DEGREE PROJECT

Audibility & Preference of DA Overload Associated with True Peak

Investigation of claims made against overload prevention

Mattias Strand

Audio Technology, bachelor's level
2023

Luleå university of technology
Department of Social Sciences, Technology and Arts

LULEÅ UNIVERSITY OF TECHNOLOGY

[This page intentionally left blank]

Abstract

The conversion of audio from the digital to analog domain has the potential to result in distortion due to converter overload. This occurs because some peaks in the signal cannot be defined digitally and only become problematic during the conversion into the analog domain, exceeding the level that can be represented by the converter, causing it to overload. Although True Peak limiting and metering can prevent and monitor this issue, some professional mastering engineers choose not to do so. The study tested claims made against overload prevention, including the adequacy of headroom in modern D/A converters and the inaudibility of the distortion caused by overload. Preference was also added to the audibility claim. Measurements show that there is not enough headroom in modern D/A converters to avoid overload, but the distortion created by overload is generally inaudible in an uncompressed WAVE format hard rock song. Additionally, there is no clear preference. The measurements found that overload only occurs when the device's volume is raised to its maximum output.

[This page intentionally left blank]

Table of Contents

1 INTRODUCTION

The loudness war has had a major impact on mastering of popular music for distribution. The objective of some modern music mastering productions is to produce the loudest music possible, using every bit available in digital full range. Loudness is believed to capture the attention of listeners and increase sales or streams (Vickers, 2010). This goal is often achieved through heavy compression, clipping, or limiting of the audio, which reduces its dynamic range and increases its potential for increased loudness. However, these methods can also distort the audio. To avoid such distortion, there are multiple level limits for signal loudness set by practice, recommendations, standards, or technical restrictions. These limits on headroom exist at multiple stages of the audio signal chain, such as in converters, external processors, digital processors, and analog consoles. Achieving desired levels of loudness therefore requires decisions and adjustments at various stages of the production and workflow. Engineers might need to adjust levels at various points to achieve one's goals of loudness.

Loud signals, close to the limit of digital headroom, may not peak in the digital domain, giving no indication of issue through conventional digital metering. But through the D/A conversion process, these loud signals have the possibility of overloading the converter, thereby creating uncontrolled distortion. A straight-forward demonstration of this concept is through a high frequency sine wave (Nielsen & Lund, 2003). The peak value of the digital representation of the sine wave may deviate from its analog peak value by up to 3 dB. When the digital signal is increased to the limit for digital headroom, the reconstructed sinewave will surpass the analog level that represents 0 dBFS. This is one of the stages where mastering engineers come up against the limits of headroom.

Digital sample peak meters can only measure the value of individual samples. To address this issue, the True Peak measuring algorithm was introduced and first published in the ITU-R Recommendation BS.1770 in 2006. This algorithm provides engineers with a dedicated meter that digitally estimates the analog peak level in dB True Peak (dBTP). In his research, Dash (2014) highlights the limitations of the True Peak measurement technique in regards to its accuracy. He argues that the oversampling rates employed by True Peak meters can compromise the accuracy of their readings, resulting in errors of up to 1 dB. This is due to the trade-off between the need for higher accuracy and the increased processing load that results from increased oversampling rates. Despite these limitations, True Peak measurement remains a widely used technique in the audio industry, particularly in the realm of modern music mastering.

Prevention of converter overload can be achieved through various techniques, with the most common and practical approach being the utilization of oversampling limiters, more commonly known as True Peak limiters. Lowering the overall volume is another technique that may seem straightforward, but this solution is often infeasible as overload often arises in conjunction with a desire for loudness. Most streaming services provide recommendations for True Peak levels in finalized productions, yet studies by Lund (2006) and Deruty & Tardieu (2014) reveal that popular music often exceed these recommended levels. This suggests that engineers either do not agree with the recommendations, are unfamiliar with their existence (which is unlikely among professional mastering engineers), or lack the capability to effectively meet these requirements.

The author has gathered anecdotal evidence from Swedish professional mastering engineers that some engineers choose to not prevent overload. Reasons for this include the belief that modern D/A converters have sufficient headroom to avoid overload, something that is mentioned as a future possibility by Lund (2006), and that distortion caused by overload may not be audible to the average consumer who uses low-quality D/A converters. Additionally, some believe that prevention methods interfere with the subjective appraisals of the material, with True Peak limiting being the main reason for it (MixbusTv, 2020). This claim regarding True Peak limiting can be explained by its effect on high frequency information. It is noted that higher frequencies tend to have higher inter-sample peaks, which are the main information that would be limited through True Peak limitation. However, there is no research to support or refute these claims.

