Application Study

Rock Band

“In Tune and in Time”

One of the first responsibilities that a producer of popular music takes on is the requirement that the final product delivered to the record company be “in tune and in time.” That is to say, the musical performances are to realize a certain standard of technical proficiency in pitch and rhythm. The legacy of this central role for a producer may be found in the many rough performances that were a part of the early history of rock and roll. Along with a heavy reliance on attitude came some rather oblique relationships to musicianship on the part of some of the musicians. Thus, especially in the “band era” of the 1960s and 1970s, came the need for some QC (quality control) and the centrality of the producer’s role as the arbiter of traditional musical standards. The long-dreamed-of tools for relatively easy pitch and rhythm “fixing” have now arrived along with the DAW. In regards to the direction of technological influence (“top-down” versus “bottom-up”), the realization of this desire suggests that the pitch and rhythm tools of computer-based audio may be seen as a striking example of agency driving technology (we needed to fix stuff, and now we can). However, as we shall see through this volume, there are always elements working in both directions. Perhaps some of what is generally considered to be excessive “fixing” of rhythm and pitch may be considered to be examples of technology driving agency (we can fix stuff, so we do). In any case, the ease and degree of control over pitch and rhythm have dramatically changed, so the new paradigm of music construction is in full bloom when it comes to realizing the producer’s dictum that performances must be “in tune and in time.”
Prior to the current computer technology the producer had relatively few options in the control of intonation and rhythmic accuracy in musical performances. The primary tool was, after a studio take that wasn’t up to the desired standard, to get on the “talk-back” to the performing musician and say something like: “That was great, but can you do it one more time for me: It was a little pitchy” (meaning either too sharp or flat for use) or “It felt a little awkward” (meaning not good enough rhythmically for use). Here the final recorded performances were created through selective repetition. Bits of performances would be captured to allow a complete, musically acceptable performance to be pieced together. If the musician was very capable, then little or no such repetition would be necessary, but in the case of the relatively inexperienced rock-and-roll band member, this could be a long and tedious process. Over time certain techniques and technologies developed that could be applied after the performance, and these aided this process in small ways—tape editing, “flying in,” and later judicious use of a harmonizer could correct problems in certain instances—but these options were time-consuming and only successful in a very limited number of circumstances. For the most part getting the required performance out of the musician, sometimes one arduous bit at a time, was the only viable option.

To explore the changes in the application of control over tuning and timing I undertook a project involving a song that I had recorded in 1994 for the band “The Jenny Thing.” I had made the original recording using the dominant professional recording technology at the time, which was a 24-track analog tape recorder. This meant I had twenty-four individual tracks for recording on which to build the music for each song. The original sessions were carried out in the typical studio production style of the time. We recorded the initial “basic” tracks of drums, bass, guitar, and vocal together, but all of the performances other than the drum track were considered “scratch” tracks—that is, they were played as guide tracks to be later “scratched” or discarded in favor of new takes of these performances. All of the instruments, as well as each element of the drum set, were recorded on their own individual tracks. All of the instruments and the lead vocal were isolated from each other so that there wasn’t “bleed” from one sound into the recording of the other, facilitating the replacement of parts later. Additional parts such as lead guitar tracks, harmony vocals, and percussion tracks were added later. By recording each part at different times I was able to focus the attention and the process of revision on the execution of each individual performance until it was considered accept-
able. This was a relatively low-budget record, so the standard for “acceptable” had to take into consideration the capabilities of the musician along with the overall time that the budget allowed for the entire recording and mixing process. This remains the dominant procedure in pop music today, though the weight of performance control has shifted from being almost completely a part of the original recording to a balance between recording and the kind of postrecording manipulation that I was now going to apply to this production. My goal in this study is to apply to this older recording the process of rhythm and pitch “fixing” currently used in pop production, allowing me to compare the original master as it was released on record to what would probably comprise the master recording if this song were produced using contemporary technology.

My first task was to transfer from the analog tape to digital audio in the DAW. For the software needed to control the digital audio stored on the computer’s hard drive I used today’s dominant professional recording software, Avid’s Pro Tools. While making the transfer and listening to the original audio I noted a slight tempo fluctuation during the song’s introduction. I remembered being continually aggravated by this when I produced the original track. This was a case where I had deemed the inconsistency to be slight enough to be acceptable, though it was significant enough to have bothered me throughout the process. I smiled to myself knowing that now I would be able to “fix” this slight problem, and then thought of the countless number of examples of such occurrences in other recordings made before the current capabilities were available. How many slight problems in recorded performances have haunted musicians and producers before there was a means of correcting them as a part of the production process? But before I explore the meaning of such musical “fixing,” I provide a narrative of the process I undertook in applying contemporary production practices to this particular piece of music.

First, a caveat: I use the terms fix, correct, consistent, and the like as technical terms, while recognizing that these also carry significant implications about the value of the alterations being made. The reality is much more complex, for value in musical performance is most often ascribed to deviations from the standard to which we are “fixing.” The implied values of such words as fixing are not necessarily a reflection of how one might value the actual effect of this process. In fact they may be completely at odds with such implications (supposedly “fixed” performances may be considered inferior to the original). I will be addressing questions of value in this more general sense later in this chapter, but for the moment I ask
the reader to temporarily indulge the use of the language for the sake of the narrative.

