Audio Engineering Society

Convention Paper 10297

Presented at the 147th Convention
2019 October 16 – 19, New York

Towards a Pedagogy of Multitrack Audio Resources for Sound Recording Education

Kirk McNally[1], Paul Thompson[2], and Ken Scott[2]

[1]*University of Victoria, Victoria, BC, Canada*
[2]*Leeds Beckett University, Leeds, UK*

Correspondence should be addressed to Kirk McNally (kmcnally@uvic.ca)

ABSTRACT

This paper describes preliminary research into pedagogical approaches to teach and train sound recording students using multitrack multitrack audio recordings. Two recording sessions are described and used to illustrate where there is evidence of technical, musical and socio-cultural knowledge in multitrack audio holdings. Approaches for identifying, analyzing and integrating this into audio education are outlined. This work responds to the recent AESTD 1002.2.15-02 recommendation for delivery of recorded music projects, and calls from within the field to address the advantages, challenges and opportunities of including multitrack recordings in higher education teaching and research programs.

1 Introduction

Guitarist, Lester William Polsfuss (Les Paul), was one of the first artists to recognize the technological and artistic value of the multitrack format. Since Les Paul's experimental recordings in the 1950s, multitrack audio recordings have become a critical component of nearly every recorded musical work. However, their acknowledgement as useful resources for researchers, and the potential to use them within audio education contexts to teach aspects of audio production and sound recording, has not been fully recognized. Historically this was due to the commercial nature of these recordings, which limited or denied access to them. Multitrack audio resources have been used in post-secondary sound recording programs to provide students opportunities to practice mixing or re-mixing, editing, vocal tuning, and drum augmentation, or to learn production skills such as recording individual voices or instruments for 'karaoke' style projects or creating sound-alike recording projects. However, the majority of multitrack audio holdings previously available, such as those found in the popular *'Mixing Secrets'* collection[1], have been edited and consolidated for ease of distribution and use, and generally don't include alternate takes, or out-takes, edit decision lists (EDL), or studio documentation.

Today, this situation is changing with new, more comprehensive multitrack sources becoming available, including commercial enterprises[2], the Sigma Sound

[1]http://www.cambridge-mt.com/MixingSecrets.html
[2]eg. https://weathervanemusic.org/shakingthrough

Archive at Drexel University [1], the EMI Music Canada Archive at the University of Calgary, the Open Multitrack Testbed [2]. These resources, in combination with calls from within the profession to revise multitrack delivery practices [3] present new opportunities for audio educators to use primary source multitracks in the classroom. In this paper we show the value of using multitrack collections that include alternative or out-takes for educational purposes; using audio analysis as a means of illustrating where there is evidence of technical, musical and socio-cultural knowledge in these multitrack holdings.

2 Background

In her 2016 study of how producers and engineers prepare, manage, and direct recording sessions, Pras concluded with a call for more studio observations and systematic research into the technical practices of recording engineers, to gain greater insight into the processes of sound recording [4]. In his study of how theoretical research can meet the challenges of pedagogy in the field of record production, Zagorski-Thomas further concluded that new, more complete (and annotated) records of studio and performance activity can enable new directions for research and fresh pedagogical approaches [5]. Recent research by McNally et. al [6] examined how multitrack audio resources are currently being used in the classroom at three post-secondary educational institutions, but underline some limitations of the publicly available holdings and the need for useful frameworks to help students and educators interpret and apply the knowledge they represent. Ethnographic methods have proven useful for such studies of music production [7][8], and do provide a more complete record of studio activity, but the focus to date has been on the video record, with the multitrack resources being constrained by standard workflows, with set-up, outtakes and alternate takes often not being preserved.

In this study, a set of multitracks were created with the specific goal of providing a more complete record of the recording session; in particular the value of having out-takes and alternative takes, not just the final versions from recording sessions. The work focuses on multitrack materials as a resource to identify and communicate information regarding three categories of interrelated and co-dependant music production knowledge: technical, musical and socio-cultural. They are only separated here for the purposes of analysis but in practice are intricately connected with one form of knowledge influencing each other within a system of circular causality.

