects of width, such as an enhanced sense of size from increased lateral separation, are grounded in psychoacoustic principles, some aspects of width in recorded music also stem from the working habits of producers held over from older format limitations or practices of consumption. For instance, some producers elected not to use stereo widening out of historical concerns with the mono compatibility of radio. Such retentions from older cultural practices can lead to panning decisions that are not sufficiently explained by cognitive thresholds and preferences. Regardless of the degree to which psychoacoustic principles or residual habits underlie producer decisions, reducing width makes sound feel less directional and diminishes the distinct spatial placement of instruments and vocals.



Figure 10: Close proximity with wide panning (left), far distance with narrow panning (right) (Waves S1 Stereo Imager plugin)

🔊 https://storage.gmth.de/zgmth/media/1222/Herbst_MixedResults_Audio04_1.mp3

Audio Example 4.1: Reference mix of “In Solitude” widened by 200%; Lead-in and Verse

🔊 https://storage.gmth.de/zgmth/media/1222/Herbst_MixedResults_Audio04_2.mp3

Audio Example 4.2: Reference mix of “In Solitude” narrowed to 25%; Lead-in and Verse

Some producers employed more advanced techniques than the panning methods discussed above to enhance width and manipulate spatial perception. For instance, Bogren (Figure 9) (Audio Example 3.5) and Getgood (not diagrammed) utilized psychoacoustic stereo processing, such as the Waves S1 Imager (shown in Figure 10), to expand the guitars beyond the maximum width permitted by the loudspeaker configuration, ensuring they do not occupy the same space as the drum-overhead microphones. In other words, the drum overheads are positioned at the edges of the regular soundstage, while the stereo-widened guitars expand the soundstage even further to enhance size and proximity. This approach to extended stereo width exemplifies how newer production techniques can create perceptual spatial dimensions that exceed physical limitations, a hallmark of hyperreal production aesthetics. Nordström (Audio Example 3.6) and Middleton widened the mix differently by slightly narrowing two of the four available rhythm guitars, stating that the brighter pair should be wider. This axiom makes sense insofar as higher or brighter sounds can be directionally located more easily than lower sound sources – a subwoofer, by counterexample, can be positioned anywhere in a room without affecting the loudspeaker system’s spatialization. In Nordström and Middleton’s mixes, the brighter guitars, which are more perceptually linked to location, are deliberately spread across the soundstage, while the darker guitars are only panned at 80%. The total arrangement is a V-shape mix, a production term derived from the symmetry of the resulting left and right panning settings (Figure 11). As the color highlights in Figure 11 show, this arrangement takes different frequency bandwidths and pans them increasingly as they get higher. Lower instruments like kick and bass are placed in the stereo center, while higher instruments are increasingly positioned to the side.

🔊 https://storage.gmth.de/zgmth/media/1222/Herbst_MixedResults_Audio03_5.mp3

Audio Example 3.5: Chorus, Bogren's mix of "In Solitude"

🔊 https://storage.gmth.de/zgmth/media/1222/Herbst_MixedResults_Audio03_6.mp3

Audio Example 3.6: Chorus, Nordström's mix of "In Solitude"

🔊 https://storage.gmth.de/zgmth/media/1222/Herbst_MixedResults_Audio03_7.mp3

Audio Example 3.7: Chorus, Scheps's mix of "In Solitude"



Figure 11: V-shape panning settings of Bogren's guitar signal with each increasing frequency band panned further to the sides (Waves S1 Stereo Imager and ADPTR Audio Metric AB, professional stereo imaging plugins)

Yet another strategy to increase width is to assign different timbres (e.g., guitar tones) to the left and right speakers. Perhaps surprisingly, no producers chose this strategy. However, Exeter (Audio Example 3.4) and Scheps (Audio Example 3.7) achieved a similar effect using different EQ settings for the left and right guitar channels. Interviews did not reveal whether this decision was motivated by a greater sense of width or whether it reflects an older production approach centered on live fidelity – traditional heavy metal bands typically have different guitar tones on both sides, which correspond to the preferred tone of each guitarist.⁷¹ This dual-guitar panning approach is less common in recorded metal of the last decade as aesthetic priorities shift towards rhythmic precision.⁷² The challenges of achieving tight synchronization between guitars encourage all rhythm guitar parts to be recorded by a single player, making tonal distinctions between separate guitar channels somewhat less individual and intuitive.

A visualization of the guitar sub-mix's stereo field in Figure 12 displays the effects of the chosen approaches at a singular slice of time. In the figure, frequency is indicated on the vertical axis with higher frequencies above. The horizontal axis indicates stereo position, with the guitar signal shown in two ways: a dark orange cloud indicates a range of how widely the signal is distributed across the stereo width within a temporal window of

71 Mynett 2017, 137–138.

72 Mynett 2019, 299–301.

one second; the bright orange line shows an averaging of the same signal distribution on a much shorter timescale of milliseconds. We annotated the visualization with a bright yellow line to indicate the degree to which the energy distribution of the guitar signals is skewed to one side or symmetrical. Exeter's (Audio Example 5.1) and Scheps's (Audio Example 5.2) guitars, shown at the top of Figure 12, are asymmetrical: the greater brightness on one channel shifts the balance towards the left and right speakers, respectively. This is not the case with Getgood's (Audio Example 5.3) and Middleton's (Audio Example 5.4) guitar mixes, shown at the bottom. These producers selected the same guitar tones on the left and right channels, resulting in an initially narrower spatial impression. Getgood and Middleton then overcame this limitation using psychoacoustic widening to create a wider, yet still symmetrical image. By contrast, Exeter and Scheps create a very wide studio image by assigning different EQ treatments to each guitar channel, indicating that an old-school approach can achieve an impression of width similar to the hyperreal aesthetics of new-school productions.⁷³

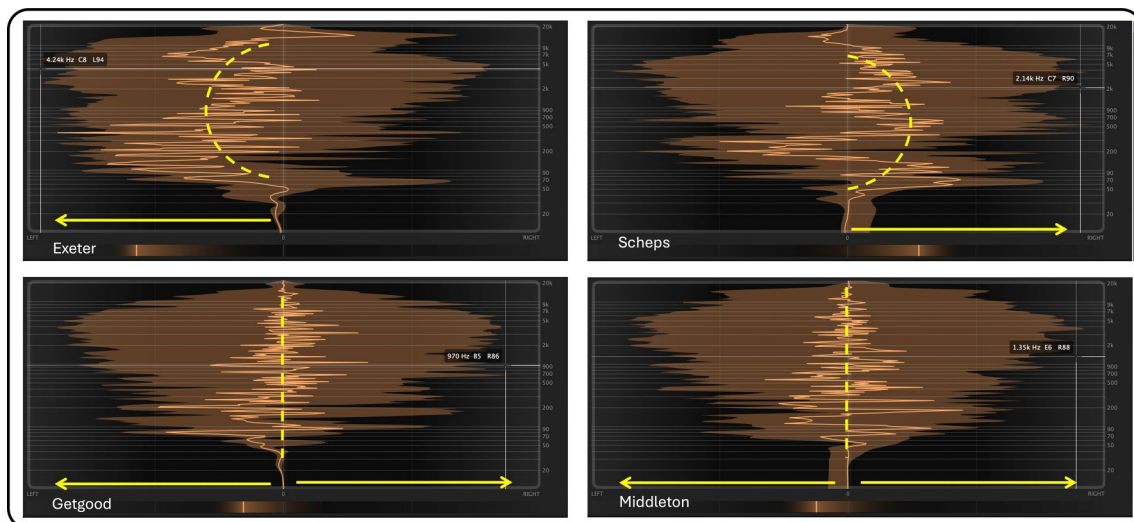


Figure 12: A comparison of panning symmetries (created with ADPTR Audio Metric AB). The dotted lines show overall trends.