**Fixing in Pro Tools**

As with most popular music, this song was recorded to a click track generated by a metronome. That is to say, the drummer listened to a click when laying his initial track while the other musicians played to the (click-informed) drummer’s performance. Using contemporary production techniques where we would be recording directly into the computer, with the click generated by the computer, the bar and beat information would already be an integral part of the recording and established before any music was played. With this historical recording I had an individual audio track with the click from the metronome recorded separately, but this was not integrated into the computer clocking function—Pro Tools wasn’t able to give a readout of the metronomic bar and beat information. In fact, because it was recorded using analog gear, the original metronomic timing was not perfect. Slight variations in the creation of the click by any analog clocking device (metronome), combined with minute variations in speed from an analog tape recorder, mean that it is not possible to simply assign the correct bpm (beats per minute) reading to the audio now in Pro Tools and have the music line up correctly with the bar and beat information. Fortunately there are tools to assist us in adjusting for these inconsistencies so that we can work within the traditional music organization of bars and beats along with the clocking precision of a computer.

In the Pro Tools program there is a plug-in tool called Beat Detective. As the name suggests, this tool investigates rhythmic qualities of audio data. It distinguishes beat information by identifying transients (high-frequency leading sound elements) that are likely candidates for marking the beginning of each beat. In this case, because I had the click track recorded on a separate audio track, it was an easy matter for Beat Detective to create a tempo map from the position of each click and thus organize the file into bars and beats. To do this Beat Detective assigns an exact tempo for each beat, to within three decimal points of bpms, thus yielding a bar and beat map that remains consistent with the original click. Beat Detective does not alter the placement of the beats, but it identifies and organizes them in a way that makes them conform to a bar and beat structure. By doing this I had a tempo map that represented the “ideal time” when the
performances were made. This was the “correct” beat structure that the drummer was conforming to when laying the initial drum track.

I then used another feature of Beat Detective to slice all of the various drum tracks into separate regions, setting a variety of parameters to help it make “intelligent” decisions about how to read the transients and divide the performance into various beat-related elements. As is typical in contemporary drum set recording, there were individual tracks for bass drum, snare drum, tom-toms, and hi-hat cymbals, as well as separate stereo recordings of overhead microphones to capture the cymbals, and room mics to capture the overall sound of the drums in the room. Beat Detective processed each track separately. Using Beat Detective on overhead and room tracks is difficult because of the complexity of the information. As sophisticated as Beat Detective is, it has trouble determining beat divisions when the audio consists of all of the drum instruments mixed together.

Once Beat Detective had created individual regions from each track of the drum performance, I used the “quantize” function to correct the timing of the drum performance. Quantizing takes the beginning of each separated region of audio and moves it along the musical timeline to the beginning of the nearest user-defined beat subdivision (in this case the smallest subdivision was eighth-note triplets, as this song used a “shuffle” or triplet subdivision of the beat). Quantizing each track individually yielded the most accurate results, but it also meant that where there were inconsistencies in the quantizing process between individual tracks I would have to make manual changes for the parts to conform to each other. While this process created a much more accurate version of the original drum performance in terms of note placement relative to the “ideal” of metronomic time, there was still considerable variation in the volume and timbre of each individual sound, as well as internal variations within the larger segments that were quantized into position. Thus the resulting performance was not the same as a performance coming from a drum machine, where every note may be metronomically placed and there is generally little or no variation in dynamics or timbre.

Quantizing these tracks took about two hours, but this is remarkably efficient considering the literally thousands of edits, adjustments of beat placements and extension of regions to close gaps, creation of crossfades in order to smooth transitions, and about twenty manual adjustments at places where the automated process produced slightly anomalous results.

In working with the drum track I made several other typical alterations
to the files in order to create cleaner and more consistent performances. There were a few weak or bad-sounding bass drum or snare drum hits that had come from inconsistent striking of the drum, and I replaced those with better-sounding hits using a basic cut-and-paste function. There is also a tool in Pro Tools called Strip Silence that allows one to create silence below a user-definable amplitude threshold. In this way it is possible to quickly eliminate leakage sound from adjacent instruments, and for drums this can create a much cleaner overall sound. For example, the tom-tom tracks had substantial off-axis\(^8\) leakage from the snare drum, bass drum, and hi-hat. By stripping away all parts of the audio file other than the actual tom-tom hits I could remove the clouding effect of this leaked audio. Strip Silence provides a very efficient means of eliminating these off-axis sounds.

After “fixing” the drum part I proceeded to work on the timing of the bass guitar part. With the first two-thirds of the performance I was able to capture and separate regions into beat-oriented sections using Beat Detective. I then quantized to eighth-note triplets and smoothed the transitions using the automated fill and crossfade function. This extends audio regions where necessary to fill in the gaps created by moving the regions to their corrected beat placement. It then creates a short crossfade between adjacent audio regions to create smooth transitions. In a few places I used the copy-and-paste function to replace a poorly played part with the same part from a different section of the song. The last one-third of the bass part is continuous legato triplets with little dynamics and enough sustain to make it impossible for the computer to discern the break points needed to create the individual beats. This portion of the music consisted of a repeated two-bar pattern, so I found the best iteration, massaged it into shape, making slight adjustments to both rhythm and dynamics, and then pasted that “fixed” two-bar phrase throughout. I had to requantize each two-bar section to the appropriate downbeat because of the slight tempo fluctuations of the click. There was one musical variation at one transition point (the same pattern played up an octave) and I left that from the original performance. Then I used the automated smoothing function to close gaps and create crossfades on this final section. The result of this work was a more rhythmically stable bass track that sounded more accurately played to the drum track.