2.1 Technical knowledge and Multitracks

Technical knowledge involved in the domain of commercial record production generally relates to the knowledge of recording technologies and the ways in which they can be deployed in the record-making process [9]. Technical domain knowledge is primarily needed for the task of engineering, and in some musical styles where music or recording technologies are central to their production or performance, technical knowledge may also be needed for the tasks of producing, songwriting or performing (i.e Hip-hop or dance musics).

Technical knowledge of the domain in the recording studio includes the broad area of acoustics, microphones, electronics and digital audio [9]. Inside the commercial recording studio, a recording engineers is expected to:

- operate the control room equipment,
- be familiar enough with electrical signal flow to trouble-shoot the inevitable broken signal path or feedback loop,
- possess at least a rudimentary knowledge of acoustics in order to make informed judgments about how sounds will translate from one listening environment to the next,
- master the intricate processes of audio recording ranging from microphone (mic) selection and placement to ‘building a mix’ step-by-step from performances captured on tape or disk.

Adapted from Porcello, 2004 [7]

Combining different types of technical knowledge, such as room acoustics, instrument acoustics, microphone characteristics, microphone placement, and audio processing is necessary to create ‘sounds’ that achieve a specific production aesthetic, or to create sonic characteristics of a particular musical style.

2.2 Musical Knowledge and Multitracks

The central element of the musical knowledge within commercial record production typically involves the contemporary Western popular song, which includes a knowledge of lyrics, melody, rhythm, harmony, the structure of songs and their arrangement in addition to "various production elements that affected the nature of the song's reception and an understanding of audiences' possible interpretations of the work produced" [10]. The application of this knowledge can be seen in action with commercially successful songwriters and producers who write with the final sound of the recording in mind. Audio engineers require a working knowledge of Western contemporary songs and their elements because their contributions can dictate the way in which the final recording sounds [9]. For example, song elements and musical features inform some of the technical processes of commercial record production, such as mixing, so musical knowledge enables engineers to "understand performers', arrangers', and producers' discussions" [7].

As well as popular song knowledge, recording engineers and producers require knowledge of the symbolic rules, traditions and practices of the related domain of musical performance, which is intrinsically part of multitrack recordings. The significant difference, however, is that these performances were captured without an audience in its traditional sense and they were assessed, rejected and selected with a consideration for tuning, timing or any other deficiency within the performance of the song. Multitracks also hold knowledge of arrangement, both in terms of micro musical elements (melody, harmony etc.) and macro musical elements (arrangement, form, etc.) and the ways in which the musical elements unfold over time. The final recording is therefore orchestrated through the combination of performance, arrangement and technical mediation in which the recording consists of the song, the musical arrangement and the track. Consequently, listening to a finished recording is the experience of hearing "both the song and the arrangement" [11], which illustrates the interconnectedness and mutual interdependence of these critical areas of knowledge [9]. Within a multitrack, the elements of musical performance and technical mediation combine through the processes of songwriting, engineering and production to create a resultant 'sound.'

2.3 Socio-cultural Knowledge and Multitracks

The sociocultural knowledge needed to create multitrack audio materials involves an array of established "practices set within a material context" [12]. This material context is the social context of the recording studio in which its traditions and cultural practices are produced and enacted. Sociocultural knowledge is enacted using other types of knowledge and underlines the superficial disconnection between the other types of knowledge involved; technical and musical, which are interconnected and codependent.

Socio-cultural knowledge isn't always immediately evident within multitrack audio materials as it often involves building and maintaining relationships with all those involved or the use of specific language or terminology to discuss musical sounds and translate sonic descriptions into technical actions. However, making a final recording is simply more than packaging a series of pre-existing sounds [13] and the ideology of a recorded musical style can be seen as a combination of the musical, technical and cultural aspects. The implementation of particular practices of ensemble recording, overdubbing, editing, splicing, compiling, altering timing or tuning, in order to create an 'ideal event' [14] are often evident in multitrack audio recordings and can show how sociocultural knowledge is enacted during a particular recording studio session and within a musical style.