🔊 https://storage.gmth.de/zgmth/media/1222/Herbst_MixedResults_Audio05_1.mp3

Audio Example 5.1: Chorus guitars, Exeter's mix of "In Solitude"

🔊 https://storage.gmth.de/zgmth/media/1222/Herbst_MixedResults_Audio05_2.mp3

Audio Example 5.2: Chorus guitars, Scheps's mix of "In Solitude"

🔊 https://storage.gmth.de/zgmth/media/1222/Herbst_MixedResults_Audio05_3.mp3

Audio Example 5.3: Chorus guitars, Getgood's mix of "In Solitude"

🔊 https://storage.gmth.de/zgmth/media/1222/Herbst_MixedResults_Audio05_4.mp3

Audio Example 5.4: Chorus guitars, Middleton's mix of "In Solitude"

⁷³ It is worth noting that many producers applied stereo widening on the master bus (i.e., the stereo track that sums together all other tracks for audio exportation to the speakers), affecting the entire mix. Accordingly, widening occurs not only on the guitars directly but also indirectly.

Earlier, we explained the V-shape mix arrangement (Figure 11), which typically achieves maximum width by widening the higher frequencies. For lower frequencies, many of the producers (Middleton, Exeter, Getgood, and Nordström) use purely, or mostly, mono bass signals across the frequency spectrum. Middleton’s bass guitar, for instance, is entirely mono (Audio Example 6.1). To see this, compare the images of Figures 13 and 14, which show the bass guitar signal using blue clouds overlayed on the orange guitar clouds. Unlike the productions in Figure 14, which generally involve visible, blue bass-clouds, Middleton’s bass signal in Figure 13 is virtually a straight blue line in the very center. In Figure 14, Odeholm (Audio Example 6.2) and Bogren (Audio Example 6.3) show the greatest bass width, followed by Scheps (Audio Example 6.4) and Nordström (Audio Example 6.5) (whose bass only widens at the very top). Further below, in our discussion of the height dimension, we examine these differences in greater detail, particularly with the concept of a “meta-instrument.”

🔊 https://storage.gmth.de/zgmth/media/1222/Herbst_MixedResults_Audio06_1.mp3

Audio Example 6.1: Chorus guitar/bass overlay, Middleton’s mix of “In Solitude”

🔊 https://storage.gmth.de/zgmth/media/1222/Herbst_MixedResults_Audio06_2.mp3

Audio Example 6.2: Chorus guitar/bass overlay, Odeholm’s mix of “In Solitude”

🔊 https://storage.gmth.de/zgmth/media/1222/Herbst_MixedResults_Audio06_3.mp3

Audio Example 6.3: Chorus guitar/bass overlay, Bogren’s mix of “In Solitude”

🔊 https://storage.gmth.de/zgmth/media/1222/Herbst_MixedResults_Audio06_4.mp3

Audio Example 6.4: Chorus guitar/bass overlay, Scheps’s mix of “In Solitude”

🔊 https://storage.gmth.de/zgmth/media/1222/Herbst_MixedResults_Audio06_5.mp3

Audio Example 6.5: Chorus guitar/bass overlay, Nordström’s mix of “In Solitude”

For now, we turn to how the stereo bass signal is generated. In Odeholm’s production, the bass guitar began as a mono signal that he duplicated and slightly modified. Using a stereo amplifier cabinet ordinarily meant for guitars, Odeholm added a widened bass to his mono signal, creating a stereo bass sound. Comparing Odeholm’s overlayed guitar and bass signals in Figure 14 reveals that the bass has almost the same stereo image as the guitars, only less distorted and less bright – features not visible in the figure but noticeable in Audio Example 6.2. Bogren employed a similar strategy, incorporating a chorus effect parallel to his main signal, which we detail below. Unexplained still are the stereo signals in Scheps’s and Nordström’s bass. Neither mentioned widening the bass, suggesting that algorithms within the audio processors are most likely responsible.

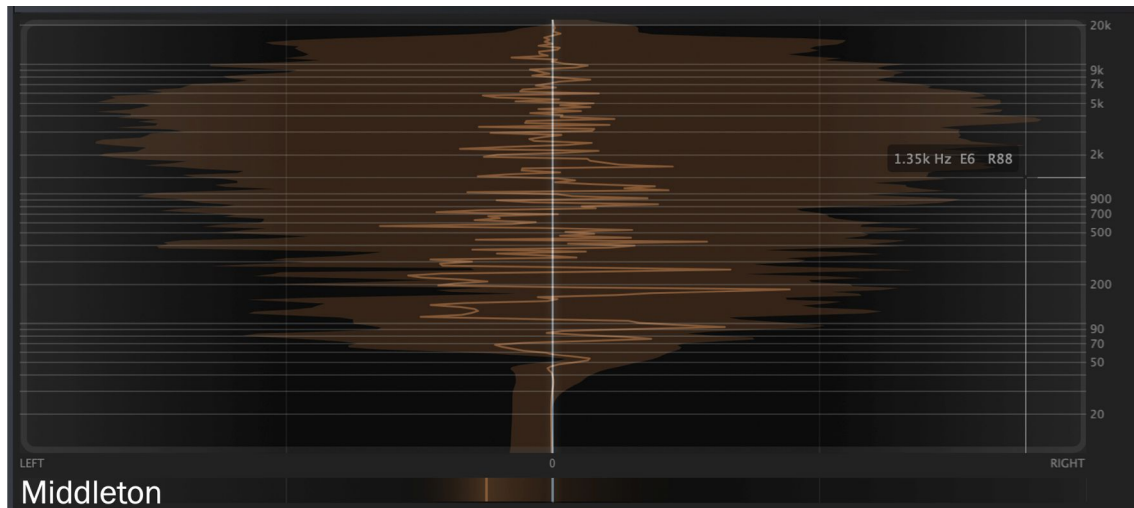


Figure 13: An overlay of guitar (orange) and bass (blue) panning symmetries, Middleton's mix of "In Solitude"

