THE 'BASS TRANSMISSION INDEX': A NEW CONCEPT FOR EVALUATING LOUDSPEAKER PERFORMANCE

LE Harris ISVR, University of Southampton, UK *Can now be contacted at the Acoustics Research Group, University of Salford, UK: L.E.Harris@salford.ac.uk*
PR Newell Consultant, Moaña, Spain
KR Holland ISVR, University of Southampton, UK

1 INTRODUCTION

Loudspeaker design is a trade-off between optimising different aspects of performance with regard to the intended application. For musical mix monitoring, the ability to accurately reproduce the temporal envelope of dynamic signals is a key factor in helping to produce a balance of instruments which will transfer well to other reproduction systems. Despite this, the industry standard still prioritises representations of the extension and flatness of the frequency response magnitude over transient accuracy, even for loudspeakers intended for professional use. This paper presents a method called the Bass Transmission Index (BTI), which aims to redress this balance and make it easier to evaluate and compare loudspeaker reproduction accuracy at low frequencies: a part of the audio spectrum that is crucial to most types of modern recorded music. Background and key information on the development of the method is presented, followed by an explanation of example data returned from analysis of some measured mix monitors.

2 FOUNDATIONS OF THE BASS TRANSMISSION INDEX

2.1 Mix Monitors and Audio Mush

Mixing is the stage of musical production where each of the recorded components are adjusted individually and then blended together. The aim of this process is to create a sonically-pleasing balance which conveys the emotion and impact intended by the musicians. This frequently takes place on smaller loudspeakers than when recording or mastering, mounted in close proximity, and often on top of the recording console itself. By this means, if the sound is relatively free from the characteristics of the of the mixing room and the 'as-mounted' response is 'flat and fast', many people have found that the method can help them to create mixes with a balance between instruments which 'travels well' to different listening environments[1,2]. Loudspeakers (or 'monitors') used for this purpose are therefore professional tools which assist the engineers to more reliably achieve their desired balances before they are sent to the final stage of mastering.

Toole[3] described neutral monitors as 'a transparent window into the art'. When the reference for the recorded sound is true, it is likely that the resulting mix will transfer well to other reproduction systems. Low frequencies in particular require accurate reproduction because, for many types of music, this is where the tonal and rhythmic foundations are laid. If these elements are not transferred coherently through the playback system in both time *and* level, then the music can lose impact, and the instruments can become confused[4.] The resulting reproduction becomes more 'soft focus' than reality; resembling a form of 'audio mush'[5].

2.2 Loudspeaker Design at Low Frequencies

Thiele[6] described a 'constellation' of parameters that combine in a specific way to determine the overall shape of a loudspeaker's response at low frequencies, but the discussion here is limited to one fundamental design feature that influences this response shape, or *alignment*: design of the cabinet. Most loudspeaker cabinets can be classed as either sealed or ported (also known as vented, or bass reflex). The ported cabinet uses a mass of air, tuned to resonate at a particular frequency, such that it reinforces the acoustic output at the point where that of the driver begins to decrease. However, resonant systems are 'slow to start and slow to stop'[7]: they take time to react when a driving force is both applied and removed. Thus, the low frequency response in a ported loudspeaker is augmented at the cost of degraded transient performance. This design trade-off is typical in small to medium size loudspeakers, where it is often desirable to get more bass from a smaller cabinet than could be achieved with a sealed box of the same size.

Mathematical models for loudspeaker design were developed in the 1970s, when Small[8] demonstrated that knowing certain parameters of any given drive unit would allow accurate prediction of the overall system response when it was placed in different types of cabinet. This removed much of the guesswork from making loudspeakers perform well at low frequencies, but it inevitably still remains the case that creating a well-behaved bass-reflex system is more difficult than for a sealed-cabinet design. It is not necessarily the case that 'sealed is good and ported is bad', because the trade-off between higher-sensitivity and transient-accuracy depends largely on the application. Loudspeakers are made to fulfil a variety of requirements depending on their intended use. Accordingly, therefore, it is the designers' job to decide which aspects of performance they want to prioritise[9]. However, poor temporal accuracy has consequences that are particularly detrimental in the case of professional mix monitoring. The impact is worsened by the fact that bass instruments are fundamental components of the rhythm section in many types of music, but their dominant content is in the region where resonances in smaller monitors are most likely to occur.

2.3 Rapid Roll-Offs and Bass Reproduction

As a rule of thumb, sealed-cabinet loudspeakers have a 2nd-order (12 dB/octave) attenuation below their cut-off frequency, and ported-cabinet designs have a 4th-order (24 dB/octave) roll-off[10,11]. The addition of protection filters to limit excessive driver excursion can increase roll off to 6th order or more[12]. There will also be varying amounts of damping, with bass-reflex systems commonly having a slight hump around the turnover frequency. It can generally be assumed that loudspeakers at low frequencies are minimum phase systems[8,10,13], which means that both parts of the complex frequency response are interdependent, such that rapid deviations in magnitude through the bass region have corresponding steep shifts in phase[14,15].

It is not intuitive to visualise how the phase component of a loudspeaker's frequency response affects a harmonically-complex signal passing through it, but Preis presented a rigorous theoretical discussion of the relation between phase anomalies and poor transient performance in various types of audio system, including loudspeakers[13,14]. A linear phase characteristic, where the phase shift is proportional to frequency, is equivalent to delaying the signal as a whole as it passes through the system because the constituent elements are delayed in correct proportion to each other[16,17]. By contrast, frequency-dependent phase shifts, such as those accompanying resonances and rapid roll-offs in a loudspeaker alignment, change the relative delay between individual components in a signal, rather than delaying it as a whole. This effect, commonly known as group delay, is also referred to as envelope delay, or dispersion[16,18], where the shape, or envelope, of a signal in the time domain is changed such that the acoustic output is not simply an amplitude-scaled version of that presented at its input terminals. In speech, this may lead to a loss of intelligibility, but in music it can produce a number of perceptible changes, such as degrading the impact of the musical presentation by affecting the timbre, rhythm, and even the pitch of the bass instruments. There is evidence that the *perception* of these effects is greater at lower frequencies and higher SPLs[19–23], so monitors that cannot faithfully reproduce the envelope of a signal may therefore make it very hard for mix engineers to create a ‘robust’ balance between bass instruments.