3 Methods

3.1 Materials

Two recording projects, one at the Warehouse Studios (WS) in Vancouver, Canada and the second at Leeds Beckett University (LBU), in the UK, provided the multitrack materials for this study. Kirk McNally was the recording engineer for the WS project and Ken Scott engineered for the LBU project. For the WS project, all audio routed to the multitrack recorder (ProTools HD, 96kHz, 24bit) was duplicated and recorded on a second system to provide a time-of-day multitrack for all session activity including microphone set-up, rehearsal and recording. Because of system and infrastructure limitations, this procedure was not duplicated for the LBU project. For both the WS and LBU projects, a time-of-day two-track audio recording was made of the console monitor output.

In addition to these audio recordings, each session was video recorded using several GoPro *Hero Four* cameras (720P, 30FPS) positioned around the studio control room and live rooms to capture the conversations, gestures and movements of the studio participants. Field notes, semi-structured interviews with session participants, and the use of autoethnography by McNally for the WS project, represents all other materials used in this study.

3.2 Video Analysis

With the sheer volume of data collected for this study (36+ linear hours of video footage), it was first necessary to identify what sections held pedagogical value. The first step in this process was identifying critical incidents, which are defined as those moments that surprise or stand out during an activity, or resonate later, upon reflection [15]. A combination of field notes, post-project semi-structured interviews and reflections were used to identify critical incidents and provide a means to focus the video analysis, and highlight the collectively recognized significant moments from the WS and LBU projects. With these sections of the project identified the video segments were then analyzed to identify what type of activity was taking place and if there was verbal communication present between project participants that could be used to help interpret the activity.

Using the work of Lefford and Thompson [16] on metacognition in the music studio as an initial framework, a coding scheme was developed to categorize how McNally and Scott evaluated their own cognition or thinking strategies during the recording projects. Language was coded into three categories, including ***valuations*** eg. 'I like it;' ***directions*** eg. 'Can you change microphone X to Y;' and ***questions*** eg. 'How do you feel about doing that again?' or 'Do you like that sound'. This analysis provided both rich pedagogical material, as well as clear direction for how the audio should subsequently be analyzed and evaluated. For example, if a *direction* was given to change a microphone, the analysis sought to measure and evaluate the result of this change. In the cases where no verbal communication was present, autoethnography (McNally) or semi-structured interviews (Scott) were used determine the intrinsic motivation or the basis for individual decisions or actions present in the selected video sections.

3.3 Audio Analysis

Audio analysis was performed using the Sonic Visualizer[3] application and associated *VAMP* plugins. The analysis conducted for this study included; time (*beat tracker* [17]), pitch (*melody extraction* [18]), and spectrum (*loudness, bark coefficients and spectral sharpness*). The spectral features used for this study were chosen because they have been identified as having perceptual relevance [19]. The purpose of the audio analysis was to provide evidence of changes made to the recorded audio as a result of either musical direction or technical mediation by the recording engineer. As such, absolute values measured for audio features between the two projects were of no consequence, and no attempt to characterize the individual sounds should be implied through this analysis.

4 Results

4.1 Illustrating the Technical

A significant difference between the recording techniques used by McNally and Scott to record kick drums was observed, with the minimalist approach of Scott being noted by students attending the LBU session and further highlighted in McNally's autoethnography. Field notes from the WS and LBU sessions show that McNally used a combination of two microphones, console equalization, hardware compression and a further external hardware equalizer, where Scott used a single microphone and external channel strip with equalization to record the kick drum.

Video analysis of the 'development of the drum sounds' sections of the sessions includes directions by both McNally and Scott regarding the patching of desired hardware equipment and routing options. However, the session activity that follows the set-up and refinement of drum recording techniques includes no further communication by Scott on the topic. In the WS project video, McNally questions the drummer about the drum sound, eliciting a positive response, "Yeah, I think they're good," and the statement that, "I won't know until the rest of the band is playing." Audio analysis of the session multitracks revealed that no further alterations were made to the drum sounds for the remainder of both the WS and LBU sessions. This supports the assessment that this activity is principally