The aim of this study is to examine the claims made against overload prevention. The headroom capability of three devices that vary in terms of purpose and quality will be tested through objective measurements, inspired by Nielsen & Lund (2003). Audibility will be tested through a controlled listening test on the device that creates the most amount of distortion through overload, evaluating the participants ability to differentiate between a version that is subjected to overload, and one that is not. The participants will be asked to indicate their preference, if any, between the two versions. Important to note is that this study will not examine claims regarding methods for preventing overload.

2 Theory

The manifestation of overload in the digital-to-analog conversion process is influenced by a range of factors, some of which may have isolated effects, while others may exacerbate each other's impact (Nielsen & Lund, 2003). The key contributors to overload include:

- Inter-Sample Peaks
- Bandwidth
- Clipping
- Codecs
- Oversampling

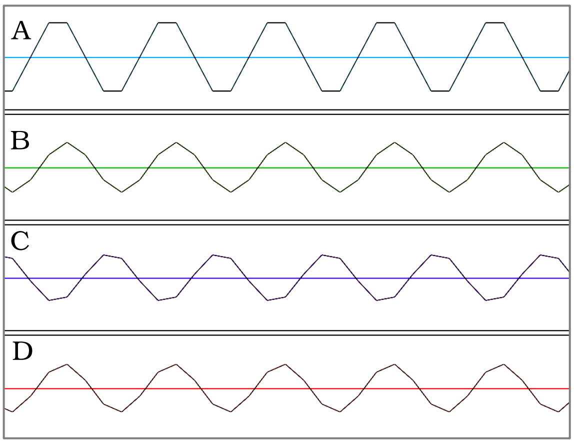

Figure 1: Pro Tools screenshot of a 7350 Hz sine wave with differing sample placement

2.1 Inter-Sample Peaks

During analog-to-digital conversion, the analog signal passes through a sample and hold circuit to define it digitally. The sample rate determines the frequency at which samples are taken. An internal or external clock is used to synchronize the timing of the samples so that they are evenly spaced and occur at the desired frequency. In devices with multiple converters, synchronization is achieved by all converters being connected to the same clock to guarantee simultaneous sample placement.

The relative positioning of the samples with respect to the signal being converted can vary significantly. For instance, recording sequential sine waves at a given frequency can result in different sample placements. This is illustrated in Figure 1, where Signal A does not have a sample to define the peak, while Signal B does. Signal C and D illustrates more possibilities for sample placement. Signal A can be digitally increased in volume to a greater extent than Signal B due to its lack of a defined peak.

When converting a signal from digital to analog, the pulse amplitude modulated (PAM) signal is converted into a continuous waveform using a smoothing filter. This reconstructs the information between digital samples and produces a more accurate representation of the original analog signal

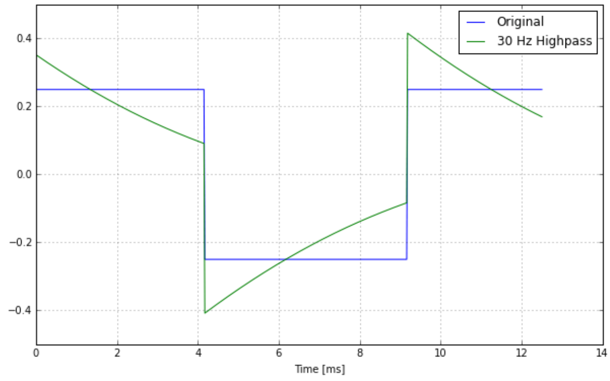

Figure 2: 100 Hz square wave with a high-pass filter with a cutoff frequency of 30 Hz and a 12dB/oct slope (Luther, 2016).

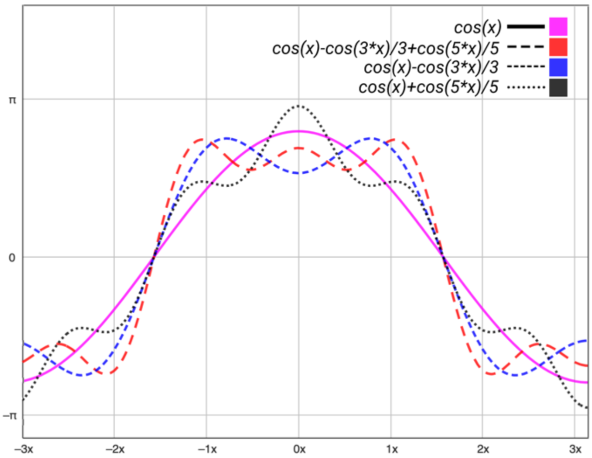

Figure 3: Square wave construction. 3rd harmonic omitted. (All multiplied by 2.5 for scaling)

(Pohlman, 2010). However, In the example of Figure 1, when Signal A's samples are at 0 dBFS, the reconstructed peak will exceed the rated 0 dBFS DC level of the converter, causing it to overload.