Perhaps the most problematic aspect of this distortion is its effect on the perceived relative levels between instruments such as resonant bass guitars and percussive kick (bass) drums. Monitors with an extended response but poor transient behaviour may lead to a mix which biased towards the bass drum because the comparatively steady-state notes of the bass guitar will sound louder due to the greater resonant enhancement, so the engineer will try to compensate in the mix by raising the level of the shorted duration, more percussive kick-drum[24]. This is very hard, if not impossible, to correct at the mastering stage, because instruments which occupy the same part of the audio spectrum cannot be adjusted independently with global equalisation. This can result in an underachieving musical product going out for distribution, or can necessitate an expensive and time-consuming stage of remixing. Thus, it can be seen why loudspeakers with significant differences in their bass alignments will not, indeed *cannot*, reproduce a piece of music with a similar level of accuracy.

2.4 The Yamaha NS10M: *Thirty-Seven* Years a Reference Monitor?

In order to produce mixes of a consistently high quality, there must be a degree of consistency across the monitoring loudspeakers, regardless of their design philosophy[25]. However, Newell *et al.*[26] presented clear evidence that there is an unacceptable lack of performance comparability between loudspeakers designed to perform the same professional function. They conducted in-depth comparative studies of small to medium-sized professional studio monitors, analysing performance in both the frequency and time domain through a combination of different measures. It was found that loudspeakers possessing flat and extended magnitude responses did not necessarily perform well in response to dynamic signals. One of the best performing systems in this respect was the Yamaha NS10M, a monitor that was small and relatively cheap (at the time it was produced), yet has remained a popular mixing reference for pop and rock recordings from the 1980s to the present day[27 ch.8]. Although its anechoic magnitude response showed it to be neither flat nor particularly extended, waterfall plots showed a rapid decay that was nearly uniform with frequency. Consideration of both amplitude and time response of this loudspeaker therefore showed that, in one respect, the NS10M was a poor performer compared to equivalent models, but in another it shared qualities with much larger, higher-quality studio monitors.

It was described in Section 2.2 that loudspeaker design at low frequencies is a design trade-off. Professional sound engineers may often prioritise accurate transient response over extension because they can learn to mentally adjust for broad deficiencies in bass output, but cannot imagine what the correct reproduction should sound like if presented with material affected by distortions due to temporal errors[28–30 ch.9]. Indeed, the British Broadcasting Corporation were selecting monitors with a particular focus on dynamic behaviour as far back as the 1940s, as they had noted that mixes produced on loudspeakers free from ‘transient troubles’ still sounded well-balanced when played back across a variety of reproduction systems[31,32]. Despite this fact, the industry standard is still focused on flatness and extension of the pressure-amplitude response.

In search of a different approach, Holland and Newell *et al.*[12,24,33,34] developed an approach for loudspeaker evaluation using the modulation transfer function (MTF). This appeared to be a promising method as it was observed that systems exhibiting long decays in waterfall plots returned low MTF scores, whereas those with a rapid decay across all frequencies scored more highly. Thus, it was concluded that this might be a quantitative way to assess the temporal performance of loudspeakers at low frequencies. This method was the basis for further investigation, and eventually led to development of the Bass Transmission Index, an MTF-based technique with parameters optimised for evaluating the bass reproduction accuracy in professional mix monitors[27,35,36 ch.11].

3 FEATURES OF THE BASS TRANSMISSION INDEX

3.1 The MTF, and Methods That Use It

The modulation transfer function exists in various forms and has a number of important applications. It is used in optics to assess image blur, in radiography to measure imaging resolution, and in audiology to quantify the limits of temporal discrimination in our hearing system[37–39]. The most prominent example of an MTF-based method in acoustic evaluation is the Speech Transmission Index (STI), a standardised method for evaluating speech intelligibility that emerged in the 1970s, and has been the subject of ongoing investigation and improvement ever since[40–42].

In practice, the MTF can be computed in several ways that do not necessarily produce equivalent results for the same system, so the computation method must be chosen with careful consideration of the intended application. However, by tailoring the analysis parameters, an MTF-based method can be developed which is optimised for a specific type of problem. As a result, a process developed for evaluating speech in rooms is not necessarily suited to assessing loudspeakers reproducing music at low frequencies, even though they may both be derived from the same concept. An MTF-based method may therefore be thought of as an application-specific version of the true MTF, computing results for a certain combination of test conditions that are known to be revealing for the type of system and signal being assessed.

3.2 Key Parameters in MTF-Based Techniques

Key parameters to consider when developing a successful MTF-based method are the range of frequencies covered (including how that range is divided up into bands for analysis), and the number and value of modulation frequencies. These ‘modulation frequencies’ are not the components of the signal itself, but the frequency-domain counterparts to its amplitude envelope; representing the rates at which any fluctuations in the shape of the test signal occur in the time domain. As described in Section 2.3, preserving the relative shape of a waveform is an essential part of accurate reproduction, so a system should logically be analysed using signals that feature temporal fluctuations similar to those found in the envelopes of programme material it would typically be expected to reproduce.