[3] https://www.sonicvisualiser.org

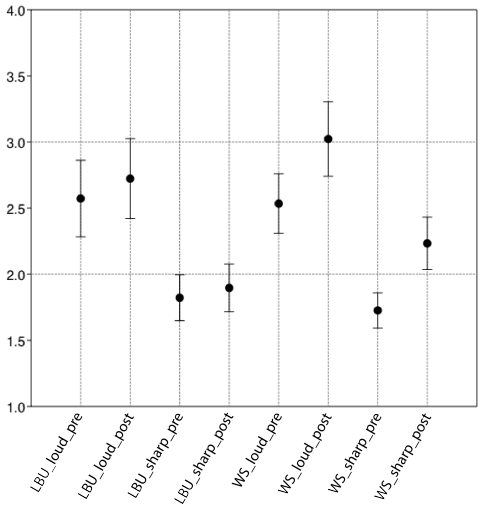

Fig. 1: Mean and whisker plot of measured loudness and spectral sharpness for LBU and WS kick drums, pre- and post-processing.

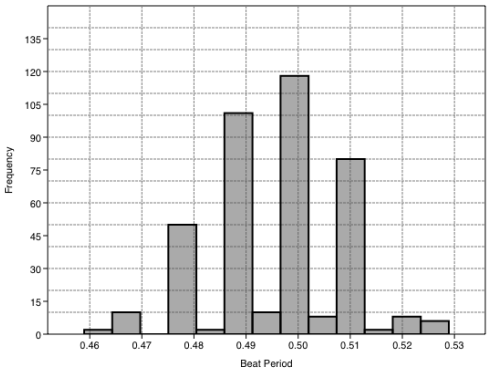

Fig. 2: Histogram showing frequency of beat periods for *Know Your Name - TK1*.

evidence of technical knowledge at work in the studio, with the lack of changes to the drum sounds indicating they were assessed as being appropriate, both musically and socio-culturally, for the music being recorded.

To illustrate the technical changes made in developing the kick drum sounds, samples were isolated and extracted from the session multitracks, pre- and post-processing. Fig. 1 illustrates the changes made, as represented by measures of loudness and spectral sharpness. In both cases, the relative mean loudness and spectral sharpness values increased, with the WS kick drum showing a greater change for both audio features, which corresponds to the nature of the processing applied by McNally.

4.2 illustrating the Musical

The decision by Scott to change the chorus tempo of the LBU project song was identified by all participants as having a significant affect on the finished work. Video analysis of the scene where this decision was made, includes Scott's direction to the drummer that,"We need to change tempo" and, "I would say two beats per minute [increase tempo]." Using a summed mix of the kick and snare drum from the LBU session multitracks and the *BeatRoot Beat Tracker*[17] plugin, it was possible to extract and analyze variation in beat periods, defined here as the time between detected beats in milliseconds, over the course of the song and across different takes.

The histograms in Fig. 2 and 3 are used to show the frequency of beat periods, following the chorus tempo change, for the first and 'master' drum performances. This analysis illustrates the synthesis of the tempo change by the drummer into his performance. The 'master' take showing a bimodal distribution that indicates a more secure and accurate realization of what Scott had requested.

4.3 Illustrating the Cultural

A small, but effective guitar overdub was identified as a useful to illustrate how cultural knowledge is enacted in the recording studio. While preparing for the guitar overdub, a band member questions the guitarist, "are you going to do your Creedence part?". Previous discussions between McNally and the guitarist had determined that a *Fender Champ* guitar amplifier would be used for this particular overdub. Upon first playback of the guitar sound in the control room, McNally can be heard stating, "Nice!" and a band member, "Perfect!." In the following video sequence, the inspirational musical example, Creedence Clearwater Revial's *Up Around The Bend* is played by band members in the control room, and discussion moves to an evaluation of

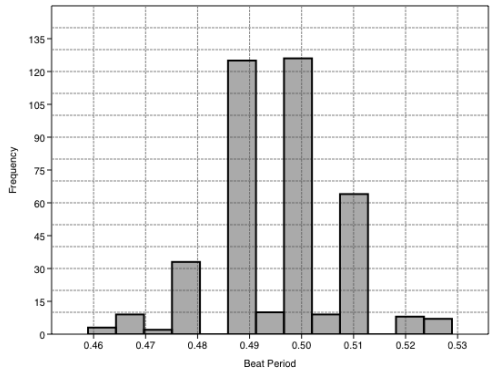

Fig. 3: Histogram showing frequency of beat periods for *Know Your Name - MSTR*.