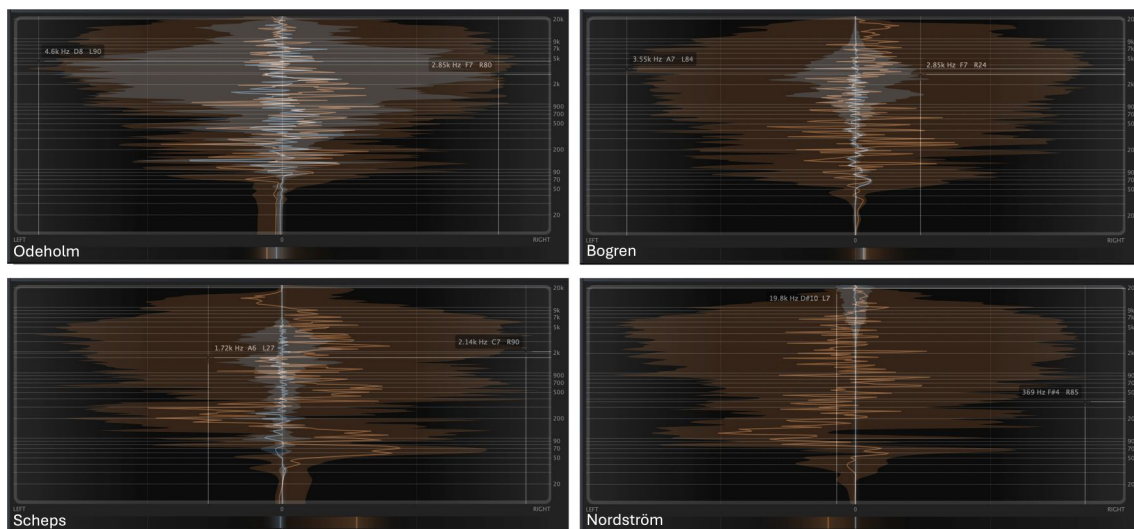


Figure 14: A comparison of guitar (orange) and bass (blue) panning symmetries

Height: High- and Low-Frequency Effects

Low-frequency components are essential for heaviness in metal, the metaphorical representation of immense sound source size and *gravitas*.⁷⁴ Since the mid-1990s, many metal bands have sought increased heaviness by downtuning their guitars or extending the low range by adding a seventh string or more. Specialized engineering treatments of the guitar sound, or kick drum, can enhance this effect. In such instances, a producer "rearranges" the expected order of frequency ranges for different instruments and sounds in a counterintuitive way, analogous to voice crossing in traditional score-based orchestration. Much like a composer might have a cello play higher than the accompanying second violins or write an alto part that crosses below the tenor, so too might a metal producer place the electric guitar lower than the bass guitar for a momentary special effect. This is especially true in subgenres like djent, which involve eight or more guitar strings extending one octave below a standard

74 Herbst/Mynett 2022a, 19–21.

electric guitar.⁷⁵ Rather than merely duplicating the bass, the extended electric guitar adds a unique timbre owing to its lighter gauge strings and different performance style.

This kind of “voice-crossing” treatment can occur in at least two ways: as described above, a songwriter could swap instruments or sound components; or an audio engineer could emphasize various frequency components in a fully written song. For instance, an engineer might emphasize the bass region of a guitar track and attenuate the lower end of an electric bass track.⁷⁶ The overall impression is akin to “voice-crossing,” only now achieved through the frequency response of different audio signals from each instrument. While the first kind, relating to song arrangement, does not occur in “In Solitude,” the second kind, relating to audio engineering, does.

As in other genres, metal requires clarity so listeners can detect individual instruments, but metal engineers operate in a particularly dense sound environment where many instruments overlap the same low-frequency bands,⁷⁷ resulting in significant frequency masking.⁷⁸ Frequency masking occurs when overlapping sound sources cause some frequencies to obscure others, making it difficult for each instrument to be heard distinctly. Producers cannot change a recorded performance, but they can alter the signal’s frequency response in ways that shape how different elements are perceived – much like how placing a pillow, earplugs, earmuffs, or water in front of a listener’s ear would selectively filter certain frequencies. In this way, producers control how the signal is mediated so that different components pass through without interference from others.

Additionally, the acoustics of an audio signal are more complex than may seem from real-time listening. Looking at loudness-normalized spectra of kick, bass, and guitar in Bogren’s mix (Figure 15), the kick drum, although perceived as a single percussive sound, simultaneously combines a lower thud around 86Hz and an upper click around 3.6kHz, along with local spectral peaks and valleys in between. This affects the overall sound compared to the guitar and bass. When a pitched instrument plays a note, the composite sound contains a fundamental frequency (e.g., 440Hz, hence the designation A440) and a series of harmonic overtones that are multiples of that fundamental (880Hz, 1320Hz, 1760Hz, etc.). Producers can modify the signal by emphasizing or attenuating these different spectrum regions to customize their timbres and balance the density of different frequency regions.⁷⁹

During the mid-2000s and earlier, the bass was often placed below the kick,⁸⁰ partly because the kick had a less pronounced low-end than would later be possible with more advanced drum triggering.⁸¹ In this older practice, the bass provides most of the mix’s

75 See Herbst/Vallejo 2023, 46–47.

76 Herbst/Mynett 2023, 11–13.

77 See Mynett 2017, 15–16.

78 Herbst/Mynett 2023, 11–14.

79 Because Figure 15 involves a single slice of time (the very beginning of the song), its appearance would change at other moments when the bass and guitars play other notes. To wit, the signal for the kick would appear the same but the fundamental frequency of the bass and guitar would change depending on their note selection at that time. Despite this variance, our overall argument about global frequency relationships between sound sources still applies.

80 Wanasek 2017.

81 Children of Bodom’s *Hate Crew Deathroll* (2003) and Soilwork’s *Stabbing the Drama* (2005) are examples of the lower-bass approach. Metallica’s *Metallica* (1991), widely referred to as the Black Album, represents an early example of the lower-kick alternative. For a more recent instance, see Gojira’s *Fortitude* (2021).

sonic weight. By contrast, Bogren’s kick appears “scooped”: in Figure 15, his kick has a strong component that peaks below the fundamental of the bass as well as another salient component in the high-end, 3–4kHz range.⁸² Specifically, the kick drum’s lower peak measures around 86Hz while the peaks of the guitars and bass overlap around 110Hz, matching their A2 pitch as notated in our transcription (Figure 16). This arrangement works well when the kick drums involve intricate rhythms with frequent rests rather than constant double-bass drumming. Otherwise, the double kick would overwhelm with unclear, booming hits. In addition to the advantage of achieving clarity within a dense sonic space, Bogren’s mix (Audio Example 7.1) has more power and excitement this way, which is a likely reason why similar approaches are becoming more common in metal production.