2.2 Bandwidth

Filtering is a common signal processing technique that involves attenuation or amplification of specific frequency components in a signal. However, the removal of spectral content from a signal can lead to unintended consequences. Specifically, it can result in an increased peak level, even if the filtered signal has less overall energy (Nielsen & Lund, 2003).

To illustrate this, Figure 2 depicts the effect of high pass filtering a low-frequency square wave. The resulting filtered signal has a potential peak level increase of 6 dB. While high pass filters are commonly utilized in analog interfaces to prevent DC offset, they can also be employed in the digital domain. It should be noted that some of the observed peak level increase may be attributed to a shift in phase, although it is clear that the application of a high pass filter to this type of signal yields a rise in peak level.

High frequency limitation presents potential consequences that are not immediately intuitive. In particular, it can result in an increase of peak level. To illustrate this, a square wave is utilized as an example. A square wave consists of a fundamental frequency and odd harmonics. By applying the Fourier series, the square wave can be deconstructed into an infinite number of cosine waves, see the equation below (notice the alternating signs).

$$\cos(x) - \frac{\cos(3x)}{3} + \frac{\cos(5x)}{5} \ldots.$$

By using a bandpass filter and omitting parts of the signal, the resulting waveform may have a higher peak level than the original square wave. This is illustrated in Figure 3, which depicts the construction of a square wave using cosine waves. Additionally, the resulting waveform when the 3rd harmonic is omitted (by removing it from the equation) is shown.

2.3 Clipping

The available bandwidth in the digital domain is subject to limitations imposed by the Shannon-Nyquist sampling theorem. The theorem states that the sampling frequency must be at least twice of the highest frequency that needs to be represented to avoid distortion. In this context, the generation of a square wave is limited by the number of available high frequency harmonics that can be utilized to define its slopes and tops. This can be seen as a smoothing effect on the waveform, reducing the sharpness of the transitions between the high and low amplitude states. Consequently, the resulting wave oscillates, overshooting and undershooting when compared to a perfect square wave, as illustrated in Figure 3. A notable point to consider is that even if an infinite amount of harmonics are added, a square wave would still experience over- and undershooting, albeit to a lesser extent, this is known as *Gibbs Phenomenon*. Digital clipping, created through either intentional or unintentional means, can be regarded as a form of square wave, albeit with some differences.

Another phenomenon that arises when sampling square waves is known as aliasing. When the sampling rate is too low, parts of the square wave signal can be incorrectly interpreted as low frequency sine waves. This is due to the requirement for an infinite bandwidth to accurately sample all of the harmonics. As a result, high frequency harmonics are aliased to lower frequencies, leading to the creation of unpleasant distortion artifacts.

2.4 Codecs

Codecs, while not a source for the creation of high peak levels, has the capacity to exacerbate the sources discussed above. Lossless codecs do not alter the peak levels of a signal, as they decode the signal back to its original state without discarding information. Lossy codecs that employ perceptual encoding techniques discard information that go unnoticed by the ear, resulting in smaller file sizes but also in an increased peak level upon decoding. This increase can be attributed to internal filtering processes and spectral shaped noise used by some lossy codecs to replace high frequency information, which are more easily subjected to peaking inter sample (Dash, 2014). In addition, bandwidth reduction also contributes to the increased peak level.

2.5 Oversampling

Oversampling is used to increase the sampling rate of a signal beyond the Nyquist rate (double the highest frequency that needs to be represented). The purpose of oversampling is to reduce the performance requirements of the anti-aliasing filter, which is used to prevent aliasing artifacts from appearing in the output signal. However, oversampling can have some drawbacks. As Dash (2014) notes, oversampling can lead to peaks increasing by up to 3 dB. One possible explanation for this is that the oversampling process can cause the signal to truncate (clip), introducing aliasing artifacts. This can increase the amplitude of certain signal components. Additionally, poor anti-aliasing filter design can also lead to increasing peaks in the oversampled signal.

Modern converters, which employ Delta-Sigma modulation, utilize oversampling to a significant extent. By oversampling the signal by a factor of 64 or more, these converters can attain a high dynamic range and a low noise floor. However, to achieve a high sample rate, the word length must be converted into a 1-bit signal. Luther (2016) notes that this conversion can result in severe overload in the converter when it is exposed to signals with inter-sample peaks. The exact cause of this phenomenon is not explained, but it could be related to oversampling or some other factor beyond the scope of this discussion.