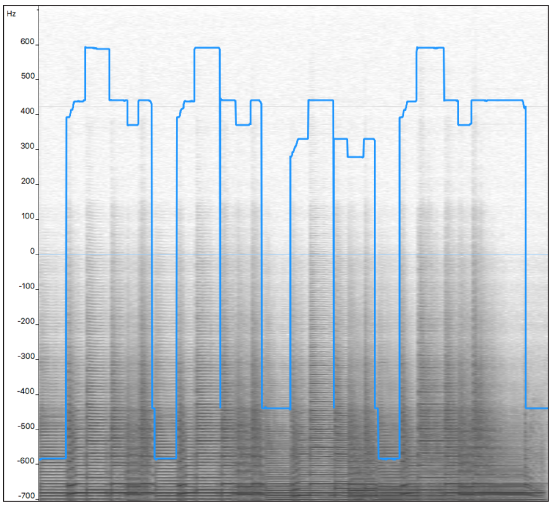

Fig. 4: Spectrogram and plot of pitch analysis for WS 'CCR' guitar.

the musical performance, with the musical quotation being played via a cellphone to the guitarist.

In the absence of either direction or questioning by session participants following the establishment of the guitar sound, the conclusion is that the musical quotation sufficiently conveyed the desired socio-cultural meaning. Fig. 4 shows a graphical representation of the guitar melody, extracted from the recording of the guitar direct-input, overlaid on a spectrogram of the guitar amplifier recording.

5 Discussion

To gain further insight into the nature of the session activity, a semi-structured interview was conducted with Scott following the LBU project. In response to a question concerning the decision to adjust the chorus tempo, as detailed in Section 4.2, Scott indicated the change, "tends to make it more human" which mirrors how musicians naturally speed up at exciting moments, and that he thought it would, "feel better." When asked about the process he used to chose a master drum performance, he answered that it should, "feel good and have all the parts right...as we wanted them." Similarly, in his response to a question regarding the process for selecting vocal takes, he stated,"It's [about] feel, it's my personal impression of what works best for the song...how do you describe personal taste? You can't really." This sentiment is echoed in his response to questioning over the microphone selection process for the LBU project vocal recording, where he states, "I picked the one I thought sounded the best" but with the caveat, "we all like different things."

Scott's responses are illustrative of tacit knowledge - the unarticulated, implicit knowledge gained from practical experience [20] [21], which Susan Schmidt-Horning applies to sound recording and describes as the ability of recording engineers to "deploy a working knowledge of the behavior of sound and the machinery of its propagation." Using only the video recordings to analyze the activity and decision-making therefore may not uncover the entirety of the knowledge and rationale involved and, so the question remains, how do we as audio educators access and use this knowledge and experience?

The analysis from Section 3 points towards a methodology for accessing this information, with the condition that alternative takes of the recordings are available. Faced with a particular recorded sound or element of a multitrack audio tape, this analysis approach allows the audio student to first evaluate if the sound has been changed technically, or musically, e.g. through the identification of changes to equalization, compression, tempo, performance, etc. The trajectory of any changes made, provides insight into the initial assessment by the original recording engineer. If the student then inserts themselves into the decision process, using the

three language categories: valuation, direction or question, they are immediately invited to critically evaluate the basis for their own decisions. For example, if the student dislikes a sound they can ask themselves if it is because of a technical feature? Can they point to a socio-cultural reference that better illustrates the sound they would like to hear? Can they reference an example from their own work that is more effective?

Of course, any decision the student makes is purely hypothetical, but it still requires them to critically engage with the three types of knowledge used in the recording studio and to practice how this knowledge is enacted during a recording session.

The study presented in this paper is part of a larger research project investigating a variety of themes regarding multitrack audio resources. This preliminary work has begun to illustrate how multitrack audio resources can be employed beyond their current use of mixing and remixing, etc. Further work is needed in this area to fully explore the multitude of ways multitrack audio resources can help to not only train and educate trainee audio engineers, but the ways in which multitrack audio archives can be mined for musical, technical and socio-cultural knowledge, and the ways in which creative decisions can be determined from the sound and arrangement of the recorded audio.