Finding the optimum combination of band and modulation frequency settings is a somewhat empirical process, and may be subject to change in response to further validation data. However, selecting a method for implementing the MTF is perhaps the most crucial decision, because it has a major impact on the nature of results, and because it requires a clear understanding of the intended application. Harris[36 ch.2.3] outlined four ways to implement the MTF, which were classified according to a fundamental difference in approach. The first pair of methods used a band-limited impulse response of the test system (loudspeaker), referred to as BLIR methods; the second pair, instead, used a band-limited input (test) signal, referred to as BLIP methods. It was shown that each class of method produces fundamentally different results. The output from each approach was compared based on two aspects identified as critical in the target application: i) inherent band-limiting error, shown by any deviation from a score of 1 in results for a simulated perfect system, and ii) responsiveness to changes in a loudspeaker's low-frequency alignment. A BLIP method was chosen for evaluating loudspeakers at low frequencies as this was found to fulfil both requirements more effectively than BLIR methods, including the Schroeder equation used in preceding work[33,35].

A detailed discussion of the features of MTF-based methods, and how these were selected for the BTI, can be found elsewhere[36 ch.2], but one particular aspect of the technique is worth highlighting briefly here: the implications of amplitude modulation at low frequencies.

3.2.1 Understanding Amplitude Modulation at Low Frequencies

The principles of amplitude modulation (AM) are found in most introductory textbooks on communications theory. It is demonstrated that a band-limited message signal modulates a much higher frequency carrier sinusoid, f_c, with the result being a message translated in frequency. In the typical form of AM, the message band is mirrored equidistantly above and below f_c. Amplitude modulating a signal in the time domain therefore changes its frequency spectrum. The modulation sidebands are also repeated either side of $-f_c$, something that can be conveniently ignored in most applications. However, standard texts do not show what happens when low-frequency bands are amplitude-modulated by very low frequencies: effectively the case where f_c is lower than the spectral content of the 'message'. Harris[36 ch.2.2.2.2] demonstrated that, in this case, the sidebands around f_c and $-f_c$ partially overlap to produce an aliasing effect which gives the impression that there is only one sideband above 0 Hz, and which apparently covers a wider range of frequencies than the original signal.

The 'effective bandwidth' issue was significant in development of the BTI because of the intended application, where the analysis bands were necessarily at low frequencies, and the sinusoids that must modulate them (common modulation frequencies in music) were found to be comparatively very low. This meant that any loudspeaker under test would effectively be excited by frequencies outside the intended analysis band for some methods of computing the MTF. Understanding this issue led to the consideration of potential redundancy in the results, and therefore influenced the choice of analysis bandwidths and arrangement.

3.3 BTI Parameters

The aim for the BTI was to develop a method that was revealing of loudspeaker performance with minimal redundancy, so that features were clear and comparisons could be made easily across different models. The focus was on accurate reproduction of the key bass-content in music. Therefore, the BTI analysis range covers approximately the first three octaves of the audio spectrum, divided into ten linear bands. Seven modulation frequencies are used, all found to be common in commercial music recordings across a range of genres. The chosen parameters were developed through a combination of the literature evidence, the evaluation of envelope spectra from 168 musical extracts, and the objective comparison of results from a group of different loudspeakers, both measured and simulated.

Selection of the MTF computation method was based not only on objective research and validation, but also on data from perceptual evaluation in listening experiments. From these, it was concluded that the chosen procedure, unlike the approach used in preceding work, was effective in predicting ordinal listener judgements of the music reproduced through a range of simulated loudspeakers, differing only in their low-frequency alignment[36 ch.8.1]. The same conclusion was reached when evaluating excess-phase systems, where the magnitude was fixed but increasing low-frequency phase shifts were simulated. The findings from this latter experiment were notable, as it was observed that phase distortion of the order introduced by loudspeakers at low frequencies is definitely audible, even to the majority of non-expert, untrained listeners when using music for audition. Although the effect is strongly programme-dependent, it appeared that audibility of increasing phase shifts in the bass response of loudspeakers depends on the temporal characteristics of the material, as well as on the overall spectral (bass) content – findings that are consistent with those reported by other authors[43–45]. This remains an important area for further investigation, as it may be possible to define which aspects of music are more likely to elicit the audible effects of temporal distortion in loudspeakers, and it may thus be possible to determine if certain styles or genres of music would be more robust if mixed on monitors that return a particular type of BTI result (discussed in Section 4).

3.4 BTI Output: Results and How to Interpret Them

A key feature of the BTI is the way it presents results. The MTF returns a score between 0 and 1, where a higher number signifies greater accuracy: i.e. the system has less of an impact on the signal passing through it. This simple and intuitive format makes it appealing for diagnostic applications because it allows direct quantitative assessment and comparison of systems. As MTF-based methods divide up the analysis into application-specific frequency bands and modulation frequencies, the result of computation is actually a matrix of individual modulation-index scores, one for every combination of the chosen test parameters. As this is hard to interpret without careful inspection, the results matrix is usually averaged, and sometimes weighted in a defined way to return an overall single figure of merit.

When developing the BTI, it was deemed essential that the results were informative whilst also being simple to interpret and compact enough to display on product-specification sheets. A distinctive visual output was developed which displays the output data in the form of a greyscale heat map, termed *intensity images*. Each ‘pixel’ in an intensity image represents one element of the numerical MTF matrix, mapped such that a score of 0 is black, and 1 is white, with various shades of grey representing the values in-between. Work preceding the BTI had used line plots of average scores within each frequency band[28,33], but it was eventually concluded that the intensity images were much more revealing of a loudspeaker’s behaviour whilst still being relatively simple to interpret.