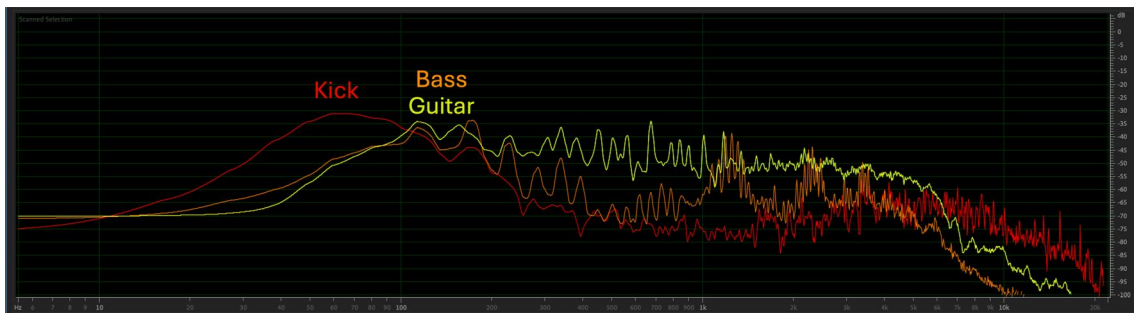


Figure 15: Kick, bass, guitar spectra, Bogren’s mix of “In Solitude,” loudness-normalized (created with Adobe Audition)

♩ = 200

D5 prolongation

Dadd⁹ Dm

Gtr. 1 & 2

P.M.-----I

Dadd⁹/A Dm/A

Figure 16: The first four bars of “In Solitude”



https://storage.gmth.de/zgmth/media/1222/Herbst_MixedResults_Audio07_1.mp3

Audio Example 7.1: Lead-in and verse kick/bass/guitars, Bogren’s mix of “In Solitude”

82 See Mynett 2017, 266.

Contrasting with Bogren, Odeholm (Audio Example 7.2) emphasized overlap between the kick and bass guitar, visible around the overlapping lowest peaks of the bass and kick in Figure 17. His primary concern was low-end impact, where the kick and bass, and to some extent guitars, reinforce each other. To maximize impact, Odeholm employed a phase interactions mixer (SoundRadix Pi), a device that automatically adjusts phase relationships between stereo signals.⁸³ While phase interactions mixers are usually employed to avoid unwanted signal cancelations, Odeholm used it to constantly adjust phase relationships between the low-frequency ranges of the kick, bass, and guitar. The resulting sound carries more impact than Bogren's lighter low-end, but it is less clear, with the kick and bass blurring together.

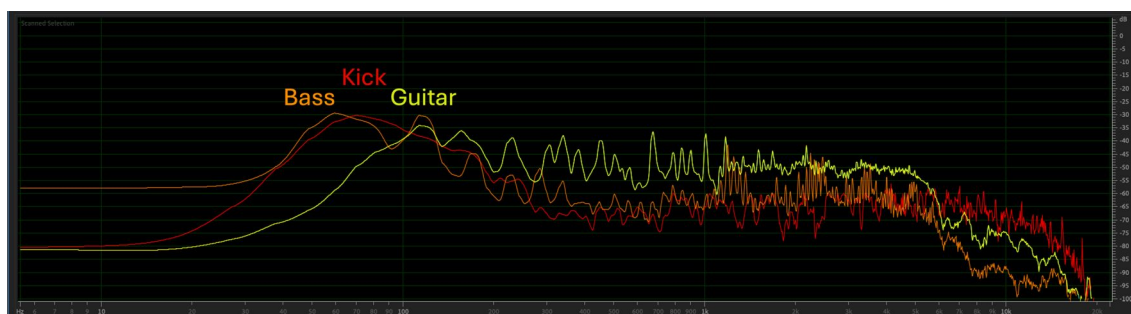


Figure 17: Kick, bass, guitar, spectra (Odeholm, loudness-normalized)



https://storage.gmth.de/zgmth/media/1222/Herbst_MixedResults_Audio07_2.mp3

Audio Example 7.2: Lead-in and verse kick/bass/guitars, Odeholm's mix of "In Solitude"

Exeter adopted a third approach, bass below kick (Audio Example 7.3). In addition to using audio processors that shape frequency response (e.g., equalizers, overdrive), Exeter programmed a sub-bass synthesizer one octave below the bass and automated the track to turn it on or off in specific sections. As Figure 18 shows, the sub-bass synthesizer is active in most of the slower sections, including the breakdown, and disabled during faster sections with double-kick drums, such as parts of the chorus. Consequently, Exeter's spectrum looks different (Figure 19). Comparing the local maxima of the bass and kick (i.e., where the peak of the lowest "hill" occurs in Figure 19), the bass and kick overlap in frequency, but the local maximum of the bass is slightly below that of the kick. When listening, the effect is subtle, but low-end details differentiate subgenres and production styles within metal and are a priority for engineers.

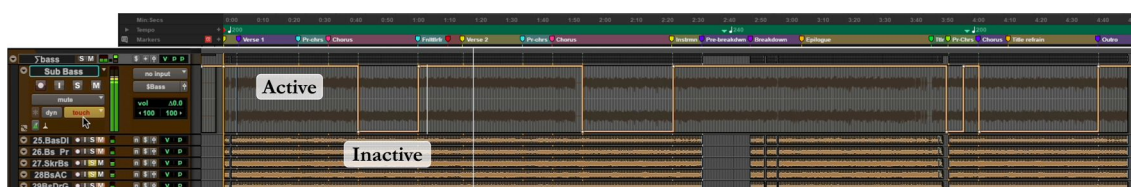


Figure 18: Automated sub-bass synthesizer in Exeter's mix of "In Solitude"

83 For an accessible explanation of phase correlation, the term for these kinds of stereo relationships, see Robjohns 2016. Phase interactions mixers are distinct from a dynamic EQ, which equalizes spectral imbalances by responding to loudness peaks or dips within defined frequency ranges.

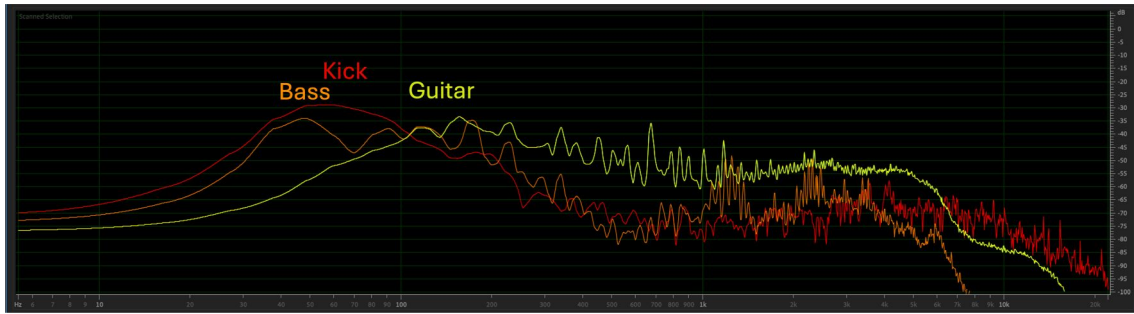


Figure 19: Kick, bass, guitar, spectra, Exeter's mix of "In Solitude," loudness-normalized

🔊 https://storage.gmth.de/zgmth/media/1222/Herbst_MixedResults_Audio07_3.mp3

Audio Example 7.3: Lead-in and verse kick/bass/guitars, Exeter's mix of "In Solitude"

We can now revisit our previous discussion of stereo width in relation to guitar and bass. In the old-school, live-fidelity tradition, a clear separation between instruments is essential for individual musicians to shine through. In new-school hyperreality, however, the ideal seems to have shifted. Getgood's interview echoes Odeholm's notion of low-frequency impact:

It's about all the things hitting at the same time and generating this oneness of the sound. So, talking about the bass, when you've got a distorted bass guitar, it really adds huge amounts of richness to that lower end. Combining with the guitars in this way makes it sound like one massive floor-to-ceiling sound, punctuated so that every attack is hit with a kick drum, and you've got this punch of the snare as well. It's like the sum of all those parts together gives me that sensation of a huge object, the heavy aspect.