6 Summary

In this paper we have used two newly generated multitrack audio holding to illustrate where recording engineers use musical, technical and socio-cultural knowledge within the studio. The value of creating or obtaining multitrack holdings for educational programs that include alternate takes or outtakes is demonstrated by way of audio analysis, and a methodology students can use to evaluate this type of resource and aid them in their personal skills development.

7 Acknowledgments

This work was made possible by the Social Sciences and Humanities Research Council of Canada Insight Development Grant *Creating Capacity for Research into Sound Recording and Music Production, 2018-2020*. The authors also wish to thank the School of Film, Music Performing Arts at Leeds Beckett University for the donation of recording studio time.

References

[1] Seay, T., "Primary sources in music production research and education: Using the Drexel University Audio Archives as an institutional model," *Journal on the Art of Record Production*, 5, 2011.

[2] De Man, B., Mora-Mcginity, M., Fazekas, G., and Reiss, J. D., "The open multitrack testbed," in *Audio Engineering Society Convention 137*, Audio Engineering Society, 2014.

[3] AESTD1002.2.15-02, "Recommendation for delivery of recorded music projects," Technical report, Audio Engineering Society, 2014.

[4] Pras, A., "What has been left unsaid about studio practices: How producers and engineers prepare, manage, and direct recording sessions," in *Music, Technology, and Education*, pp. 45–62, Routledge, 2016.

[5] Zagorski-Thomas, S., "How is theoretical research meeting the challenges of pedagogy in the field of record production?" in *Music, Technology, and Education*, pp. 85–100, Routledge, 2016.

[6] McNally, K., Seay, T., and Thompson, P., "What the masters teach us: Multitrack audio archives and popular music education," *The Bloomsbury Handbook of Popular Music Education: Perspectives and Practices*, p. 113, 2019.

[7] Porcello, T., "Speaking of sound: language and the professionalization of sound-recording engineers," *Social Studies of Science*, 34(5), pp. 733–758, 2004.

[8] Thompson, P. and Lashua, B., "Getting it on record: Issues and strategies for ethnographic practice in recording studios," *Journal of Contemporary Ethnography*, 43(6), pp. 746–769, 2014.

[9] Thompson, P., *Creativity in the Recording Studio: Alternative Takes*, Palgrave Macmillan, 2019.

[10] McIntyre, P., "Creativity and cultural production: A study of contemporary Western popular music songwriting," *Creativity Research Journal*, 20(1), pp. 40–52, 2008.

[11] Zak, A., *The poetics of rock: Cutting tracks, making records*, Univ of California Press, 2001.

[12] Barker, C., *Making sense of cultural studies: Central problems and critical debates*, Sage, 2002.

[13] Zagorski-Thomas, S., *The musicology of record production*, Cambridge University Press, 2014.

[14] Eisenberg, E., *The recording angel: Music, records and culture from Aristotle to Zappa*, Yale University Press, 2005.

[15] Messner, M. A., "Barbie girls versus sea monsters: Children constructing gender," *Gender & Society*, 14(6), pp. 765–784, 2000.

[16] Lefford, M. N. and Thompson, P., "Naturalistic artistic decision-making and metacognition in the music studio," *Cognition, Technology & Work*, 20(4), pp. 543–554, 2018.

[17] Dixon, S., "Automatic Extraction of Tempo and Beat from Expressive Performances," *Journal of New Music Research*, 30(1), pp. 39–58, 2001.

[18] Salamon, J. and Gómez, E., "Melody Extraction from Polyphonic Music Signals using Pitch Contour Characteristics," *IEEE Transactions on Audio, Speech and Language Processing*, 20, pp. 1759–1770, 2012, doi:10.1109/TASL.2012.2188515.

[19] Peeters, G., "A large set of audio features for sound description (similarity and classification) in the CUIDADO project," *CUIDADO IST Project Report*, 54(0), pp. 1–25, 2004.

[20] Polanyi, M., *Personal Knowledge: Towards a Post-Critical Philosophy*, University of Chicago Press, 1958.

[21] Polanyi, M., *The Tacit Dimension*, Doubleday, 1966.