*The Wonderful World of Auto-Tune*

Having started to work with audio that contained pitch information—in this case the bass guitar track—it was now time to employ some pitch cor-
rection. The most frequently used tool for pitch correction in the digital
domain is a plug-in called Auto-Tune. Auto-Tune and its various succes-
sors employ pitch detection algorithms that are capable of reading very
small variations in pitch in real time (single-voice only, it cannot read mul-
tiple notes played at the same time). Once the software has determined the
continuous pitch information for a segment of audio, it creates a graphic
representation of that pitch on a grid where the vertical axis is pitch and
the horizontal axis is time. The user can then redraw the pitch representa-
tion on the graph to alter the pitch. When the original audio is played back
through the plug-in, it adjusts the pitch to the redrawn graphic informa-
tion. In this way variations in pitch deemed incorrect can be “corrected” in
exactly the way the user desires. This may be gentle correction to move
pitch variations closer to the actual note, or aggressive changes that lock
the pitch to the desired note. There is also an automatic mode that corrects
pitch in real time as the audio is fed through the plug-in. In this mode the
audio is gently moved toward whatever note the original audio is closest to,
though you can dictate which notes are “valid” by indicating scale function
or even by designating your own custom scale.9

I tried the Auto-Tune’s auto-mode on the bass guitar and found that it
nudged the pitch into a slightly more stable-sounding place. After all, the
bass guitar is a fretted instrument so most of the pitch information was
pretty accurate in the first place. However, variations in each string’s tun-
ing, and pitch shifting caused when the string is stretched slightly by the
pressure of the finger on the string against the fretboard were reduced by
the application of Auto-Tune.

I then moved onto the lead vocal, where Auto-Tune is frequently used
for pitch correction. Many rock singers have a less than exacting ability to
execute accurate intonation. The relative merits of variations from the
ideal pitch may be argued, but for this exercise, and for most of the vocals
heard in rock production today, Auto-Tune is used to “improve” the accu-
rracy of a singer’s pitch. In this instance I used the graphical mode, which
allows for more aggressive retuning of each vocal line than the automatic
mode. Although I sometimes used a straight line in the graphic window to
“flatten” the pitch to the exact note, the program has settings for how
quickly and completely it “corrects” the singer’s performance in line with
the graphical model that the user has created. For this project I have the
parameters set to the fast side of the “retune” continuum and the choosy
side of the “tracking” continuum, so corrections are made pretty quickly to
conform quite accurately to the graphic representation. Thus, even though
I may graphically indicate a flat line for the pitch on the note desired, there will still be some pitch variation in the final audio as the program does not track the note immediately or retune it completely. “Correcting” the vocal pitch performance took about two hours. Although the pitch accuracy of the original vocal was quite good, the vocal intonation sounds more stable after this process.

Although I do not feel the need to adjust the timing of any of the vocal lines, this is something that is also done in contemporary production. Vocal lines (or words) may be moved slightly to “improve” their rhythmical accuracy. They may even be moved by beats or bars to fit into the overall composition in a different way, creating a different vocal arrangement. Certain functions, such as vamps at the end of songs where there might be a long passage of vocal ad-libbing, are often the result of extensive editing. In the past we might painstakingly construct an “ad-lib” section by recording one “riff” at a time, with the singer performing and reperforming certain phrases as we worked through the section of the song—composition through recording. Now we are more likely to record several versions of the vocal vamp and then construct the ad-lib sequence through the editing process—composition through editing. In one recent project of mine an entire ending ad-lib section was created by repurposing pieces sung at completely different locations in the arrangement and using them for a newly constructed vocal ad-lib ending.

Finally I work on the guitar tracks. Because these are primarily chordal parts it is not possible to use Auto-Tune on them. It is also difficult for Beat Detective to determine beat information because the transients at the beginning of beats are not necessarily that pronounced. However, because of the repetitive nature of these parts, it is often possible to find iterations of particular phrases that are played particularly well and to use these in places where the same phrases are not as well executed. I do this with the guitar tracks, eliminating any rough spots and generally creating more accurate performances. I am done “fixing”: total time, about twelve hours.

I then mix the song—balancing levels between instruments and voice and optimizing the sound relative to frequencies, dynamics, and ambiances according to my personal aesthetic. I run one version using the original “unfixed” audio and one version using the “fixed” tracks (audio clips 1 and 2). It is not my intention to make a formalized survey of responses, but I did take an informal reading by playing the two mixes for some of my recording classes. I find that in general students are not able to identify what separates the two recordings. In fact they often misidentify the dif-
ferences between the two versions as changes in arrangement or instrumentation that did not occur. This, for one thing, suggests that people listen very differently when they are asked to listen.\textsuperscript{10} For the most part I do get responses that identify the “fixed” mix as being “clearer,” more “polished,” more “stable,” and more “professional,” but I also get some responses that identify the “unfixed” mix as “cleaner” or that otherwise apparently reverse the nature of the two versions. I do not pursue listening tests, as my interest is in the implications of the process rather than in gathering statistical data from such tests.

\textit{Meaning in Manipulation}

The process described above—from “fixing” to mixing—raises many questions regarding the cultural, social, and philosophical position of contemporary music creation. Before considering each of these areas in the following sections, I want to consider the root idea of audio manipulation, its epistemology. In doing so I will address some basic objections that arise in relation to the kind of manipulation of audio here described. Have we become so removed from the original music as to have lost the essence of its meaning? Do critiques of recordings in regards to performance technique (Barthes) or musical expression (Robert Philip) or rhythmic subtlety (David Epstein) miss the point by failing to identify or acknowledge the new forms of musical creativity that have emerged within the contemporary technological context?\textsuperscript{11}

There is first the musical act, and this is profoundly human. In some form human agency is required in order to produce what it is that we call music. All music is once removed from the act that created it. That is to say that music is already a secondary product; it is, in Anthony Gritten’s words, the “trace left by an act.”\textsuperscript{12} Musical instruments themselves are technological objects that are inserted between the musical act and the music. Equally, vocal technique is necessary for vocal expression to become singing. Music is not a natural phenomenon but is necessarily a product of human activity. Musical construction begins before there is music. It begins in the musical act, but the constructive nature of music is then an inherent part of its very call from act into existence. Thus construction, most often through a musical instrument, means that \textit{technological mediation is part of the entire history of musical creation}.