Validation of the BTI method, with both simulated and real loudspeaker responses, revealed characteristic behaviour in the intensity images. It seems that certain BTI shading patterns are associated with particular features of a loudspeaker's low-frequency response. These have not yet been formally characterised according to loudspeaker design parameters, but several rules have been established on an empirical basis. These are discussed further in Section 4, but interpreting the trends within a BTI intensity image can be broadly summarised in three main points:

1. Consistent shading in the *horizontal* direction (frequency band axis) indicates consistent reproduction level with frequency (a flat response).
2. Consistent shading in the *vertical* direction (modulation frequency axis) indicates faithful envelope transmission (wave shape is preserved).
3. Lighter shading represents higher scores.

4 EXAMPLES: EVALUATING REAL LOUDSPEAKERS

This section describes some examples of BTI results for a number of professional monitoring loudspeakers. The focus is on demonstrating the differences in behaviour for the key design decisions that were described in Section 2, which affect the responses of loudspeakers at low frequencies.

4.1 Low-Order Alignments

Figure 1 shows BTI results for three sealed-cabinet loudspeakers with very similar low-frequency alignments. The frequency bands and modulation frequencies are on along the *x*- and *y*-axis respectively. The intensity images tell us about the extension and flatness of the steady-state frequency response, as well as indicating how these loudspeakers respond to dynamic signals passing through them. In general, the results for sealed-cabinet monitors produce characteristic BTI shading patterns similar to that shown in Figure 1, with consistent vertical banding and a gradual linear transition, horizontally, from black to white, through increasingly lighter shades of grey. This shows that the response has a very gradual rate of attenuation, typical of that seen for second-order roll-offs, and also demonstrates that the decay is relatively smooth, with few bumps or ripples. The transition to grey starts in the upper bands, a sign that none of these monitors exhibit very extended low-frequency responses.

It is important to note the almost total consistency in vertical shading for all of these loudspeakers. This direction reveals how the performance varies with the modulation-frequency within a given band. The absence of variation in vertical shading within a given frequency band indicates that all of these loudspeakers are able to accurately reproduce amplitude modulation across the range of values reasonably expected for musical signals. Some older, under-damped sealed box designs could exhibit a more resonant behaviour, but in general, the dominating factor in the alignment of modern, professional sealed-cabinet loudspeakers is the cut-off frequency. The BTI images show little to no vertical variation in modulation index (and therefore shading); the only feature of interest is therefore the frequency band in which white starts transitioning to grey, i.e. the region where the output begins to decrease. However, the behaviour in the intensity images becomes more complex when other types of low-frequency alignment are considered, as explained in Section 4.2.

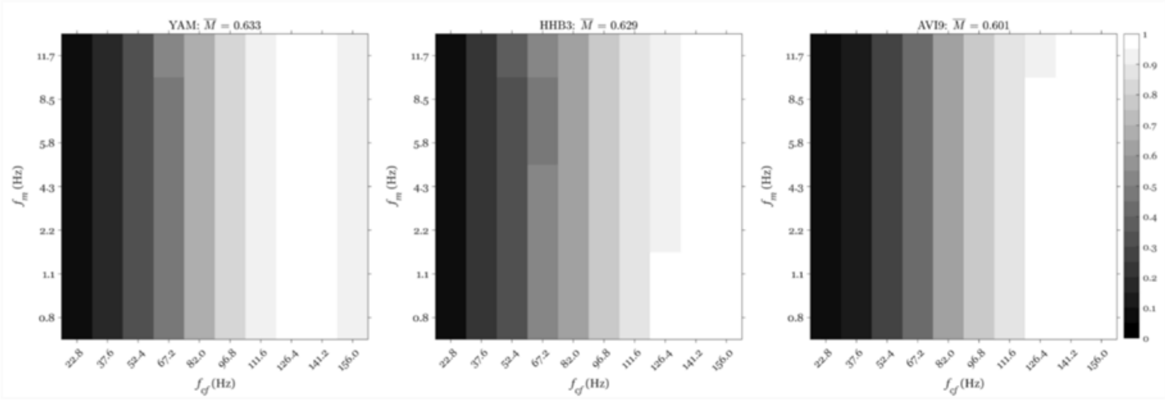


Figure 1. BTI results for three sealed-cabinet professional monitors: Yamaha NS10M (mean score = 0.633), HHB Circle 3P (0.629), and AVI Pro9 (0.601).

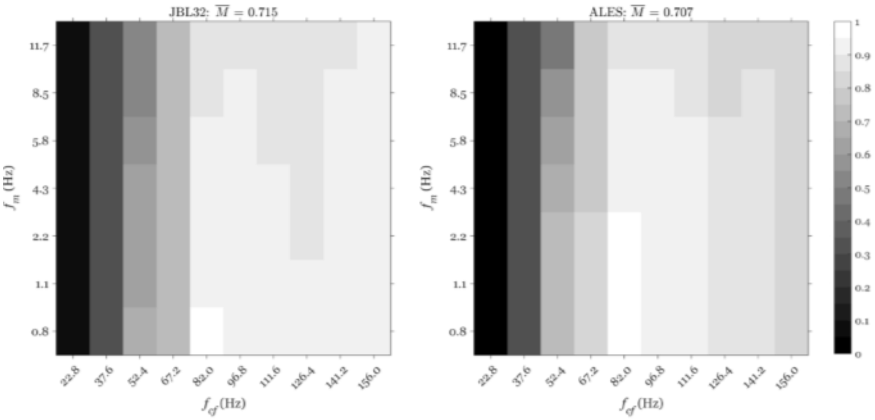
4.2 Higher-Order Alignments

Figure 2 shows the BTI results for six loudspeakers that all have bass-reflex cabinets, and a generally similar range of design parameters, with minus 10 dB points between 30 to 40 Hz, although the roll-off shapes and slopes are quite varied. In the figure, they are grouped into pairs that have certain similarities in the pattern of their intensity image shading.

