Audio Engineering Society

Convention Paper

Presented at the 137th Convention
2014 October 9–12 Los Angeles, USA



# Listener preferences for analog and digital summing based on music genre

Eric Tarr[1], Jane Howard[1], and Benjamin Stager[1]

[1] *Curb College of Entertainment and Music Business, Belmont University, Nashville, TN, 37212, USA*

Correspondence should be addressed to Eric Tarr (eric.tarr@belmont.edu)

**ABSTRACT**

The summation of multiple audio signals can be accomplished using digital or analog technologies. Digital summing and analog summing are not identical processes and, therefore, produce different results. In this study, digital summing and analog summing were performed separately on the audio signals of three different recordings of music. These recordings represented three genres of music: classical, pop/country, and heavy rock. Twenty-one listeners participated in a preference test comparing digital summing to analog summing. Results indicated that listeners preferred one type of summing to the other; this preference was dependent on the music genre.

## 1. INTRODUCTION

The recording and processing of audio signals can be accomplished by both digital and analog technologies. One process in multi-track audio recording is the combination of the individual signals into a format for playback (e.g. mono, stereo, 5.1, or 7.1 signals). Individual signals in a multi-track audio recording can be combined by using digital summing (DS) or analog summing (AS). These different types of summing technologies produce different signals.

Digital Audio Workstations (DAW) provide DS as one aspect of the recording and processing system. However, even if individual signals were recorded with a DAW, it is not necessary to combine them digitally. Rather, they can be combined using AS.

The following describes a common system of AS, substituted in place of DS, with the assumption that the final result should be a digital signal. First, the individual digital signals from the DAW are separately and simultaneously converted to analog signals. Then, the individual analog signals are combined using an AS device into a format for playback. Finally, the combined analog signal is converted into a digital signal and recorded into the DAW.

Depending on the analog device used for AS, there will be differences between the signal created using AS and a signal that could be created using DS by the DAW. These differences may be a result of signal distortions in AS due to the conversion process or from nonlinear components of the signal path. However, it was not the purpose of this study to determine the specific cause of the differences.

The purpose of this study was to examine whether listeners perceptually prefer music that has been processed using DS or AS technologies. An additional purpose was to investigate whether the listeners' perceptual preference was consistent across different music genres. It is common for recording and mixing techniques to vary between different music genres. Therefore, it is possible that DS or AS may be a preferrable processing technique for certain music genres.

Previous studies have investigated listener discrimination for various DS algorithms [1, 2]. In these studies, the DS of five different DAWs was tested to see if trained listeners could discriminate between the various DS algorithms. Results showed that listeners were not able to discriminate between DS algorithms in a statistically significant manner. This study aims to extend these previous studies by measuring listener preferences, and comparing the results of signals created using DS and AS.

## 2. METHODS

A perceptual experiment was completed to analyze listener preferences for DS and AS based on music genre.

### 2.1. Subjects

Twenty-one subjects participated as listeners in the perceptual experiment. Subjects were undergraduate students at Belmont University in Nashville, TN; ranging in age from 20 to 23 years old. Subjects had received a minimum of three years of formal critical listening training.

### 2.2. Stimuli & Equipment

Three recordings comprised of multi-track audio were used to create the stimuli for the experiment. Each recording was selected to represent different genres of music: classical, pop/country, and heavy rock. All individual audio tracks were recorded and mixed using Avid Pro Tools 10 at Oceanway Studios in Nashville, TN. Original audio files were recorded using a 48 kHz sampling rate and 24 bit-depth. To prepare the original audio files for summing, sub-mix groups were created within the Pro Tools session for different instrument sections: drums/percussion, bass, guitar, vocals, etc.

To create the DS stimuli, a stereo audio file was created by mixing together the sub-mix groups within the Pro Tools session. To create the AS stimuli, each sub-mix group was routed separately through a Focusrite Rednet 4 digital-to-analog converter. After the sub-mix groups were converted to analog signals, they were routed to an ADM 780M console. As the separate sub-mix signals were routed through the console, all channel faders were set to unity gain. A stereo mix of the audio signals was created by summing the sub-mix groups within the console. This stereo signal was recorded back into the Pro Tools session using the Focusrite Rednet 4 analog-to-digital converter.

To calibrate the ADM console prior to summing, a test signal (1 kHz sine wave tone) generated from within Pro Tools was routed to 14 console channels at a nominal level of 1.23v RMS. The input and output levels were then matched on each channel separately.

After recording the summed signals, the amplitude of the AS signal was matched to the amplitude of the DS signal based on the RMS energy. Five-second segments of the stereo files were selected for use in the perceptual experiment. The length of the stimuli was selected to provide enough time for the subjects to perceive the recording, while balancing the required memory load. The same time window was selected in the DS stimuli and AS stimuli.

### 2.3. Procedure

During the perceptual experiment, 10 trials of each music genre were presented. In each trial, the five-second segment of DS audio and AS audio were both presented, separated by one second of silence. The order of summing type (AS-DS or DS-AS) was randomized across trials. Listeners responded after each trial by indicating whether they preferred the first or second stimulus without knowledge of the summing type used in the stimuli. A total of 30 trials were used in the experiment, blocked by musical genre. Stimuli were presented over headphones in an acoustically controlled listening environment.