3 Measurements

3.1 Method

Objective measurements were made upon three devices. The purpose of the measurement is to gather data regarding how much distortion is created when converters are subjected to overload. The data obtained from the measurement will be used to assess the validity of the headroom claim made against overload prevention. Additionally, the converter that exhibits the highest level of distortion will be selected for use in a subsequent listening test. The test was conducted using Total Harmonic Distortion + Noise (THD+N) measurements.

3.1.1 Devices

The following three devices measured:

- Avid HD I/O
- Macbook M1 Air (2019)
- Huawei P20 Pro

Table 1: Test signal frequencies, True Peak and digital sample peak values

Frequency	True Peak	Sample Peak
997 Hz	-0.1 dBTP	-0.1 dBFS
5512,5 Hz	+0.7 dBTP	-0.1 dBFS
7350 Hz	+1.4 dBTP	-0.1 dBFS
11025 Hz	+2.8 dBTP	-0.1 dBFS

These three devices represent commonly used equipment for music playback by both consumers and mastering engineers. The Avid HD I/O is a professional-grade audio interface used in recording studios and music production environments. The Macbook M1 Air (2019) is a popular laptop computer among music producers, DJs, and musicians. The Huawei P20 Pro is a smartphone, phones are currently the most common device used for music playback on-the-go.

3.1.2 Test Signals

In order to measure the THD+N of the devices, four sine waves of varying frequencies were generated with the goal to specifically create converter overload during playback. Frequency selection was based on previous research conducted by Nielsen and Lund (2003). Three of the frequencies are integer fractions of the sampling rate and 997 Hz has no relation to the sampling rate. The signals were generated in the analog domain using a Neutrik TT402 and were recorded into Pro Tools using a sample rate of 44.1 kHz. The files were normalized to a digital sample peak output level of -0.1 dBFS to avoid truncation (digital clipping).

Through this analog method, the optimal sample placement to create inter-sample peaks was achieved. Due to slight frequency oscillations inherent in analog generators, their frequencies will very slightly oscillate up and down in frequency. Because of this, parts of the recording lacking a sample to define the peak could be extracted through visual analysis of the Pro Tools timeline (as illustrated in Figure 1) and duplicated to create a longer test signal. The True Peak value of each signal was subsequently measured using Voxengo SPAN[1].

As sine waves have a simpler waveform, the measurement error of True Peak is reduced in comparison to complex waveforms. Furthermore, the obtained values were similar to the theoretical peak values of those stated by Nielsen and Lund (2003). The slight level variations can be attributed to differing methods to generate the signals. The test signals, their dBTP and digital sample peak values are shown in Table 1.

3.1.3 Setup

For the THD+N measurements, a Neutrik TT402 device was employed as the measuring instrument, using the 80 kHz low-pass (LP) filter. To conduct the tests, each device was separately connected to the Neutrik, and the test signals played back. The left and right channel of the stereo output was measured separately. To ensure consistency, multiple outputs was tested on Avid's HD I/O, and the results were compared to identify any significant differences. During initial testing, the output volume of the devices was set to the maximum level. This was followed by a subsequent test in which the output levels of the devices were lowered to examine volume controls impact on THD+N. The reduction level varied between devices, as each one utilized different systems to achieve reduction. For instance, the master fader was lowered for the HD I/O, whereas the device volume control was manipulated for the phone and laptop.

3.2 Results & Analysis

THD+N values for maximum output and decreased output levels are shown in Table 2. The highest THD+N value from the left and right channel was chosen. These results refute the claim regarding headroom as the most drastic increase in distortion occurs directly when the level exceeds the rated 0 dBFS DC level. The distortion between converters are quite close in THD+N level up until the 2.8 dBFS where the Macbook M1, the youngest device in the sample, performs better under overload than the Avid HD I/O,

[1] Voxengo SPAN https://www.voxengo.com/product/span/

Table 2: THD+N in dB (80 kHz LP-Filter). Output level of devices shown in brackets.
* Devices cannot be compared with each other in the decreased output test.

Frequency	True Peak	Avid HD I/O (0 dBFS)	Macbook M1 Air (100%)	Huawei P20 Pro (100%)	Avid HD I/O (-6 dBFS)*	Macbook M1 Air (70%)*	Huawei P20 Pro (70%)*
997 Hz	-0.1 dBTP	-81.5	-73.8	-67.2	-75.7	-64.0	-55.3
5512.5 Hz	+0.7 dBTP	-31.7	-32.3	-32.0	-75.6	-62.1	-56.1
7350 Hz	+1.4 dBTP	-25.9	-27.1	-27.1	-75.4	-69.5	-56.4
11025 Hz	+2.8 dBTP	-17.4	-26.2	-22.9	-71.6	-57.6	-59.2

which produced the highest level of distortion. However, the HD I/O produces the least amount of distortion when not subjected to overload.