Technology is, of course, essential to the act of recording. The advent of recording created a universe of music that was necessarily mediated
through complex technology. While recording did not actually alter the fundamental dependence of music on technology, the kind of nostalgic values that Adorno and Benjamin express seem to draw on a distinction between “natural” and “technologically mediated” music. If there is to be such a distinction, it is more accurately understood as between “naturalized” and “mechanized” technology. That is to say that earlier, prerecording music technology was viewed as natural because the technological aspects had already been deeply integrated into the cultural process. Audio recording technology inserts itself into the musical process between the music creation and the music preservation. It is the technology that is needed to produce the audio preservation of music, just as flutes, pianos, and the like were the needed technology to produce the sound of music. However, it is not the case that live performance remains the arena of music unaffected by the technology of audio recording. In fact technological effects on so-called live performance have made it more of a relative to recordings than a stranger.13

Technology not directly connected to the musical act manipulates and mediates virtually all music in contemporary culture—both live and recorded. The musical object is already a product of technology, its “trace” left not only by the act but by the mediation of technology as well. The application project I have described represents a dramatic shift in the way technology mediates our relationship to music creation. It is indicative of new kinds of music creations. As production increasingly takes on the qualities of both performance and composition, all of these activities blur in the final recorded presentation. Contemporary production techniques greatly expand the manner in which people may be involved in musical construction, and in doing so redefine the very nature of what it means to be a musician. Théberge notes that “an understanding of the basic technologies, routines and practices of studio recording has gradually become an essential part of every musician’s store of knowledge and skill.”14 However, technological expertise does not obviate the need for some of the traditional relationships of people to music creation. When Cook says that the creative imagination in music production requires “some degree of specifically musical training,”15 he may not have been anticipating the kind of production I am employing here, yet musical training is required to effectively edit and manipulate music even in highly automated circumstances. Has the technology trumped musical knowledge? Certainly it has not completely done so, as basic principles of music remain relevant to application of the creative imagination in these circumstances as well. The
more developed the practitioner’s musical ear, the more sophisticated the relationship to the editing and “fixing” process.\textsuperscript{16} The manner in which such sophistication may be acquired, however, may be far from those recognized in the traditional music classroom.\textsuperscript{17}

As to musical content, the same kinds of considerations apply. Frederic Jameson indicates that music follows the Sapir-Whorf hypothesis, which posits new thought as the result of expanded language capabilities, and this suggests how “music of a given period is able to express new kinds of content.”\textsuperscript{18} We may extrapolate the same for technological applications to music, allowing them to express new kinds of content as well. The vocabulary of music is as much wrapped up in digital technology now as it was in previous times with new instruments or new cross-cultural influences. And these new kinds of musical content are not to be understood as cold, objective applications based on some kind of mathematical treatment. As Arnold Pacey argues, technology participates on all levels: “No aspect of human life, be it music, medicine, or technology, can be adequately discussed if we are always restricted to a scientific mode of discourse.”\textsuperscript{19} So even the most audacious forms of contemporary manipulation of audio data are not to be completely divorced from the technology of musical instruments or of live sound production. That is to say that Gritten’s “traces of the act” may be understood in technological terms—as a part of the postrecording process whereby music is the result—in the same way that he describes them in the terms of live musical production.

New technologies continue to participate in music’s ability to express new kinds of content. The fact that there is no music without technology (instruments) or technique (voice) undermines the conservative forces that have sought to minimize the qualities of musicianship and the creative forces of composition that are dependent on the technologies of recording. It is nostalgia disguised as the “natural” that fuels the conventional critiques of contemporary music and its practices. The new techniques and expressions of musical creativity are today becoming naturalized in contemporary culture. It is in this context that I begin a more detailed look into the application project described above.

\textit{Variations and “Fixing”}

The bulk of the work on the pop song as described above was done to “improve” on the extent of deviations from “perfect” rhythm and pitch in the original performances. In contrast it is common to make claims for incon-
sistencies of musical performance as being essential to the emotional content of the performance. These inconsistencies may be random (as reflected in the randomization functions of some musical software), or an intentional and consistent kind of note placement that is systematically applied by a musician as a part of an overall performance structure. This systematic approach is most often associated with popular music forms and is thought of in terms of rhythmic “groove.” Performance inconsistencies may also take the form of nuanced deviations from nominal (score-based) values; in this case the deviations are not consistently applied, rather they vary the expressive intention over different passages and sections of the music. This approach is widespread in Western art music. Although the two approaches may coexist, they are fundamentally different. They share a reliance on inconsistency for effect, but differ completely in how such deviations are applied. In either case, most observers acknowledge the value of inconsistencies as essential elements of musical performance. However, many also argue for the inability of a machine, or a human programmer, to reconstruct such inconsistencies in a way that matches the emotional impact of a singular human performance. Others are busy constructing automated feel factors, groove templates, and randomizing algorithms in order to emulate and control, and even it may be argued, improve upon the quality of deviation in a musical performance in order to provide a greater emotional impact. These programmable approaches may seek to emulate either the groove-based or the expressive-based approach to rhythmic variations. The objective of both forms of variation, and both modes of implementation (the human and the programmed), is a greater emotional impact.