Both Getgood and Odeholm allude to the idea of a "meta-instrument," where the guitar, bass, and possibly the kick and snare appear as one inhuman and larger-than-life entity with hyperreal size and impact.⁸⁴

Odeholm exemplifies this new-school approach by blending instrumental sources in two ways: merging them in frequency space and overlapping them in stereo location, effectively dissolving traditional instrumental boundaries to create a unified sonic entity that transcends conventional band instrumentation. Comparing Figures 17 (Odeholm) and 19 (Exeter), Exeter's guitar is clearly distinct from his bass. While the guitar signal frequency contour remains relatively constant, the bass dips into a valley in the mid-frequency range before peaking around 1–2.5kHz. By contrast, Odeholm's guitar and bass are fairly uniform, neither exhibiting a pronounced dip nor peak. By overlapping in the same frequency area, Odeholm's guitar and bass signals reinforce each other and enable the impression that the two sources act as one. In fact, Odeholm's kick overlaps in much the same way, contributing to a meta-instrument effect. As interviews revealed, Odeholm's guitar and bass share the same amplifier and loudspeaker system and distortion treatment, making bass and guitars a single, coherent perceptual entity.

A shared amplifier and loudspeaker system impacts the impression of stereo separation, Odeholm's second technique for creating a meta-instrument sound. In one approach, Odeholm widens the bass to occupy the stereo extremes (Audio Example 8.1). As Bogren does currently (Audio Example 8.2), Odeholm previously used a parallel chorus

84 See Herbst/Mynett 2022b, 646.

effect in his workflow, which involves two separate signals that transform the bass from a mono to a stereo instrument. He altered his approach to exaggerate the stereo impression. As Figure 14 shows, the blue image representing the bass signal extends almost to the sides, much like the orange image of the guitar. By contrast, Bogren’s bass, also shown in blue, is mostly confined to the center, not to Nordström’s extent, but with much less stereo width than Odeholm’s. In another approach, Odeholm’s addition of a third guitar in the center reinforces the impression of a meta-instrument within the stereo space normally given to the bass. Whereas we previously discussed Odeholm’s moving the bass to the outside, where guitars traditionally reside, his addition of a third guitar in the center moves the guitars to where the bass ordinarily appears. The overall result is a spatial cross-merging that makes everything one coherent entity. This seems quite deliberate – of all the producers involved, Odeholm was the only producer to choose a three-guitar setup. Not only does his bass overlap with the guitar’s frequency spectrum and stereo image, but the guitars also occupy the same central (i.e., mono) space as the bass. The result is a sound we previously likened to a sonic “sledgehammer.”

🔊 https://storage.gmth.de/zgmth/media/1222/Herbst_MixedResults_Audio08_1.mp3

Audio Example 8.1: Pre-chorus stereo bass, Odeholm’s mix of “In Solitude”

🔊 https://storage.gmth.de/zgmth/media/1222/Herbst_MixedResults_Audio08_2.mp3

Audio Example 8.2: Pre-chorus stereo bass, Bogren’s mix of “In Solitude”

Conversely, Scheps audibly separates his bass and guitars in terms of height and width (Audio Example 7.4). The bass is considerably less distorted, with melodies and articulations clearly perceptible, especially when transitioning into higher registers during fills. Exeter’s bass and guitars are similarly distinct (Audio Example 7.3), partly because of his additional sub-bass and partly due to his loud bass volume, the center position in his stereo field, and the timbral difference between his bass and guitars. Bogren’s mix (Audio Example 7.1) falls into this camp, but not Nordström’s (Audio Example 7.5) – the other “in-between” producer. Nordström described the role of the bass rather enigmatically: “Bass in metal music is not an important instrument, but it is, at the same time, super important. But it’s not important if it is audible.” Nordström distinguishes between the bass extending the guitar’s frequencies downwards to enhance its impact (“super important”) and the possibility of hearing the bass as an instrument in its own right (“if it is audible”). His mix favors impactful blend, following the concept of a meta-instrument. While Nordström did not reflect on this approach to the same extent as Odeholm or Getgood, he expressed that he aims to “get the band to sound like one unit. It’s just like this should be like one big tank going forward.” All producers blended guitars, bass, and drums – particularly kick, snare, and toms – except Scheps and Exeter, suggesting that such a blend is integral to contemporary metal production.

🔊 https://storage.gmth.de/zgmth/media/1222/Herbst_MixedResults_Audio07_4.mp3

Audio Example 7.4: Lead-in and verse kick/bass/guitars, Scheps’s mix of “In Solitude”

🔊 https://storage.gmth.de/zgmth/media/1222/Herbst_MixedResults_Audio07_5.mp3

Audio Example 7.5: Lead-in and verse kick/bass/guitars, Nordström’s mix of “In Solitude”

One final consideration of the height domain concerns the drums and reveals how production choices can obscure or highlight genre idioms within metal. The blast beat, a staple of extreme metal drumming, exists in different variations but is typically intended to convey energy and aggression. For producers, dense and often intricate drumming makes achieving a clear separation of sounds at high speed within an equally dense arrangement both crucial and challenging. Among the drumbeats in Figure 20, the blast beat and what we call the punk beat present challenges to clear production. Readers can compare these drumbeat types in Audio Example 9.3.

Vast differences between the mixes are evident, with varying levels of tonal separation (e.g., an audible distinction between kick and snare) and methods of conveying musical expression. Scheps's drum mix (Audio Example 9.1) involves what is sometimes called "machine-gunning," a fusing of the blast beat's kick and snare into a single snare sound. Generally speaking, kick drums are more forgiving – able to withstand heavier processing without sounding unnatural – when it comes to signal processing, but snare drums require careful deliberation to resemble a human drummer. One consequence of Scheps's machine-gunning is a mechanical snare sound that audibly differs from acoustic drums, even though he did not use drum samples, typically associated with mechanical timbres. The relative uniformity between his kick and snare risks reducing the impact of the drumming during this climactic section of the song where the drums take up a blast beat (especially during 0:00–0:06 and 0:30–0:38 in Audio Example 9.1). Surprisingly, Exeter (Audio Example 9.2), whom we grouped with Scheps under the old-school paradigm, took the opposite approach. Exeter emphasized the kick (and bass guitar) while still keeping the snare audibly present. The combination of audible and distinct double kick and snare drum drives the music and captures the performance's energy. To our surprise, the new-school producers seemed to most successfully capture the feeling of a human-performed drum kit. In Otero's (Audio Example 9.3) and Bogren's mixes (Audio Example 9.4), the blast beat balances performative authenticity – a realistic depiction of live drumming that sacrifices sonic intelligibility – with maximum impact, an optimization of rhythmic accuracy and synchronization. Odeholm's mix (Audio Example 9.5) exemplifies the latter. His strategy of maximum impact goes, in his words, "beyond the sound of a band." This deliberate transcendence of reality is controversial within metal communities, which encompass a diverse range of subgenres and aesthetic ideals.