As output level is decreased through the volume control (*Decreased output test), we see that the amount of distortion does not increase in the same magnitude as the results from the maximum output level test, indicating that there is no amplification circuit after the D/A converter. Rather, volume is manipulated by adjusting the digital level that enters the D/A converter. Consequently, overload, associated with True Peak, can only be created if the output level is set at or near the maximum. It is important to note that the results of the decreased output test (indicated with *) for the device cannot be directly compared with one another. This is because the measurements did not capture the magnitude of volume attenuation, making it impossible to establish a reliable baseline for comparison.

The THD+N value for the Huawei P20 Pro during the decreased output test (70%*) is decreasing instead of increasing. A likely explanation for this is the signal increasing further above the noise floor. Macbook M1 Air experiences an anomaly during the +1.4 dBTP test, suddenly decreasing in distortion before subsequently increasing in the more suspected pattern. Why this behavior occurs is unknown, although it could be a "sweet spot" for the converter.

Nielsen and Lund's (2003) study revealed significant variation in overload handling among the tested devices. The authors explain that there are methods to circumvent overload through converter design. Two methods for this are creating additional headroom in the D/A converter itself or adding digital headroom before the signal reaches the D/A converter through allocating bits, dedicated for headroom. Based on the tested devices of this study, it appears that such prevention methods were not employed. Though it should be noted that the device's age range, span from 13 to 4 years old (as of the writing of this paper), and these methods could be more widely adapted in current devices. To draw more generalized conclusions regarding the prevalence of these techniques in commercial devices, a larger sample size would need to be tested.

4 Listening Test Method

4.1 Audibility & Preference

To evaluate the audibility claim made by mastering engineers regarding the prevention of overload, a listening test was conducted using the ABX method. However, relying solely on the parameter of audibility may present a skewed view. This is because the perception of distortion may not necessarily be indicative of a level of concern that necessitates disregarding the claim. Additionally, the distortion artifacts created by overload may be perceived by some participants to subjectively improve the quality of the music or listening experience. To address these concerns, the parameter of preference has been added to the test.

In the listening test, participants will be presented with three stereo channels, A, B and X. A and B will always be different versions and one of them will always be the same as X. Participants will be asked to identify if A or B corresponds with X and which they

Table 3: Stimulus measurement

Length	04:06
Format	WAVE
Sample Rate	44.1 kHz
Bit Depth	24-Bit
Integrated LUFS	-4.5
dBTP Max, L	2.0
dBTP Max, R	2.3
Occurences over 0 dBTP, L	40118
Occurences over 0 dBTP, R	40648

prefer. To this question participants get three choices, A or B, as well as the opportunity to answer that there is no preference. By including both audibility and preference parameters in the test, the study will provide a more comprehensive understanding of the potential impact of overload on listener perception and preference.

4.2 Stimulus

There are limitations in the use of static sine waves as stimuli to study audibility. Firstly, static sine waves do not adequately represent mastered music, which is the area of interest for this study. Secondly, static sine waves continuously cause the converter to overload, which is not reflective of the way in which mastered music overloads.

To address these limitations, the author of this study obtained a mastered song in the genre of interest from a professional mastering engineer. The song was then digitally measured in Pro Tools using VoxengoSPAN, and YOULEAN LOUDNESS METER 2[2] to ensure that it was suitable for the study. Specifically, the measurements were used to verify that the song had the desired levels of True Peak. Table 3 provides detailed information on the stimulus used, including its length, sample rate, bit depth, integrated LUFS, dBTP Max and number of occurrences of levels above 0 dBTP.

The song underwent editing to produce six distinct sections, each of which was used during the listening test. These sections were derived from the song's intro, chorus, last chorus, verse, solo, and outro. The musical intensity of each section varied, with the highest level of intensity occurring in the last chorus and the lowest level in the outro. The decision to use multiple sections rather than playing the song in its entirety was based on the desire to randomize the order to avoid order error.

The song consisted of drums, bass guitar, keyboard, and distorted guitars in all sections except for the outro. Vocals were present throughout the song, except for the initial part of the intro and the entirety of the solo. The two choruses featured extensive use of backup vocals/choir parts. The outro section was solely comprised of keyboard, vocals and backup vocals/choir.

In addition, the first, second and fifth section in the order was repeated at the end of the test. The specific section varied between participants because of the randomized order.