## 3. RESULTS

Chi-squared tests were performed on the listeners response data to analyze preferences. Individual listener preferences were analyzed, along with group preferences. A small percentage of individual listeners showed a statistically significant preference ($p < 0.05$) for one type of summing throughout the entire experiment. Across all music genres, 19% of listeners preferred DS and 5% of listeners preferred AS, while 76% of listeners showed no preference.

When separated by music genre, a higher percentage of listeners showed a statistically significant preference ($p < 0.05$) for one summing type. These results are shown in Fig. 1 for DS, AS, and no preference (NP). For classical music, 57% of listeners preferred DS and 5% of listeners preferred AS, while 38% of listeners showed no preference. For pop/country music, 14% of listeners preferred DS and 0% of listeners preferred AS, while 86% of listeners showed no preference. For heavy rock music, 19% of listeners preferred DS and 57% of listeners preferred AS, while 24% of listeners showed no preference.

Across all musical genres, compiled group responses indicated no significant preference for summing type. However, a significant preference ($p = 0.001$) for DS was found for classical music and a significant preference ($p < 0.001$) for AS was found for heavy rock music. No preference was found for pop/country music.

Paired-sample t-tests were performed between the results of the three different musical genres. Statistically significant differences were found between classical and pop/country ($p < 0.001$), classical and heavy rock ($p < 0.001$), as well as pop/country and heavy rock ($p = 0.03$).

## 4. DISCUSSION

In summary, results from this experiment suggest listeners prefer DS for one music genre (classical) and AS for a different music genre (heavy rock). No preference was found between the two summing methods for the third music genre (pop/country). Therefore, listener preferences for summing technology depend on the music genre.

To further investigate this result, this study could be extended by including multiple songs of the same music genre. Additionally, more genres of music could be compared. Data on the perceptual quality for a music genre may also be important to analyze when concluding why one signal is preferred to another.

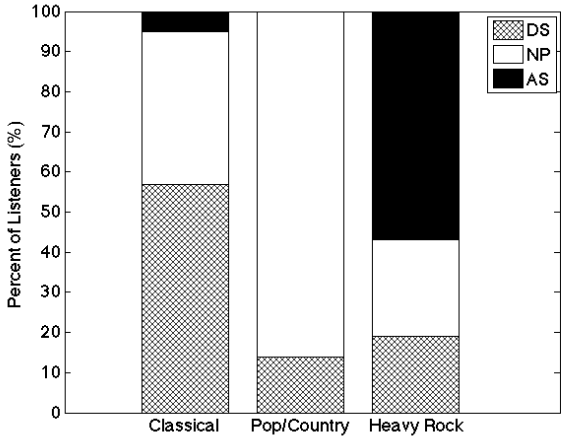


**Fig. 1:** Percent of listeners preferring summing type based on music genre.

The results is this study were obtained using one AS system - Focusrite Rednet converter and ADM 780M console. Many other AS systems exist, and could be included to generalize the conclusions from this study. Although previous studies have found no subjective differences for various DS systems [1, 2], it may be possible that subjective differences can be found for different AS systems.

The entire AS system was treated as a single processing system in this study. It may be beneficial to understand what aspect of the entire AS system changes the signal to how it is perceived. It is possible that a perceptual preference can be found for signals that are merely converted between digital and analog representations, or processed by individual analog circuits without using AS. However, it may be the actual summing aspect of the system that is important.

Listeners in this study were asked to choose which of the two stimuli in each trial they prefered. It can be concluded that if a listener achieved statistical significance for a preferred type of summing, then they can detect a difference between the signals. However, it cannot be concluded that listeners with no preference could detect a difference between the signals. Given the number of listeners that showed no

preference in certain music genres, it may be useful to study listener discrimation along with perceptual preference.

It was previously found that listeners could not discriminate between various DS algorithms [1, 2]. One explanation for this is that "any high-quality digital mixing system [...] is designed to provide a result that is sonically transparent" [3]. Conversely, AS systems are not transparent and can be detected perceptually. Therefore, AS can be used to consciously add subjectively pleasing elements to a mix.

In conclusion, the choice of using DS or AS technology is subjective. Preceptual preferences are not consistent across all audio signals. Nonetheless, preferences exist for a majority of listeners in some cases. The optimal summing method can be chosen by determining the highest percentage of listeners with a preference. When this percentage of listeners is combined with the percentage of listeners that showed no preference, only a small percentage of listeners will typically remain with an opposing preference.

## 5. REFERENCES

[1] B. Leonard and P. Buttner-Shnirer, "Subjective Differences in Digital Audio Workstation Math," AES Convention e-Brief, AES 132, Budapest, Hungary, April 2012.

[2] B. Leonard, S. Levine, and P. Buttner-Shnirer, "Objective and Subjective Evaluations of Digital Audio Workstation Summing," presented at the AES 133 convention, San Francisco, USA, October 2012.

[3] G. Kashiwa, "The Pro Tools 48-bit Mixer," Digidesign Technical White Paper, Daly City, USA, October 2005.