In *Performing Rites*, Frith notes Charles Keil’s model of participatory discrepancies: “A PD is ‘a slight human inconsistency’ in the way that a musician executes rhythm, pitch and timbre,” and goes on to provide a good summary of various positions regarding PDs. As an example of a highly restrictive outlook on the application of PDs he references David Epstein’s claim that session musicians can’t play “naturally or ‘musically’” to a click track. Epstein suggests that the mechanized influence of the metronomic click track robs the potential for a truly “musical” performance. It is difficult to divorce Epstein’s negative assertion from the fact that professionally his work has been as a conductor. He seems to be reflecting a very specific viewpoint regarding rhythmic practice that may be appropriate to his own work but is dislocated from many other approaches to performance. Just as the concert musician must balance his or
her own expressive inclinations against the conductor’s lead, so might a
studio musician balance his or her expression against the click track. The
studio musician might engage the click track as another musical part just
as concert musicians collaborate with the conductor as their musical part-
ner. It is true that the conductor’s lead is personalized rather than mecha-
nized, and it may be that strict adherence to metronomic time would rob
most classical music performances of their musicality. But Epstein’s com-
ment is aimed at the studio musician who often performs to metronomic
time. This practice is so widespread that Epstein appears to be condem-
nning virtually all of contemporary popular music, and indeed one might
better interpret his comment as a reflection of this broader judgment,
rather than the more specific condemnation of the relationship of popular
music to metronomic time. Musicians in popular music are frequently in-
teracting with the equivalent of a click track (a strictly metronomic ele-
ment) as an essential element in locating rhythmic interest. With the use
of drum machines (and other sequenced parts) as part of finished record-
ings, this practice extends beyond the more hidden relationship of playing
to click track. If we accept popular music as “natural” or “musical” at all,
then we must accept the active relationship to metronomic time.

In contrast to the Epstein reference, Frith acknowledges Prögler’s de-
scriptions of how “digital instrument makers have sought to ‘humanize’
their programs, have become concerned with ‘imperfections,’ ‘inaccura-
cies,’ ‘perturbations,’ ‘offsets,’ ‘adjustments,’ ‘shifts’ and ‘feel.’” An early
and comprehensive description of the kind of PD associated with Western
art music performance is found in Leonard B. Meyer’s 1956 work Emotion
and Meaning in Music. Meyer argues that emotional responses to music
may be analyzed through “a process in which the relationship of deviation
and norms to affective aesthetic responses can be examined and dis-

cussed.” By describing and analyzing musical inconsistencies this work
suggests ways that one may intervene in the application of PDs. However,
having been written prior to our ability to manipulate the details of per-
formances after they have been recorded, Meyer’s work doesn’t address the
questions of rhythmic programming or “fixing.” One of the early writings
to address the possibility of creating emotional responses by rhythmic pro-
gramming was Michael Stewart’s 1987 article, “The Feel Factor.” Stewart
catalogs the effects of various departures from metronomic note placement
in popular music. He suggests how a trained musician may create a more
compelling performance when playing with intention against a click track,
using the application of PDs to create specific effects such as “nervous” or
“heavy.” He uses this analysis as a basis for suggesting ways that programmers may re-create or alter performances to generate or heighten such emotional responses. Other important work in this area has been done by Ernest Cholakis in developing the technology for creating groove templates. His work also suggests ways that programming may “improve” upon original performances by altering PDs in specific ways to produce predictable emotional responses.

I have had extensive conversations with studio drummers about their interaction with the click track. Many of them have become expert at adjusting their performances’ relationship to metronomic time to suit the music or artist that is being recorded. Some artists prefer a drum track that pushes the beat, where the drums tend to land slightly ahead of the metronomic pulse (more exciting or more nervous, depending on one’s point of view) and some artists prefer a drummer that lays the beat behind the click track (heavy or “grooving,” though again, subject to some subjective interpretation). Although much of the work analyzing PDs has been motivated by programming considerations, it has served to guide many musical performances as well. This is a typical example of technologies working on both sides of the spectrum—from the traditionally technological back to the more traditionally performative.

To extend this question further is to consider the many ways that location of rhythmic interest may vary. Changing aesthetics might find the focus of rhythmic interest on compositional complexity, on note placement, or on even less obvious elements such as harmonic rhythm. Relocation of the center of rhythmic attention is an inherent part of shifting musical interests in various cultures. The notion of “humanizing” rhythmic inconsistencies in mechanically generated performances may provide significant rhythmic interest in a localized cultural environment, as it seems to in contemporary popular music. At the same time, inconsistency may be “built into” the relations between levels in hierarchically “deep” music—such as in the relationship of performer to conductor in Western art music. Similarly the physical relationship of rhythm to the body may vary from culture to culture, and over time in any given culture, and this in turn may affect the location of rhythmic interest. These physical relationships may exist within the performers themselves, in collaboration with a conductor or in an exchange with dancers; and each may dictate a differing location of rhythmic interest.

But aesthetics are hardly the only consideration. For practitioners such as myself we must find a balance between the aesthetic and the practical.
The nature and degree of “fixing” are influenced by a variety of real-world considerations such as available time and budget, the wishes of the artist and others involved in the production, and one’s own aesthetic regarding conforming to theoretical ideals of pitch and time. The fallacy that computers are time-savers is really exposed when it comes to “fixing” audio. The vastly expanded ability to correct small inconsistencies in pitch and timing is matched by the vastly expanded time it takes to make such corrections. We are almost always confronted with a situation where performances could be more completely “fixed,” but where either budget or aesthetic judgment restricts further editing or processing. Although ideally the aesthetic and collaborative processes dictate the extent of “fixing,” it is sometimes the purely practical time and money considerations that end up setting the limits on these activities.