4.2.1 Genre

Initially, multiple genres were considered for inclusion in the study. However, to reduce the number of tests each participant needed to do, the decision was made to focus solely on hard rock/metal music. This genre is often mastered to be loud and is therefore suitable for the test. It is worth noting that the use of hard rock/metal music as the sole genre in this study may limit the generalizability of the findings. Specifically, other genres that do not utilize distortion to the same extent as hard rock/metal may produce different results. Future research could expand on this test by including a broader range of musical genres.

[2] YOULEAN LOUDNESS METER 2 https://youlean.co/youlean-loudness-meter/

4.3 Participants

A total of twenty-one participants was recruited for the listening test, all students at the School of Music at Piteå. Of the twenty-one participants, twelve were studying sound engineering while nine were studying various musician focused educations. This population represents intermediate listeners, being more aware of concepts such as distortion than the average listener, but not pertaining the expertise of professional listeners with multiple years or decades of experience. The year of study for each participant varied between year 1 and year 3. No data was collected regarding participants previous listening experience and will therefore vary as some enter education programs with a lot of previous experience, or non. Participation in the study was voluntary, with no compensation provided to the participants.

4.4 Setup

The equipment used in the test included a Pro Tools system, an SSL Duality console, Avid's HD I/O and ATC SCM50ASL speakers. The choice of converter was based on the results of the previous objective measurements, which showed that the HD I/O distorted the most when subjected to overload. The equipment was located in Control Room 1 at the School of Music in Piteå, where the test was conducted.

The stimulus used in the test was duplicated into two stereo tracks: one for playback at 0 dBFS, which created overload in the converter, and another at -6 dBFS that did not cause any overload. Four different outputs from the HD I/O converter were utilized, two channels for stereo playback at a -6 dBFS output and two channels for stereo playback at a 0 dBFS output. These channels were routed to the SSL Duality console, where volume normalization was performed (see the section below).

To enable participants to switch between the channels labeled A, B, and X, the DAW Control function on the SSL Duality console was used. Three VCA groups were created in Pro Tools for A, B, and X. Participants could use the DAW Control to solo between the three groups at their discretion. The Pro Tools X-OR solo function was also employed to eliminate the need to activate and deactivate solo functions simultaneously.

4.4.1 Volume Normalization

Several methods were attempted for volume normalization. Due to the nature of the test, being converter overload, digital normalization was not feasible. Initially, a nulling method was employed using sine waves obtained from the objective measurement tests. This method involved manipulating the faders and flipping the polarity of one stereo pair until the lowest output was achieved. However, this method was problematic as the stimulus does not continuously overload the converter, resulting in poor volume normalization.

In the final test, a combination of three methods was employed, nulling the stimulus, dB metering, and subjective evaluation by ear. Nulling the stimulus was utilized as an initial, rough method. In the second stage, a MiniLyzer ML1 using an A-Weighted SPL measurement, was used to fine-tune the levels. Finally, the author relied on subjective evaluation by ear to correct significant apparent issues before refining the levels using the previous methods. Despite these efforts, perfect normalization could not be achieved, which implies that a small volume difference between A and B will always be present.

4.5 Procedure

Participants were invited to a control room where they were informed about the test protocol. Specifically, the ABX-method was introduced, and participants were instructed to indicate if A or B was identical to X, as well as to express their preference for either A, B, or neither. Participants were also requested to provide additional information regarding their preference choice and were given a question about test difficulty after completing the entire test.

Prior to the main experiment, a pre-test was conducted to determine whether two pieces

Table 4: ABX results for each section

	Intro	Verse	Chorus	Solo	Last Chorus	Outro
Trials	31	30	28	36	31	33
Successes	13	13	13	23	17	12
Cumulative Probability	85.9 %	81.9 %	71.4 %	6.6 %	36.0 %	96.0 %

Table 5: Preference results for all trials

	-6 dBFS	0 dBFS	None
Count	58	64	67
Percent	30.7 %	33.9 %	35.4 %

of information, in addition to an instruction, were necessary for the main experiment. It was evaluated if informing the participants about the difference in distortion and volume was necessary. Additionally, if participants should be able to decide master volume themselves.

The difference between A and B proved to be small, participants in the main test were therefore informed that the difference between A and B was related to distortion. In addition, participants were told about potential issues with volume normalization, being asked to focus on distortion, this was done to decrease the chance of participants getting the correct answer based on volume differences alone. Individual adjustments of master volume during the pre-test, proved to pose a significant challenge for volume normalization, as the differences increased with master volume, it was therefore decided that all participants would listen to conduct the test at the same volume.

5 Listening Test Results & Analysis

Null hypothesis: There is no audible difference between versions.