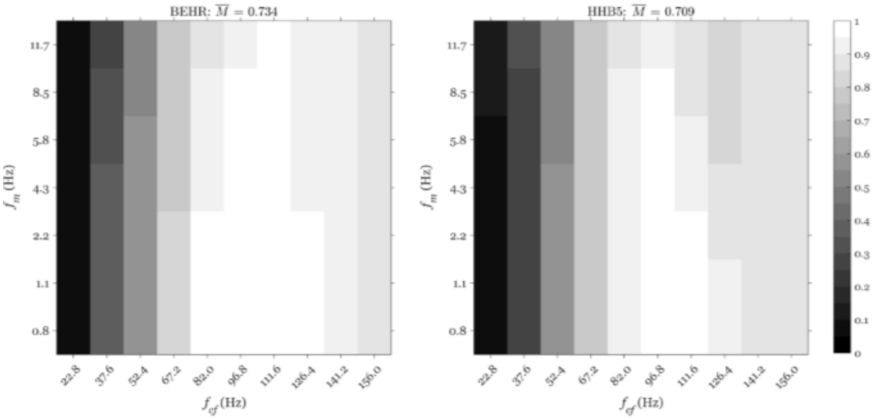
There are some general observations to be made about these loudspeakers as a group. The first feature to note is the rate at which the performance changes along the frequency band (*x*-) axis. Models with steep roll-offs below their cut-off frequency show sudden horizontal transition from light to dark, with few shades of grey separating them. This transition is most abrupt in designs that use electrical protection filters which increase the rate of roll-off to 6th-order or more.

The models in Figure 2 also exhibit a common characteristic that is found in any loudspeaker lacking a well-controlled bass response: shading inconsistency in the vertical direction. Deviations in vertical shading show that the modulation index scores are varying with modulation frequency (*fm*) inside a given test band. The examples in Figure 2(c) demonstrate this most clearly in the 38 Hz and 56 Hz bands, respectively, with the shading growing steadily darker as *fm* increases. This implies that the system may perform adequately for slowly varying input signals, but less well for more rapidly modulated inputs. As described in Section 2, a loudspeaker exhibiting this type of behaviour is unlikely to be a reliable reference when used for the purpose of musical mix monitoring.

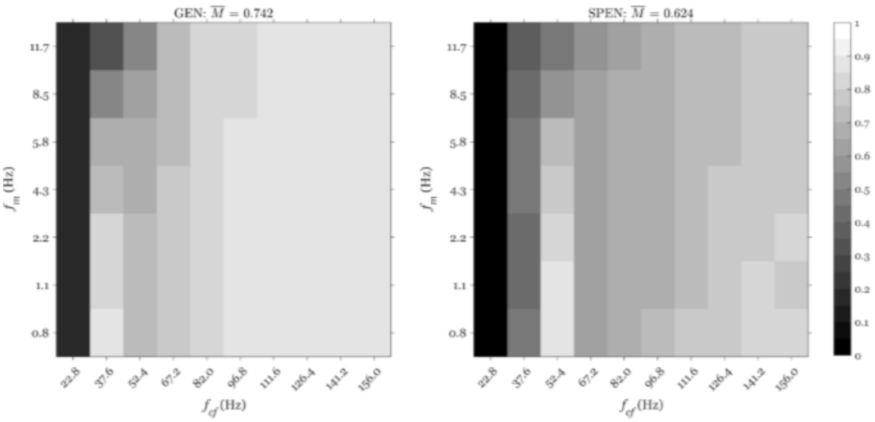
Comparison of Figures 1 and 2 illustrates that the assessment of vented-box loudspeakers is much more complex than that of sealed-cabinet systems. Performance may vary across both axes, changing in ways that appear unpredictable. There may be isolated 'bright spots' within a dark region, protrusion-like areas of constant high or lows scores, or large sections of consistently mediocre performance until a sudden descent into unacceptable temporal distortion. So, it seems unlikely that a single, simple mean-score is sufficient to adequately summarise this behaviour and allow reliable comparison with other loudspeakers. Section 4.3 explores this in further detail.



a) JBL LSR32 (0.715) and Alesis M1 Active (0.707)



b) Behringer TRUTH B2031 (0.734) and HHB Circle 5A (0.709)



c) Genelec S30D (0.742) and Spendor SA300 (0.624)

Figure 2. BTI intensity image results for six reflex-loaded monitors, paired by similar alignment. The mean BTI scores are given next to the model names.

4.3 Limitations of the BTI Mean Scores

4.3.1 Low Scores for Well-Behaved Loudspeakers

Comparison of Figures 1 and 2 shows a general tendency for well-controlled sealed-box loudspeakers to get relatively low scores. This is because they usually start rolling off toward the upper end of the bass region (unless they are extremely large: as is the case with some models designed for use in mastering). Conversely, loudspeakers with responses extended by using resonances produce results that exhibit the boost in output in all but the lowest frequency bands. That is to say, they show higher MTF scores across a greater number of bands than would be seen in comparable sealed-cabinet systems, which therefore results in a higher mean when averaged across the entire BTI matrix. However, because the bass extension is gained via design strategies that degrade the temporal accuracy, the apparent increase in performance does not extend to higher modulation frequencies. As a result, loudspeakers of this type exhibit the characteristic mottled (or scrambled) appearance in their intensity images, where the shading lacks vertical consistency. In such a case, if viewing only the averaged numerical result, there is no way to tell where any compromises have been made, so the full intensity images would be required in order to determine the details.