How Do We Decide What Is “Right”? 

The collaborative process has taken on entirely new dimensions as a result of the technological capabilities I have described here. The ability to “fix” affects the making of the initial recording as well as the process of editing and mixing. There is a standard Pro Tools joke that reflects the recording side of these new capabilities:

A vocalist is recording his part for a song and singing rather poorly. When he finishes his “take,” the producer gets on the talkback to the singer and says, “That sucked. Come on in.”

The producer is suggesting that the vocal is complete and that the artist can come back into the control room, even though the producer has noted that the performance was not at all good. The assumption is that the quality of the performance doesn’t matter because it’s simply going to get corrected later anyway. Of course this is an exaggeration in order to create humor, but what is implied is now an everyday part of the recording process in many circumstances. Performers have heard how their performances can be corrected for pitch or rhythm and they, along with the other collaborators in the process, often take this into account in judging a performance. A singer or a producer may well say something like, “That take was great. One word was a little out of tune but we can just fix that.” Thus “fixing” becomes a creative part of the recording process. There’s a telling comparison with the much older studio saying, “We’ll fix it in the mix,” which was intended as satirical—an excuse used when a performer was un-
able to create a satisfactory performance and the producer was trying to avoid endlessly unsuccessful attempt to get it right—because the fact was that it really couldn’t be “fixed in the mix.” It is precisely that kind of “fixing” which is now a reality.

The same inclinations that lead to the desire to “fix” performances affect the control exerted over the initial performance. In either case it is some form of aesthetics that drives those decisions, although technology may dictate the means of implementation. The aesthetics come down to how those making the decisions hear things. Do we like the way it sounds or do we wish for it to sound differently? However, more than the simple aesthetics or personal taste may affect these decisions—they may be colored by the perceived demands of the marketplace or by the fear of how one’s performance may be judged by one’s colleagues. Ultimately it is a personal choice that is often made by the producer. The following rambling comment by prominent contemporary rock producer John Goodmanson indicates that the difficulty one may have in balancing personal aesthetics with success, and with the perceived demands of the marketplace, may affect the musical process in the studio:

I wind up doing it instrument by instrument more now I think than I used to. I think that just has to do with how bands play together [since] I work with younger bands and now it’s more major label stuff, so being closer to perfect is sort of more important than it used to be. Stuff that’s more commercial, more like regular pop music, is definitely all about the energy, but then there’s also a component of it where it needs to be pretty well locked, especially if you’re shooting for FM radio. Really it comes down to what the players can pull off in the live take and what they can’t pull off, then it gets rebuilt instrument by instrument.  

Such choices are also the product of cultural influences that are a reflection of the time. The cultural ear shifts, representing a new intersection of music and technology. A simple example is the use of drum machines in popular music. When first introduced in the 1970s drum machines were widely ridiculed as horribly stiff and unnatural sounding. They were considered useful for the songwriter in sketching ideas or making demos but hardly appropriate for final production. Nonetheless they started to be used in certain genres of music—especially disco, which initially adopted highly repetitive and “mechanized”-sounding drum performances—and became accepted in certain limited musical circles. Over time their use has grown exponentially, and we rarely hear about their stultifying effect on
music or musical feel any more. People are still involved in programming
music using the idea of participatory discrepancies (see Cholakis above),
and the use of sampled drum loops that incorporate actual performances
has become very common, but there is a lot of popular music that uses the
“perfect” performance of a drum machine as its rhythmical basis. It is not
just that the cultural ear has developed a high tolerance for the “stiff” per-
formance of the metronomic drum machine: over time this tolerance has
developed into a required and desired element for certain genres of popu-
lar music, such that much of disco, rap, or hip-hop would not sound styl-
istically correct without the drum machine’s presence. This standard also
affects the desire to fix human performances.

Just as early recordings seemed to make musicians want to play in a
more technically correct manner, to adjust what sounded to them like er-
rors in performance when they heard themselves played back via the
recordings,26 so contemporary musicians may wish to hear rhythmic per-
formances that adhere more closely to the ideal of the drum machine or
other quantized performances. More will be said on this later, but for now
I observe that we have these new capabilities and we may choose to use
them or not, to please our ear for whatever reasons. The aesthetics of “fixing”
is simply the aesthetics of hearing, of how we wish for things to sound. But our
ear is influenced by the capabilities of these technologies as well as by aes-
thetics. This is the essence of the “debate” between the effects of technol-
ogy and the forces of cultural practice—technology and cultural practice
are both continually feeding off of each other. And there is a synergy be-
tween the practitioner’s ear and the public’s ear as well. That is to say that
production techniques (such as the use of drum machines) become ac-
ceptable and even expected by listeners, and this drives practitioners to
conform to the expectations of previously produced music—drum ma-
machines might be necessary for a recording to really qualify as “disco,” for
example. And beyond this, applications in one genre cross over into oth-
ers, so that technologies begin to appear in genres that might have previ-
ously rejected them as inappropriate—for example, drum machines in
punk rock or synthesizers in folk music.