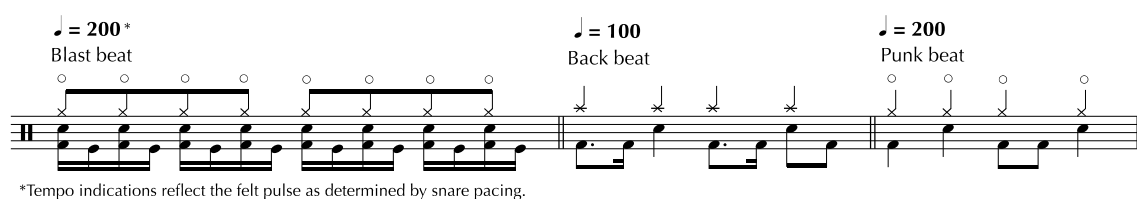


Figure 20: Drumbeat types heard throughout "In Solitude"

🔊 https://storage.gmth.de/zgmth/media/1222/Herbst_MixedResults_Audio09_1.mp3

Audio Example 9.1: Final chorus blast beat, Scheps's mix of "In Solitude"

🔊 https://storage.gmth.de/zgmth/media/1222/Herbst_MixedResults_Audio09_2.mp3

Audio Example 9.2: Final chorus blast beat, Exeter's mix of "In Solitude"

🔊 https://storage.gmth.de/zgmth/media/1222/Herbst_MixedResults_Audio09_3.mp3

Audio Example 9.3: Final chorus blast beat, Otero's mix of "In Solitude"

🔊 https://storage.gmth.de/zgmth/media/1222/Herbst_MixedResults_Audio09_4.mp3

Audio Example 9.4: Final chorus blast beat, Bogren's mix of "In Solitude"

🔊 https://storage.gmth.de/zgmth/media/1222/Herbst_MixedResults_Audio09_5.mp3

Audio Example 9.5: Final chorus blast beat, Odeholm's mix of "In Solitude"

This discussion of blending vs. separation regarding kick/bass/guitar and kick/snare combinations suggests that both strategies can be valid. One's preference may depend on particular instruments, their roles (e.g., soloist vs. rhythm section), and overall compositional context (e.g., chorus vs. intro). These findings complicate our initial framework by showing that the old-school/new-school dichotomy, while useful as an analytical tool, cannot fully capture the complexity of producer choices that often transcend simple categorization. A producer's alignment with a particular tendency may correlate with a producer's familiarity and experience in mixing certain subgenres of rock and metal. Scheps spoke about being less familiar with extreme metal styles, which, in his opinion, led to less idiomatic and potentially less effective outcomes. Conversely, Otero specializes in technical death metal, where blast beats and rhythmically dense guitar riffs are commonplace. Listening to how blast beats are engineered makes this difference salient.

5. CONCLUSION

We hope this case study of metal music production serves as a primer that will help popular music analysts to include audio engineering in their analyses. Such work need not radically reinvent the wheel; it can integrate production qualities into traditional observations such as harmony, melody, rhythm, and form. How that can be done is reflected in the reasons why such work is advantageous. As we argued in the second section, producers' timbral nuances reflect and inform the norms and expectations of individual subgenres as they exist on the record, the definitive text within popular music. Accordingly, analysts might ask how those norms are served or thwarted by production choices as much as songwriting elements. Exeter's rendering the sub-bass inactive during the chorus, for instance, opened space for a more energetic double kick characteristic of climactic sections.

It is for this reason that our article highlights loudness and virtual space as two major areas for inquiry. With loudness, we noted how global loudness levels do not always tell the whole story. Even with loudness normalization, producers can use distortion to sound brighter and seem louder, shedding light on the importance of high-frequency concentrations across a mix and within individual instrumental signals. With virtual space, broad comparisons between mixes can use tools like the sound-box or soundstage to communicate subjective impressions of width and depth. Such spatial representations use cognitive metaphors and embodied experiences to connect production choices to aspects of music cognition. They show how producer decisions are not arbitrary nor wholly relative but instead, tap into an intercultural need to interpret sounds according to a multi-dimensional environment of sound sources.

Given access to the individual recorded tracks of an instrument, analysts can use stereo-width overlays to reveal panning splits and asymmetries – sometimes with separate guitars with different tones, sometimes with one instrumental signal, split and modified. Annotated screenshots of digital audio workstations (DAWs) and audio processor plugins show professional techniques and allow empirical comparisons of audio treatments on micro and macro scales. Even without access to the tracks and stems that a producer uses, analysts can employ spectrograms, two-dimensional spectra, and waveform diagrams to draw their attention to nuances of sound that might otherwise be missed. Add to all this the possibility of ethnography and one has a powerful cross-reference between sound traces and artist statements to corroborate analytical intuitions and observations.

Because listening impressions are contingent on playback equipment, individual experience, and personal idiosyncrasies, learning about the conventions and working patterns of music production can strengthen analysts' identifications of production strategies used on a record. As Moylan states, one might not be able to hear or listen for certain aspects of sound if one does not know those exist.⁸⁵ We thus encourage analysts to ask questions about width, depth, and height in non-traditional ways. What kind of panning occurs? How is it divided among sounds, and are purposeful asymmetries present that draw out certain aspects of musical expression? Which instruments or sound effects are most salient at what times and in what combinations? In addition to global observations about harmony and texture, one can add detailed observations about timbral blending and relationships of foreground/background, particularly regarding whether a recording recreates a live experience or seeks to transcend it. Height may be familiar through metaphors of high and low pitch, but it can also direct analysts to genre-oriented metaphors such as heaviness, particularly in relation to how producers fill out the frequency regions occupied by different signals. For instance, we highlighted the similarity to voice-crossing that can sometimes occur when producers design guitar tones to sound below the bass. We also highlighted the “meta-instrument” approach of Odeholm, who sought to take over from a live band and treat its component parts as ingredients for a single, transcendent, heavy sound.

Odeholm's hyperreal meta-instrument, comprising kick, bass, and guitar, is a clear example of how the studio can be an instrument in its own right. The V-shaped panning settings are another, showing how a producer can shape not just the signal of individual instruments but also the track as a whole. Altogether, these opportunities for strategically altering and enhancing a recorded song show how much power and control a producer can assume over the final product. In his efforts to transcend the traditional rock band, Odeholm becomes akin to a spirit possessing the musicians.