With this in mind, it is worth considering the intended mounting arrangement for a given loudspeaker. As described in Section 2.1, small monitors, intended for use at very close listening distances, are likely to be located on top of the mixing console. This provides a general boost to the level at low frequencies in a way that does not affect the temporal performance – quite unlike the introduction of a resonant port to reinforce the output when the drive unit naturally begins to roll off. The results of BTI analysis in this mounted condition, rather than the free-standing case, would then show an increase in the overall score, and a general ‘lightening’ in the intensity images, without introducing inconsistencies along the vertical axis. It may therefore be the case that loudspeakers such as those shown in Figure 1 could benefit from equalisation, or from reinforcement from console-top locations, whereas this may only further confuse the reproductions produced by systems displaying results resembling those in Figure 2. In the case of loudspeakers which did not have enough headroom to have the roll-offs boosted for practical use, it may be possible to predict the likely BTI score when flush-mounted, or placed on top of a sturdy console, by equalising the response to the ‘reinforced’ condition during anechoic measurements (albeit at lower sound levels).

4.3.2 Same Score, Different Behaviour

One more feature of the BTI numerical scores must be addressed here, because they have the potential to be misleading if they are presented independently from their intensity-image counterparts. In the case of loudspeakers with very similar types of alignment, the mean scores can reliably indicate which is the more accurate reproducer of bass content. However, it is not so straightforward when comparing models with contrasting low-frequency design characteristics. In this case, it is possible that two loudspeakers may get virtually identical overall scores, even though they behave quite differently at low frequencies. Figure 3 shows an example.

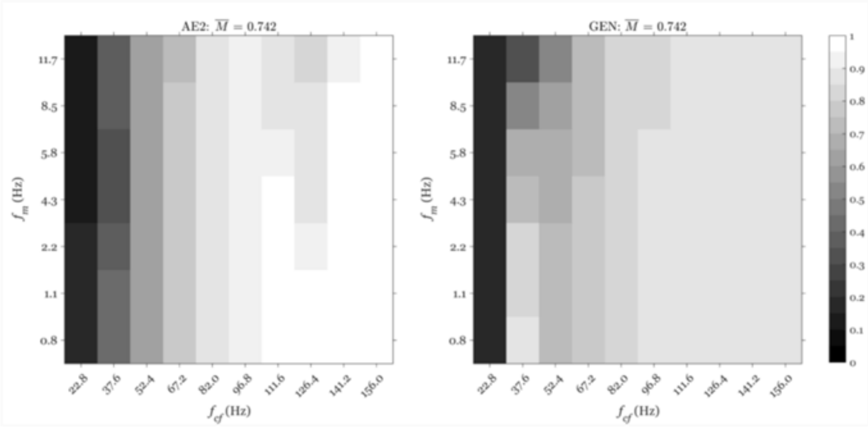


Figure 3. An example where the BTI returns an identical overall score (0.742) for loudspeakers with contrasting low-frequency alignments. Models: Acoustic Energy AE2 and Genelec S30D.

It may be deduced from the intensity images that the loudspeakers in Figure 3 have contrasting alignments. One will produce bass that is loud and low, but not necessarily accurate. Conversely, the other may sound bass-light in comparison, but what is actually there will otherwise be true to the recording. Therefore, the BTI *intensity images* give us a better understanding of how loudspeaker design trade-offs are more likely to alter a musical waveform, and a potential customer may favour one aspect of reproduction over another, depending to a large degree on their order of priorities. For example, for mix monitoring, a foremost concern often tends to be temporal fidelity, so in this particular application, vertical consistency will often be favoured over the horizontal 'whiteness' in the intensity images. Clearly though, without the visual information, it would be impossible to decide which of the monitors would be most suitable for mixing musical recordings if, based on only the numerical average, they appeared to be equally accurate reproducers of bass content. In practice, the only way to decide which of the monitors with such contrasting alignments would be most appropriate for any specific purpose would be by the audible evaluation by the intended users, but the BTI images can provide very useful, objective insight.

For the further development of the BTI concept, the judgements of a wide range of professional mix engineers would yield a greater understanding about the correlation between the perceived impression of bass reproduction accuracy and BTI results, and, in particular, the usefulness of the averaged numerical scores. However, this relationship has already been explored with non-expert listeners, and has produced promising results, but it was concluded that they struggled to compare alignments where the differences were more complicated than changes in the cut-off frequency[36 ch.7.5]. It is suggested that professional mix engineers, who are highly critical and experienced listeners due to the nature of their work, would be capable of the discrimination of more subtle differences in reproduction. Following their input, it may eventually be possible to map the numbers returned by the BTI against meaningful terms for bass reproduction accuracy.

5 CONCLUSIONS

The BTI is an MTF-based measurement method, developed to evaluate the low-frequency 'accuracy' of loudspeakers reproducing musical signals; specifically, with the application of professional mix monitoring in mind. The graphical representations (intensity images) do not fall into any easily definable 'order of merit' in the same way as the numerical scores, but they do provide a good representation of the relative merits of different performance characteristics in any given loudspeaker. Indeed, they have some distinct advantages over the numerical scores; at least in the present, unweighted, averaged form. They are also essential in understanding the differences between monitors that return similar or identical mean numerical scores, and have been demonstrated to be very effective in this case. The BTI concept undoubtedly still needs more validation to match scores with listener judgements, and in particular with its target audience, but the experiments, to date, have been very encouraging.