We could project into the future effects of such cultural imperatives but
we don’t necessarily have to: we can reference a popular science fiction
book to do it for us. In Neal Stephenson’s Snow Crash an aspiring rapper,
Sushi K, confronts a crowd of teenage music fans: “He stares at the crowd,
five thousand potential market shares, young people with funkiness on
their minds. They’ve never heard any music before that wasn’t perfect. It’s
either studio-perfect digital sound from their CD players or performance-perfect fuzz-grunge from the best people in the business.” 27 Yet even “perfect” drum machine performances have very small fluctuations, as there is really no such thing as “perfect” clocking. 28 So, in spite of the futuristic reference, the continuum of deviation is inherent in all music—more or less “perfect” but never perfect. What has changed more dramatically than the level of adherence to the “perfect” performance is the people who now may contribute to that aspect of the performance. The performer used to be the primary arbiter of that continuum; the recordist is now a frequent contributor and often even the final arbiter in these aesthetic decisions.

The position of the performer’s hegemony over PDs has changed in both popular music and Western art music. There is considerable overlap between the role of the producer in popular music and the conductor in concert music: the dynamic between the producer/conductor and the performing musicians is very similar in the sense that it may be highly interactive, but there is often the tension between the fact that authority is placed with the producer/conductor yet the actual production of sounds rests with the musician. The primary difference for the producer and the conductor is their relationship to musical time: for the most part the conductor exerts musical control prior to the performance—in rehearsal—and in real time as part of the performance of the music, while the popular music producer often exerts control after the fact as well. Similarly classical producers have found their own means of independent influence after the fact through the process of editing: sometimes multitudes of painstakingly chosen edits are used to construct a final recording. While this doesn’t represent the same degree of intervention as when a popular music producer actually moves sounds around to alter the original performance, it is an evolution of the same impulse whereby people other than the original performer are making direct and significant alterations to the final recorded “performance.”

Mixing and Mixing Metaphors

The final stage of this application project was the creation of the stereo mix. The considered mixing of audio elements has been practiced since the advent of recording, beginning with simple techniques such as the positioning of musicians relative to the primitive horns first used to capture sound. This process reached a whole new level of complexity and sophistication with the advent of multitrack recording. With these tape recorders
(and now through the DAW) many disparate audio elements can be recorded at different times and even in different places, each synchronized to the same timeline. Along with having multiple audio elements comes the necessity of mixing these elements to create a final (generally stereo) audio program. For our purposes, without reviewing the entire history of mixing, it is valuable to note the major changes in mixing that have arrived with the use of the DAW.

Whereas pitch and time “fixing” have seen advancements that encompass wholly new capabilities, the art of mixing has been altered by more subtle improvements. The two primary advances in mixing technology of the past twenty years, automation and recall, have seen significant enhancements as a result of the DAW production techniques, but have not seen substantive changes in basic functionality. Mix automation is the ability to program (automate) changes in the relationship between elements dynamically, as movement over time. This means that volume relationships can be altered sequentially—for example, the vocal can be made louder in the chorus than in the verse—and this change in volume is stored in memory and reproduced automatically in reference to a consistent timeline. This kind of volume automation was first introduced on analog mixing consoles (using computer-assisted functionality) in the 1970s. Early systems were cumbersome and unreliable, but this process was refined and by the late 1980s very reliable systems were available for analog consoles.

Earlier techniques of mixing by altering the position of musicians relative to the horns or microphones, or using multiple microphones and altering their position to capture different instrumental balances, or even changing instrumentation for the sake of the clearer reproduction on primitive recording systems, had already established the precedent whereby recordists had significant influence over the final presentation of the musical performance. Multitrack recording, combined with volume automation, allowed the recordists even greater control over the blend of elements in the final audio program. With sounds retaining their separate identity up until the final mix, and with dynamic automation of volume, producers have the kind of control over the final sound of the music, the relationship of the individual elements, that influences aspects of the arrangement and ultimately even the composition. Placement and balance of melodic, harmonic, and rhythmic elements can alter the listener's experience profoundly—as anyone who has heard certain contemporary remixes can attest.

DAWs greatly expand the implementation and functionality of au-
tomation, thereby further increasing control over the presentation of the final audio program. On the analog systems only volume (and muting, which is simply another means of volume control) was accessible through automation. When working entirely within the digital realm virtually every parameter of signal processing is available for automation. Panning effects, parameters of ambience effects (such as reverb), and subtle alterations in timbre (through equalization) can all be dynamically altered over the course of an audio program, and these changes can be memorized and automatically reproduced for each playback. Additionally the computer allows for a level of fine control over volume automation that was not possible when linked to an analog console. Through manipulation of graphic representations of volume (and every other parameter) against the waveform of the sound, the mixer can select the exact portion of the audio that s/he wishes to automate and control the variations to within one-tenth of a decibel. All of this is quickly and easily done with immediate playback of the results available to allow for an aural judgment of the effect. The enhanced ease and speed of automation technology assist in heightening intuitive responses to music construction. This in turn expands the creative capabilities of the recordist responsible for building the final mix.

This kind of intervention in the sound of the final audio represents a significant expansion in the capability of altering the arrangement (or orchestration) of music after it has been recorded. It is interesting to note that when complex mixing became a standard part of making audio recordings through the proliferation of multitrack recorders in the 1960s, it was often referred to as remixing. This was a reference to the idea that the original “mix” of the elements was the sound when they were played together (as though they had been played for a stage performance) and the studio process was a revision of this “live” mix by the musicians, a remix. Remix differentiated the aesthetic of reproduction from the aesthetic of live performance. As the process became ubiquitous (and the aesthetic less often based on live performance), the re- was dropped and the process was simply called mixing. Shortly thereafter the idea of remixing was reborn as new mixes that were different versions of the original mixed elements—most often for different intended audiences such as dance clubs. These remixes now frequently contain different elements than the original recordings, involving additional people in the compositional side of music creation.