Odeholm is perhaps an extreme case of agency, but Moylan has strongly argued that one should generally assume producer agency on a record:

Given the high degree of control and scrutiny available in the production process, we should believe that what is sonically present on a record is what was intended, even if seemingly arbitrary or flawed. [...] [T]he sounds of the final record are carefully crafted; *what is there was put there, or was allowed to remain.*⁸⁶

85 Moylan 2002, 98.

86 Moylan 2020, 11, emphasis added.

Our findings offer an important qualification to this perspective, highlighting the tension between intentional design and pragmatic compromise that characterizes professional production practices. We do not believe that every aspect of the recording is intentional, and we see examples during “In Solitude” where producer differences were not necessarily deliberate or strategic but rather accidental or the consequences of production moves they made elsewhere in the song. While those differences were certainly “allowed to remain,” ethnography helps nuance that agency. During interviews, producers revealed that sometimes they meant to add automation but later forgot or could not be bothered, reflecting a familiar degree of aleatoricism built into our imperfect human tendencies. At times, the relatable element of deadlines introduced imperfections. “If I had another week, I would have done...” was a common refrain. Producers have a certain quality standard, but no one was perfectionistic.

We foresee several potential directions for future work. Genre conventions seem especially fruitful. In our case study, we observed several differences between how producers treated vocal layers during the breakdown section that variously highlight relationships to different genre influences. The analytical tradition of *Formenlehre* is another. Music analysts concerned with how listeners experience sectional boundaries and formal archetypes can investigate how the different instrumental layers within a track influence the perception of beginning, continuing, or ending. Finally, we see the nuances involved in production comparable to the artistic decisions involved in orchestration. Future research might build on the cognition-oriented taxonomy of orchestral grouping effects by Stephen McAdams and colleagues.⁸⁷ Following their work, it may investigate how producers blend or separate audio signals to achieve special effects, play with subgenre codes, or achieve other aesthetic ideals. It could consider timbral modifications by audio engineers as well as compositional and arrangement factors involving instrument choice, tuning, voice-leading, rhythm, and synchronicity. Because music production is such a common contributor to popular music recordings, an amalgamation of production strategies and traditional components of analysis carries a broadly versatile utility across popular music as a whole.

Acknowledgment

This research was generously funded by the Arts and Humanities Research Council (AHRC) in the United Kingdom under grant AH/T010991/1.

Data Availability Statement

The data from the entire HiMMP research project, including the multi-track of the song “In Solitude” and the producer versions (full bounces and vocal/instrument stems), is available in a publicly accessible dataset: <https://doi.org/10.34696/9s05-wv03>.

87 McAdams et al. 2022.

References

- Antares. 2024. "A Beginner's Guide to Dynamic Range in Audio Production." *Antares Tech* (23 Jan). <https://www.antarestech.com/community/a-beginners-guide-to-dynamic-range-in-audio-production> (27 Sep 2024)
- Benjamin, Walter. 1969. "The Work of Art in the Age of Mechanical Reproduction." [1935] In *Illuminations: Essays and Reflections*, edited by Hannah Arendt. Reprint, translated by Harry Zorn. New York: Schocken Books, 1–26.
- Brackett, David, ed. 2014. *The Pop, Rock, and Soul Reader: Histories and Debates*. Third edition. Oxford: Oxford University Press.
- Chanan, Michael. 1995. *Repeated Takes: A Short History of Recording and Its Effects on Music*. London: Verso.
- Cunningham, Mark. 1996. *Good Vibrations: A History of Record Production*. Chessington: Castle Communications.
- Dockwray, Patricia / Allan F. Moore. 2010. "Configuring the Sound-Box 1965–1972." *Popular Music* 42/3: 181–197.
- Fenton, Steve / Hyunkook Lee. 2019. "A Perceptual Model of 'Punch' Based on Weighted Transient Loudness." *Journal of the Audio Engineering Society* 67/6: 429–439.
- Frith, Simon / Simon Zagorski-Thomas. 2012. *The Art of Record Production: An Introductory Reader for a New Academic Field*. Farnham: Ashgate.
- Herbst, Jan-Peter. 2014. *Netzwerk Sound: Eine didaktische Herausforderung der populären Musik*. Augsburg: Wissner.
- Herbst, Jan-Peter. 2016. *Die Gitarrenverzerrung in der Rockmusik: Studien zu Spielweise und Ästhetik*. Münster: LIT.
- Herbst, Jan-Peter / Karl Bauerfeind. 2021. "Teutonic Metal: Effects of Place- and Mythology-Based Labels on Record Production." *International Journal of the Sociology of Leisure* 4: 291–313.
- Herbst, Jan-Peter / Mark Mynett. 2021. "Nail the Mix: Standardization in Mixing Metal Music?" *Popular Music and Society* 44/5: 628–649.
- Herbst, Jan-Peter / Mark Mynett. 2022a. "Toward a Systematic Understanding of 'Heaviness' in Metal Music Production." *Rock Music Studies* 10/1: 16–37.
- Herbst, Jan-Peter / Mark Mynett. 2022b. "What is 'Heavy' in Metal? A Netnographic Analysis of Online Forums for Metal Musicians and Producers." *Popular Music & Society* 45/5: 633–653.
- Herbst, Jan-Peter / Mark Mynett. 2023. "'I Just Go with What Feels Right.' Variance and Commonality in Metal Music Mixing Practice." *El Oído Pensante* 11/1: 4–31.
- Herbst, Jan-Peter / Mark Mynett. 2025a. "Aesthetic Tensions in Metal Music Production: Genre Expectations, Technological Mediation." *Popular Music and Society* 49/1. <https://doi.org/10.1080/03007766.2025.2530807>
- Herbst, Jan-Peter / Mark Mynett. 2025b. *Heaviness in Metal Music Production, Vol. 1: How and Why it Works*. Abingdon: Routledge.