A 'reliable' loudspeaker (in terms of mixing consistency), even if it is short on overall bass level, will tend to give rise to mixes which can easily be corrected by equalisation at the mastering stage – or even by the tone controls on domestic music systems. That is to say, the mixes can easily be adapted to other reproduction systems without losing their essential character. On the other hand, many resonant loudspeaker designs, no matter how flat their 'frequency responses' may appear, can give rise to 'idiosyncratic' mixes which are the results of misleading deceptions in the time domain, and the 'smudging' of fine detail in the sounds being reproduced. Such loudspeakers are likely to lead to mixes which are not correctable by means of equalisation in the later stages of production, and 'lost detail' is something that simply will not be heard. The BTI response images present a significant new way of indicating such differences, and any loudspeaker which shows a poor BTI response will be unlikely to accurately represent the 'true' low-frequency balances between the bass instruments. This is especially so between mixtures of bass instruments with transient (percussive) sounds, and those with more resonant tones. Whilst these instruments may occupy the same frequency range, they tend to have very different temporal characteristics.

Obviously, no reliable artistic judgements can be made about things which cannot be accurately heard, so deceptions at the time of mixing can lead to either the failure to realise the potential of a mix, or the disappointments at the failure of a mix to translate well to other systems. In either case, the point to be made is that a loudspeaker which exhibits a poor BTIs will be scrambling the musical information, whereas a loudspeaker with a good BTI, even with a higher low-frequency roll-off, will simply be reducing the *level* of the musical signal at those frequencies: hence the possibility of a correction by subsequent low-frequency equalisation during mastering.

The BTI therefore offers a means of visually evaluating the likelihood of a loudspeaker being capable of the accurate reproduction of a musical signal in both the pressure-amplitude and temporal domains. The temporal response has been shown, especially since the 1980s, to be very important for achieving consistent balances between bass instruments of transient and resonant natures, such as between the bass drums and bass guitars which are essential parts of so much 'rhythmic' music, and which have suffered so much from 'unreliable' mixing conditions.

6 REFERENCES

1. Philip Newell and Keith Holland, *Loudspeakers: For Music Recording and Reproduction*, Chapter 8: Form Follows Function, Focal Press, 1st edition, 2007.
2. Andrew Goldberg, Aki Makivirta and Ari Varla, 'Compensating the Response of Near-Field Loudspeaker Monitors to Minimise the Effects of Desktop Acoustic Loading', *Journal of the Audio Engineering Society*, **54**(5), 401–411, May 2006.
3. Floyd Toole, 'Art and Science in the Control Room', *Proceedings of the Institute of Acoustics*, **25**(8), 243–252, 2003.
4. Robert Harley, 'A Guide to Better Bass', *The Absolute Sound*, **TAS 197**, November 2009. Available online: http://www.theabsolutesound.com/articles/a-guide-to-better-bass-tas-197-1/; accessed 18/02/2016.
5. Tony Andrews, 'Reality? Or Soft Focus?' *Presented at PLASA for the Gottelier Award Master Class*, 2008. Available online: http://www.funktion-one.com/dl/files/1819.pdf; accessed 18/02/2016.
6. Neville Thiele, 'The Loudspeaker Parameters and Their Evolution', *Proceedings of the Institute of Acoustics*, **31**(4), 27–49, 2009.
7. Phil Ward, 'Are Closed-Box Speakers Really Better Than Ported Ones?' *Sound On Sound*, October 2018. Available online: https://www.soundonsound.com/sound-advice/q-are-closed-boxspeakers-better-than-ported-ones.
8. Richard Small, 'Direct-Radiator Loudspeaker System Analysis', *J. Audio Eng. Soc.*, **20**(5), 383–395, June 1972.
9. Anthony Lucas, 'Sealed vs. Ported Enclosures', Available online: https://www.eminence.com/sealed-vs-ported-enclosures/; accessed October 2018.
10. Richard Small, 'Closed-Box Loudspeaker Systems– Part I: Analysis', *J. Audio Eng. Soc.*, **20**(10), 798–808, December 1972.
11. Richard Small, 'Vented-Box Loudspeaker Systems– Part I: Small-Signal Analysis', *J. Audio Eng. Soc.*, **21**(5), 363–372, June 1973.
12. Keith Holland, Philip Newell, and Peter Mapp, 'Steady State and Transient Loudspeaker Frequency Responses', *Proceedings of the Institute of Acoustics*, **25**(8),105–118, 2003.
13. D. Preis, 'Phase Distortion and Phase Equalization in Audio Signal Processing - A Tutorial Review', *J. Audio Eng. Soc.*, **30**(11), 774–794, November 1982.
14. D. Preis, 'Linear Distortion', *J. Audio Eng. Soc.*, **24**(5), 346–367, June 1976.
15. Marshall W. Leach, 'The Differential Time-Delay Distortion and Differential Phase-Shift Distortion as Measures of Phase Linearity', *J. Audio Eng. Soc.*, **37**(9), 709–715, September 1989.
16. Alan Oppenheim , Alan Willsky and Hamid Nawad, *Signals & Systems (2nd edition)*, Chapter 6: Time and Frequency Characterisation of Signals and Systems, Prentice Hall, 1997.
17. B.P. Lathi, *Modern Digital and Analogue Communication Systems (3rd edition)*, Chapter 3: Analysis and Transmission of Signals, OUP USA, 1998.
18. Richard Heyser, 'Loudspeaker Phase Characteristics and Time Delay Distortion: Part 1', *J. Audio Eng. Soc.*, **17**(1), 30–41, January 1969.
19. R. Mathes and R. Miller, 'Phase Effects in Monaural Perception', *J. Acoust. Soc. Am.*, **19**(5), 780–797, September 1947.
20. R. Plomp and H. Steeneken, 'Effect of Phase on the Timbre of Complex Tones', *J. Acoust. Soc.Am.*, **2**(2), 409–421, 1969.
21. Stanley Lipshitz, Mark Pocock and John Vanderkooy, 'On the Audibility of Midrange Phase Distortion in Audio Systems', *J. Audio Eng. Soc.*, **30**(9), 580–595, September 1982.
22. Henrik Moller, Pauli Minnaar, Soren Krarup Olesen, Flemming Christensen, and Jan Plogsties, 'On the Audibility of All-Pass Phase in Electroacoustical Transfer Functions', *J. Audio Eng. Soc.*, **55**(3), 115–134, March 2007.
23. Mikko-Ville Laitinen, Sascha Disch, and Ville Pulkki, 'Sensitivity of Human Hearing to Changes in Phase Spectrum', *J. Audio Eng. Soc.*, **61**(11), 860–877, November 2013.