The other advancement in mix technology involves the process of mix
recall. Recall refers to the ability to re-create all of the settings involved in
the mix process at a later time. Every knob, fader, and function (or their
digital equivalent) must be remembered and reset in order to reproduce a
particular audio mix. In the world of analog mixing boards and outboard
processing gear this was accomplished through a combination of snapshot
memory that allowed for a graphic recall of the mixing board’s settings that
could be then reset manually by the engineer, and extensive note taking
and subsequent resetting of the parameters of the outboard gear—the ad-
ditional processing gear used as an adjunct to the processing power of the
mixing console. In the most disciplined studios this process took at least a
half an hour and was about 90 percent reliable. When working entirely
within the computer this process now takes less than a minute (the time it
takes to open the file) and is 100 percent reliable.

Immediate, reliable recall allows for both a more extended and a more
spontaneous relationship to the mixing process. Mixes can be listened to
and critiqued over time and then easily and quickly changed without hav-
ing to reset a mixing console. This increased functionality of automation
and the ease of recall have created new opportunities that encourage the
extension of the mixing process. It has made it easier for more people to be
involved as mixes can be sent (via the Internet most frequently) to almost
any location, feedback given via email, and then changes quickly made and
resent for approval. It is now not uncommon for the mix process to be as
lengthy as (or more lengthy than) the recording process.

Mixing is the one of the recordist’s avenues to participating in the ar-
ranging and composing in a direct fashion as described above. It also pro-
vides opportunities for other creative collaborations that merge the musi-
cal and the technical. Typically in a mix session an artist may ask the
recordist to make the audio sound more aggressive, or dreamier, or more
magical or even more purple! The recordist has to interpret these
metaphorical soundscapes and translate them into specific technical appli-
cations. Mixing is the most creative part of the technical aspects of making
recordings because it requires interpretation and technical application
within the context of what is often a very complex interrelationship be-
tween a wide variety of pitches, rhythms, dynamics, timbres, and ambi-
ences. The tools available for manipulation of these elements have become
enormously expanded within the realm of digital audio. “Fix it in the mix”
is now so much of a reality as to have left the old satiric implication of this
phrase meaningless.
**From the Mixer to the Listener**

Mixing is once removed from the recording process, which is once removed from the musical performance. It is a complex construction of sound and a convoluted process. Digital re-creations of acoustical spaces allow for simulations of ambiances not actually found in nature. Very accurate and discrete delayed versions of the sound source, reverb with tails that end abruptly, and perfectly regular modulations represent unnatural acoustical phenomena that are frequently employed in mixes. Sonically the combinations of tonalities and ambience often create acoustical impossibilities—sonic landscapes that could not actually exist in the physical world. The isolation of the instruments allows the mixer to place each sound in its own acoustical environment. No ensemble could possibly occupy the number and variety of spaces created in many mixes. The mixing landscape is often a product of creative spatial imagination.

The process is also unnatural in its dislocation from the music’s temporal timeline. Even prior to graphically based mix automation, the mixing engineer would frequently replay short sections of audio over and over—adjusting the mix until the engineer was satisfied with the small section. Now, adjustments are made off-line—that is, graphically against the representation of the waveform—and then listened to for effect. In either case, an outside listener at a mix session would hear bits of music played repeatedly and sections of music played out of sequence for long periods of time before any kind of recognizable playback of the entire audio program might be heard. The music is “dressed up” for final presentation, one article of clothing and one accessory at a time, each design carefully tailored for fit and finish.

Interestingly this prompts a comparison between composing and mixing—the two far ends of the process that brings most music from initial creation to the form made available to the listener—as there is a dislocation for a composer between the process of creation and the final music as well. The composition process that is a primary means of bringing music into existence bears a close relationship to the process by which a mixer makes the final preparations for music to be heard. Because the composer is modeling when composing at the piano (or any instrument), “he hears the music as he imagines it in his ‘inner ear.’” If the composer tries to perform with too much emphasis on this part of the process—the inner ear—she may lose the audience who isn’t hearing those connections. Similarly if a mixer focuses too much on the momentary relationships of
sound, he may lose the flow of musical ideas and lose the listener. In either event, the actual compositional process (at the piano, say) and the mixing process are typically not very musical sounding when being done.

One might expect that both composing and mixing reflect a progression from the simple to the complex. Yet Beethoven often moved back and forth between simpler and more complex versions of his symphonies, finding what for him were the best choices by building from the bottom while at the same time he was editing from the top.\textsuperscript{32} Similarly with mixing, the mixer frequently moves back and forth between focusing on an individual sound, and the building of the relationship between all the complex tonalities of an ensemble. The timbre, stereo positioning, and ambience of the flute (for example) may be set individually and then reset when placed in the context of the ensemble and then reset again when listened to in isolation at a later point. The composer imagines their creations with a kind of detail that most listeners are oblivious to, unless they study the music analytically. Like the composer, the mixer has a very intimate knowledge of the music, and the choices made in creating the mix are far beyond the understanding of the majority of listeners.

So in the end the “time-disordering operations found in a recording studio . . . become a means of exploring the temporal conditions of research and ethnographic representation,”\textsuperscript{33} yet the same disordering operations are found in the compositional process. Thus the musicology of the creation of the contemporary recording has some direct parallels to the musicology of composition. Ultimately, as the processes of recording and composing become increasingly intertwined, we might also draw comparisons between DAWs and musical instruments, though the relationship to real-time performance is fundamentally different. But despite the dislocations of process in recording, it is the final ordering of music that may be its strongest attraction, its most basic correlation to the human condition: “What most authorities seem to agree is that music reflects in some way the order—the organization—that is necessary for the human nervous system to function.”\textsuperscript{34} And the use of a DAW is now frequently a major contributor to that ordering process.