To be able to determine if the null hypothesis can be accepted or rejected, cumulative probability analysis was utilized. The threshold required to reject the null hypothesis is set to $p < 5$ %. This threshold is commonly used in scientific research to indicate statistically significant results.

Table 6 presents the results of the analysis where all sections with both sound engineer students and music students are included. Results show that the cumulative probability values for each section were well above the $p < 5\%$ threshold required to reject the null hypothesis. This suggest that any perceived differences between the two versions were likely due to chance, rather than participants being able to perceive a difference and the null hypothesis can therefore be accepted.

The results for preference, as presented in Table 5, indicate a lack of significant preference for either version. The participants' inability to differentiate between the two versions in the audibility test implies that their preference is based on their ability to detect a difference. It is possible that participants may have preferences for factors beyond audibility for the distortion created by overload. Additionally, it is important to note that while no significant difference was found in overall preference, no clear individual preference was captured either.

Cumulative probability analysis for the separate sections of the stimulus is shown in Table 4. Notably, the solo section exhibited a near statistically significant result. With exception of the initial parts of the intro section, the solo section is the only section where vocals are absent.

Table 6: ABX results for all trials

	All	Sound Engineers	Musicians
Trials	189	108	81
Successes	91	51	40
Cumulative Probability	72.0 %	75.0 %	58.8 %

Table 7: Preference results for each section

Count	Intro	Verse	Chorus	Solo	Last Chorus	Outro
Trials	31	30	28	36	31	33
-6 dBFS	7	7	9	15	7	13
0 dBFS	11	9	10	11	14	9
None	13	14	9	10	10	11

Table 8: Changes in repeat sections

Count	Audibility	Preference
Unchanged	31	24
Changed	32	39

The lead guitar in the solo section possesses a high amount of treble to distinguish itself and capture the listener's attention. Consequently, the more focused sound, increase in treble, coupled with the absence of vocals, may offer insights into the results.

Furthermore, the overall preference for each version was analyzed for the separate sections, with results indicating no clear preference between the two versions, as indicated in Table 7. However, the free-text responses of participants who preferred the -6 dBFS version on the solo section revealed that higher clarity was the primary reason for their preference, accounting for 11 out of 15 responses. This, in conjunction with the cumulative probability, suggests that some participants could have perceive a difference. It should be noted, however, that the sample size for each section was relatively small, and the findings may still be attributed to chance.

The number of times participants changed their answers in the repeated sections was analyzed and reported in Table 8. Results indicated that participants changed their answers for the audibility test seemingly randomly, which reinforces previous indications that participants were unable to perceive a difference between versions. However, participants' preference changed to a higher extent than in the audibility test. A possible explanation for this is as that the original sections come at the start of the test and the repeats at the end. Results could be due to fatigue or a refinement of preference as participants had more exposure to the material, it is however not reflected in the audibility test changes.

At the end of the test, participants were asked to rate the difficulty of the test. The majority of participants (19 out of 21) reported that the test was extremely hard, while the remaining two participants reported that it was neither hard nor difficult. Notably, participants who reported that the test was neither hard nor difficult did not produce statistically significant results, indicating that they did not perceive a difference.

6 Discussion

The act of intentionally leaving mastered material with the potential for overload may appear counterintuitive when considering relevant theories. However, mastering is a subjective and artistic process in music production that involves collaborative decision-making among relevant parties, such as artists, mix engineers, and labels. Although parts of mastering involve a more theoretical process, some may believe that there are right and wrong approaches, yet the goal is to achieve the desired artistic vision.

Some professional mastering engineers choose to ignore the True Peak value because they believe that the prevention of overload results in a sound quality that's worse than leaving it with overload. This is supported by the theory that high frequencies are primarily responsible for overload. The act of True Peak limiting therefore means limiting the high frequencies. Additionally, some mastering engineers may ignore the True Peak value due to the claims made, without any subjective opinions about True Peak limiting.

The objective measurement test conducted in this study showed that consumers who do not maximize the volume of their devices (with digital volume control) will never experience overload. Therefore, it is recommended that consumers refrain from doing so. Consumers who do max their volume often do so in noisy environments, such as in public transport, where the small effect of overload on sound quality would be negligible. From the limited amount tested devices, the study saw that consumer D/A converters do not have built-in headroom for overload prevention, which is a feature that manufacturers could easily add by bit allocation. However, adding this feature would result in an overall loss of volume.

The listening test conducted in this study revealed that participants with intermediate listening experience generally cannot distinguish between a version of a song that is subjected to overload and one that is not. Additionally, no clear preference for either version was observed. However, participants were close to statistical significance between versions in one of the songs sections. It is difficult to draw reliable conclusions from the limited sample size and unclear preference regarding the section, possible explanations include the lack of vocals and focused guitar solo.