24. Philip Newell, Keith Holland and Peter Mapp. 'The Perception of the Reception of a Deception', *Proceedings of the Institute of Acoustics*, **24**(8), 2002.
25. C. Mathers, 'On the Design of Loudspeakers for Broadcast Monitoring', Technical Report BBC RD 1988/14, BBC Research Department, December 1988.
26. Philip Newell, Keith Holland and Julius Newell, 'The Yamaha NS10M: Twenty Years a Reference Monitor. Why?' *Proceedings of the Institute of Acoustics*, **23**(8), 29–40, 2001.
27. Philip Newell and Keith Holland, *Loudspeakers - For Music Recording and Reproduction*. Routledge, New York and London, 2nd edition, 2018.
28. Philip Newell and Keith Holland, *Loudspeakers - For Music Recording and Reproduction*. Focal Press, Amsterdam, 1st edition, 2007.
29. Sakari Tervo, Perttu Laukkanen, Jukka Patynen, and Tapio Lokki, 'Preferences of Critical Listening Environments Among Sound Engineers', *J. Audio Eng. Soc.*, **62**(5), 300–314, May 2014.
30. Bruno Fazenda, Matthew Stephenson and Andrew Goldberg, 'Perceptual Thresholds for the Effects of Room Modes as a Function of Modal Decay', *J. Acoust. Soc. Am.*, **137**(3), 1088–1098, March 2015.
31. R. King, D. Shorter, and T. Somerville, 'The Selection of a Wide-Range Loudspeaker for Monitoring Purposes', Technical Report M.008, BBC Research Department, February 1948.
32. T. Somerville, 'High Fidelity Sound Reproduction: A Brief Survey', Technical Report A.024, BBC Research Department, March 1948.
33. Keith Holland, Philip Newell and Peter Mapp, 'Modulation Depth as a Measure of Loudspeaker Low-Frequency Performance', *Proceedings of the Institute of Acoustics*, **26**(8), 2004.
34. Keith Holland, Philip Newell, Sergio Castro and Bruno Fazenda, 'Excess Phase Effects and Modulation Transfer Function Degradation in Relation to Loudspeakers and Rooms Intended for the Quality Control Monitoring of Music', *Proceedings of the Institute of Acoustics*, **27**(8), 1–8, 2005.
35. Lara Harris, *Can the Modulation Transfer Function be Used to Predict the Low Frequency Reproduction Quality of Musical Signals?* Master's thesis, ISVR, University of Southampton, MSc Dissertation, 2007.
36. Lara Harris, *Development of a Modulation Transfer Function-Based Method for Evaluating Bass Reproduction Accuracy in Professional Monitoring Loudspeakers*, Ph.D. thesis, University of Southampton, July 2015. Available online: https://eprints.soton.ac.uk/397270/.
37. L. Baker, 'Status of OTF in 1970', *Optica Acta: International Journal of Optics*, **18**(2), 81–92, February 1971.
38. Neal Viemeister, 'Temporal Modulation Transfer Functions Based Upon Modulation Thresholds', *J. Acoust. Soc. Am.*, **66**(5), 1364–1380, November 1979.
39. Ehsan Samei, Egbert Buhr, Paul Granfors, Dirk Vandenbroucke and Xiaohui Wang, 'Comparison of Edge Analysis Techniques for the Determination of the MTF of Digital Radiographic Systems', *Physics in Medicine & Biology*, **50**(15), 3613, July 2005. Available online: http://stacks.iop.org/0031-9155/50/i=15/a=009.
40. T. Houtgast and H. Steeneken, 'The Modulation Transfer Function in Room Acoustics as a Predictor of Speech Intelligibility', *Acustica*, **28**(1), 66–73, January 1973.
41. Tammo Houtgast, Herman Steeneken, Wolfgang Ahnert, Loius Braida, Rob Drullman, Joost Festen, Kenneth Jacob, Peter Mapp, Steve McManus, Karen Payton, Reiner Plomp, Jan Verhave and Sander van Wijngaarden, *Past, Present and Future of the Speech Transmission Index*, TNO Human Factors, Soesterberg, 2002.
42. 'BS EN 60268-16: Sound System Equipment: Part 16: Objective Rating of Speech Intelligibility by Speech Transmission Index', 2011.
43. J. Deer, P. Bloom and D. Preis, 'Perception of Phase Distortion in All-Pass Filters', *J. Audio Eng. Soc.*, **33**(10), 782–786, October 1985.
44. D. Preis, F. Hlawatsch, P. Bloom and J. Deer, 'Wigner Distribution Analysis of Filters with Perceptible Phase Distortion', *J. Audio Eng. Soc*, **35**(12), 1004–1012, December 1987.
45. Bruno Fazenda, Matthew Wankling, Jonathan Hargreaves, Lucy Elmer and Jonathan Hirst, 'Subjective Preference of Modal Control Methods in Listening Rooms', *J. Audio Eng. Soc.*, **60**(5), 338–348, May 2012.