- Herbst, Jan-Peter / Mark Mynett. 2025c. *Heaviness in Metal Music Production, Vol. 2: Learn from the Masters*. Abingdon: Routledge.
- Herbst, Jan-Peter / Alexander Vallejo. 2023. *Rock Guitar Virtuosos: Advances in Electric Guitar Playing, Technology, and Culture*. Cambridge: Cambridge University Press.
- McAdams, Steven / Meghan Goodchild / Kit Soden. 2022. "A Taxonomy of Orchestral Grouping Effects Derived from Principles of Auditory Perception." *Music Theory On-line* 28/3. <https://doi.org/10.30535/mto.28.3.6> (29 Jun 2025)
- Moore, Allan F. 2001. *Rock: The Primary Text: Developing a Musicology of Rock*. Second edition. Farnham: Ashgate.
- Moore, Allan F. / Patricia Schmidt / Ruth Dockwray. 2011. "A Hermeneutics of Spatialization for Recorded Song." *Twentieth-Century Music* 6/1: 83–114.
- Moorefield, Virgil. 2010. *The Producer as Composer: Shaping the Sounds of Popular Music*. Cambridge (MA): MIT Press.
- Moylan, William. 2002. *The Art of Recording: Understanding and Crafting the Mix*. Amsterdam: Elsevier.
- Moylan, William. 2020. *Recording Analysis: How the Record Shapes the Song*. London: Routledge.
- Moylan, William / Lori Burns / Mike Alleyne, eds. 2023. *Analyzing Recorded Music*. London: Routledge.
- Mynett, Mark. 2017. *Metal Music Manual: Producing, Engineering, Mixing, and Mastering Contemporary Heavy Music*. Abingdon: Routledge.
- Mynett, Mark. 2019. "Defining Contemporary Metal Music: Performance, Sounds and Practices." *Metal Music Studies* 5/3: 297–313.
- Mynett, Mark. 2020. "Maximum Sonic Impact: (Authenticity/Commerciality) Fidelity-Dualism in Contemporary Metal Music Production." In *The Bloomsbury Handbook of Music Production*, edited by Simon Zagorski-Thomas / Andrew Bourbon. London: Bloomsbury, 293–302.
- Robjohns, Hugh. 2016. "Q. What are My Phase-Correlation Meters Telling Me?" *SOS: Sound on Sound*. (October) <https://www.soundonsound.com/sound-advice/q-what-are-my-phase-correlation-meters-telling-me> (31 Mar 2025)
- Schubert, Emery / Wolfe, Joe. 2006. "Does Timbral Brightness Scale with Frequency and Spectral Centroid?" *Acta Acustica* 92: 820–825.
- Smialek, Eric. 2016. "The Unforgiven: A Reception Study of Metallica Fans and 'Sell-Out' Accusations." In *Global Metal Music and Culture: Current Directions in Metal Studies*, edited by Andy Brown / Karl Spracklen / Keith Kahn-Harris / Niall Scott. London: Routledge, 106–124.
- Strachan, Robert. 2017. *Sonic Technologies: Popular Music, Digital Culture and the Creative Process*. London: Bloomsbury.
- Théberge, Paul. 1997. *Any Sound You Can Imagine: Making Music/Consuming Technology*. Middletown: Wesleyan University Press.

- Théberge, Paul. 2001. "‘Plugged In’: Technology and Popular Music." In *The Cambridge Companion to Pop and Rock*, edited by Simon Frith, Will Straw and John Street. Cambridge: Cambridge University Press, 3–25.
- Thomas, Niall / Andrew King. 2019. "Production Perspectives of Heavy Metal Record Producers." *Popular Music* 38/3: 498–517.
- Thompson, Paul. 2019. *Creativity in the Recording Studio: Alternative Takes*. Basingstoke: Palgrave.
- Walther-Hansen, Mads. 2020. *Making Sense of Recordings: How Cognitive Processing of Recorded Sound Works*. Oxford: Oxford University Press.
- Waksman, Steve. 1999. *Instruments of Desire: The Electric Guitar and the Shaping of Musical Experience*. Cambridge (MA): Harvard University Press.
- Walser, Robert. 1993. *Running with the Devil: Power, Gender, and Madness in Heavy Metal Music*. Middletown: Wesleyan Press.
- Walser, Robert. 2003. "Popular Music Analysis: Ten Apothegms and Four Instances." In *Analyzing Popular Music*, edited by Allan F. Moore. Cambridge: Cambridge University Press, 16–38.
- Walser, Robert. 2007. "The Poetics of Rock: Cutting Tracks, Making Records by Albin J. Zak." [Review] *Journal of the American Musicological Society* 60/1: 253–257.
- Wanasek, Joel. 2017. "Low-End Mastery." *Unstoppable Recording Machine*. <https://members.urm.academy/courses/low-end-mastery> (27 Aug 2024)
- Zagorski-Thomas, Simon. 2010. "The Stadium in Your Bedroom: Functional Staging, Authenticity and the Audience-led Aesthetic in Record Production." *Popular Music* 29/2: 251–266.
- Zagorski-Thomas, Simon. 2014. *The Musicology of Production*. Cambridge: Cambridge University Press.
- Zagorski-Thomas, Simon / Katia Isakoff / Sophie Stévançe / Serge Lacasse. 2020. *The Art of Record Production. Vol. 2: Creative Practice in the Recording Studio*. Abingdon: Routledge.
- Zak, Albin III. 2001. *The Poetics of Rock: Cutting Tracks, Making Records*. Berkeley: University of California Press.

Discography

AC/DC. 1980. "Back in Black." *Back in Black*. New York: Atlantic.

Data

- Bogren, Jens. HiMMP interview, in-person, 11 January 2023. Fascination Street Studios, Örebro, Sweden, transcribed by Jan-Peter Herbst.
- Exeter, Mike. HiMMP interview, in-person, 26 February 2023. Greenroom Studios, Huddersfield, United Kingdom, transcribed by Jan-Peter Herbst.

Getgood, Adam. HiMMP interview, in-person, 5 June 2023. J and J Studio, Bristol, United Kingdom, transcribed by Jan-Peter Herbst.

Middleton, Josh. HiMMP interview, in-person, 31 May 2023. MDL Tone Studio, Reading, United Kingdom, transcribed by Jan-Peter Herbst.

Nordström, Fredrik. HiMMP interview, in-person, 9 January 2023. Studio Fredman, Gothenburg, Sweden, transcribed by Jan-Peter Herbst.

Odeholm, Buster. HiMMP interview, in-person, 12 December 2023. Atlantis Studios, Stockholm, Sweden, transcribed by Jan-Peter Herbst.

Otero, Dave. HiMMP interview, in-person, 18 December 2023. Flatline Audio Studios, Denver, United States, transcribed by Jan-Peter Herbst.

Scheps, Andrew. HiMMP interview, in-person, 31 August 2022, Rockfield Studios, Monmouth, United Kingdom, transcribed by Jan-Peter Herbst.

© 2025 Jan-Peter Herbst (j.herbst@hud.ac.uk, ORCID iD: 0000-0001-7453-0141), Eric Smialek (E.Smialek@hud.ac.uk, ORCID iD: 0000-0002-2111-2322)

University of Huddersfield

Herbst, Jan-Peter / Eric Smialek. 2025. "‘Mixed’ Results: An Introduction to Analyzing Music Production through Eight Commissioned Metal Mixes." *Zeitschrift der Gesellschaft für Musiktheorie* 22/1, 61–97. <https://doi.org/10.31751/1222>



Dieser Text erscheint im Open Access und ist lizenziert unter einer Creative Commons Namensnennung 4.0 International Lizenz.

This is an open access article licensed under a Creative Commons Attribution 4.0 International License.

eingereicht / submitted: 29/09/2024

angenommen / accepted: 18/02/2025

veröffentlicht / first published: 10/07/2025

zuletzt geändert / last updated: 22/07/2025