It is possible that converters may exhibit different distortion characteristics. The listening test may have employed a converter with distortion characteristics that were inaudible. However, since the test utilized the converter with the highest THD+N, it is likely that this possibility was avoided.

Participants had a higher chance to answer correctly than incorrectly. This is attributed to the flaw in volume normalization during the test and the fact that the versions corresponding to X and X itself, were always played through the same fader. Meaning that a difference in volume between A and B exists, but not in the version corresponding to X and X itself. Therefore, a correct answer could be due to chance, hearing a difference between versions, or simply hearing differences in volume. Conversely, an incorrect answer is mainly attributed to chance. As a result, weighing the cumulative probability to 0.5 may have been unrepresentative of the realistic chances in the test.

Participants were informed that the difference between the versions was the amount of distortion between them. This information was intended to make it easier for participants to identify any audible differences between the two versions. However, it is possible that some participants may not have been familiar with the concept of distortion. Despite this, no comments or feedback were received from participants suggesting a lack of understanding about the concept of distortion. Furthermore, it is unlikely that the results would have been significantly different if participants had not been informed about the difference in distortion. This is because no audible difference was perceived. However, it could explain the results of the solo section. In a hypothetical situation where participants were able to perceive a difference, repeating the test without informing about the distortion might have yielded a different answer.

The inclusion of only one genre and song in this study yielded limited results, and therefore, conclusions regarding the subject of overload prevention cannot reliably be made. Genres with less distortion and lower musical intensity may be subject to more distinguishable distortion characteristics from overload. However, genres with very low intensity, such as jazz or orchestral music, are less likely to be affected as they are generally not mastered for high loudness, and issues with overload are therefore less likely to occur.

This study did not take codecs into account, nor did it investigate how popular streaming services handle music. As codecs have the potential to significantly increase the peak value, overload through streaming services and perceptual codecs may be audible. It is also unclear how wireless formats, such as Bluetooth, affect peak values. Further research is recommended on these topics, given the popularity of these formats among consumers.

7 Conclusion

Based on the limited amount of tested devices, the claim of modern D/A converters having sufficient headroom to never overload, is false. For a hard rock song in an uncompressed 44.1 kHz sample rate WAVE format, overload is generally not audible, nor is there a clear preference between a version with overload and one without. Devices with digital volume control will never overload unless their volume is maximized.

Conclusions regarding the relevance of overload prevention cannot be made without further research into multiple genres, the effects of codecs and streaming services, Bluetooth and an investigation into overload prevention methods.

8 Acknowledgement

I would like to express my sincere gratitude to my supervisor, Natanya Ford, for her support, guidance and constructive feedback throughout this research.

I am also grateful to my professors, Jan Berg and Nyssim Lefford, for their insightful advice, support and information.

I would like to give a special thanks to Thomas "Plec" Johansson, who not only inspired me, but also introduced me to the subject of overload.

9 References

Dash, I. (2014, April). True Peak Metering—A Tutorial Review. In *Audio Engineering Society Convention 136*. Audio Engineering Society

Deruty, E., & Tardieu, D. (2014). About dynamic processing in mainstream music. *Journal of the Audio Engineering Society*, 62(1/2), 42-55.

ITU-R. (2015, October). *BS.1770-4: Algorithms to measure audio programme loudness and true-peak audio level.* International Telecommunication Union.

Lund, T. (2006, October). Stop counting samples. In *Audio Engineering Society Convention 121*. Audio Engineering Society.

Luther, C. (2016, August 3). Intersample Peaks And Beyond *[Graph]*. *The Science of Sound*. https://science-of-sound.net/2016/08/sh-html/

Luther, C. (2016, August 3). Intersample Peaks And Beyond. *The Science of Sound*. https://science-of-sound.net/2016/08/sh-html/

MixbusTv. (2020, November). *Inter-sample Peaks in Mastering: Why Nobody Cares? [Video]*. YouTube. https://www.youtube.com/watch?v=SskN8cRNzkI&ab_channel=MixbusTv

Nielsen, S. H., & Lund, T. (2003, May). Overload in signal conversion. In *Audio Engineering Society Conference: 23rd International Conference: Signal Processing in Audio Recording and Reproduction*. Audio Engineering Society.

Pohlmann, K. C. (2010). *Principles of digital audio* (6th edition) (pp.77-124). McGraw-Hill Professional Publishing.

Vickers, E. (2010, November). The loudness war: Background, speculation, and recommendations. In *Audio Engineering Society Convention 129*. Audio Engineering Society.