24th INTERNATIONAL CONGRESS ON SOUND AND VIBRATION
23–27 July 2017, London

A COMPARISON OF THE SUBJECTIVE AND REFERENCE METHODS FOR EVALUATING THE TONAL CHARACTER CORRECTION IN BS4142:2014

Matt Torjussen

noise.co.uk Ltd, The Haybarn, Newnham Grounds, Bretford, Warwickshire, UK
email: matt@noise.co.uk

The significance of an impact at a residential receiver due to environmental sound from a commercial sound source can be affected by certain acoustic features. BS 4142:2014 allows for a character correction for these features, one of which is a correction for tonality. The standard provides three methods for determining this character correction, two of these allow graded corrections to be made; the subjective method and the reference method. The standard also makes reference to considering the uncertainty in any assessment; however, no studies have been carried out to investigate the equivalency of the subjective method and reference method in the context of the standard. A comparison of the two methods has been made by conducting a series of subjective user tests that present typical users of the standard, noise consultants and environmental health practitioners, with typical situations to assess using their professional judgement. The subjects were presented with audio recordings and their subjective assessment has been compared to the results obtained using the reference method.
Keywords: BS 4142, Tonality, Commercial, Industrial, Assessment

1. Introduction

British Standard 4142:2014 ("BS 4142") [1] provides a method for assessing the impact of industrial or commercial sound on nearby sensitive receivers. The method is based on the relative difference between the rating level, the sound of the industrial or commercial specific source plus certain acoustic feature corrections, and the background sound level in the absence of the specific sound. The greater the excess of the rating level above the background sound level, the greater the likelihood of adverse impact. This assessment is made at the location of the nearest noise sensitive receiver.

1.1 Background

BS 4142 was first published half a century ago in 1967, with four subsequent revisions culminating in the latest in 2014. The standard has evolved over the past 50 years due to the development of new ways of quantifying and measuring sound. The two most obvious examples of this are the introduction of the A-weighted equivalent continuous sound pressure level, L_{Aeq}, and the advent of the digital sound level meter.

It was with the introduction of L_{Aeq} in the 1990 revision to quantify the specific sound that tonal sound character was first considered as part of the assessment.[2] Identification of tonal character, and therefore the attraction of a 5dB character correction, relied on a subjective impression of the noise as described in BS 4142:1990:

"If the noise contains a distinguishable, discrete, continuous note (whine, hiss, screech, hum, etc,..."

It was acknowledged during a comprehensive review of this standard that the omission of an objective procedure for evaluating the tonality of industrial sound resulted in problems applying the standard.[3] This was exaggerated by the digital approach to applying the 5dB correction; misapplying a 5dB character correction could easily make a critical difference to the assessment outcome, which may have financial or social implications.

The revision of the standard in 1997[4] saw a subtle change to the treatment of tonality by referring to "acoustic features" rather than noise "characteristics", which appears to have been based on a research effort to build an acoustic feature model. The character of the noise being defined as the noise perceived by the listener, based on a combination of acoustic features.[5] It is noted that many of the recommendations about acoustic feature corrections were not absorbed into the 1997 revision. Despite the need for an objective means of assessing tonality being identified prior to the revision of the standard, a suitable means of doing so wasn't available at the time of its publication.[6]

The 2014 edition was a significant revision of the standard with greater provision for the treatment of noise "character", the word now having been reintroduced interchangeably with the term "acoustic feature". The standard now includes three means of assessing the prominence of the tonal character of the sound, with provision for a graded correction: a subjective method, an objective method and a reference method.

The inclusion of the reference method in particular addresses the implementation problems that were highlighted as early as 1990; however, the subjective method of assessment remains part of the standard. The continued presence of the subjective method has inevitably led to a slow uptake of the new reference method, a problem exacerbated by the fact that the new reference method requires narrowband analysis which is not commonly practiced by environmental noise consultants and environmental health professionals in the UK. This is an anecdotal observation made by the author, based on discussions with the subjects that participated in the study.

The choice of the reference method is based on studies carried out in the lead up to the 2014 revision that identified it as the most extensively tested method currently available.[7,8] Its inclusion in ISO 1996-2[9] is likely to have provided a level of confidence in the method, despite having not currently been adopted as a British Standard.

1.2 Tonality correction

BS 4142 acknowledges that tonality can increase the significance of an impact beyond that accounted for by the excess of the specific sound level over the background sound level. A character correction is applied to the specific sound level to obtain the rating level where tonality is present at the assessment location. The three methods for deriving this correction are:

- Subjective method
  Whether the sound is not tonal (0dB correction), contains a tone that is 'just perceptible' (+2dB correction), contains a tone that is 'clearly perceptible' (+4dB correction) or contains a tone that is 'highly perceptible' (+6dB)
- Objective method
  Tones are identified as highly perceptible if the one-third octave band containing the tone exceeds both adjacent bands by some constant level difference. The level difference is based on the frequency band containing the tone
- Reference method
  A character correction is applied on a sliding scale between 0dB and 6dB based on the second Joint Nordic Method[10] ("JNM"), which can be found in ISO 1996-2

The objective method doesn't allow for a graded character correction and doesn't offer any more detail than that provided by the reference method. Therefore, it has not been included in this study.

1.2.1 Subjective method

BS 4142 indicates that the subjective method should be used where a new specific sound source is proposed but not yet installed, thus making it difficult to make a rigorous evaluation of tonality at the assessment location using the reference method.

In the description of how to rate the tonality subjectively, the standard transposes the character corrections that would be applied under the reference method. The following tonality corrections are applied:

- +0dB where the source is not tonal
- +2dB where tonality is 'just perceptible'
- +4dB where tonality is 'clearly perceptible'
- +6dB where tonality is 'highly perceptible'

These corrections are applied based on the tonality of the sound at the assessment location. No additional definitions of 'just', 'clearly' or 'highly' are provided but there are examples provided in the standard that can help the user gauge their own assessment situation.

1.2.2 Reference method

A brief description of the reference method is provided in the following section to describe how it was implemented in the context of this study; however, it is not the intention of the author to provide a comprehensive description of the method. The reader is directed to the original technical document, which is freely available online.

The principal of the reference method is based on the audibility of tones in a critical frequency band, *CB*. This critical frequency band is centred on the tone centre frequency, f_c, and contains the energy that would mask the tone's audibility. The width of the critical band is constant up to f_c = 500Hz and then equal to $0.2f_c$ at higher frequencies.

To carry out the tone assessment, an autospectrum of the sound pressure level of the source is required. For the majority of the typical frequency range under consideration, a frequency resolution of approximately 3Hz is required. Therefore, a time-series of at least 1-minute is recommended. The autospectrum is created implementing a Hanning window, in dB referenced to 20µPa, which should be presented A-weighted. For the autospectra used in this study, a frequency resolution of 1.46Hz has been used.

An overview of the features in the autospectrum that are required is illustrated in Figure 1.

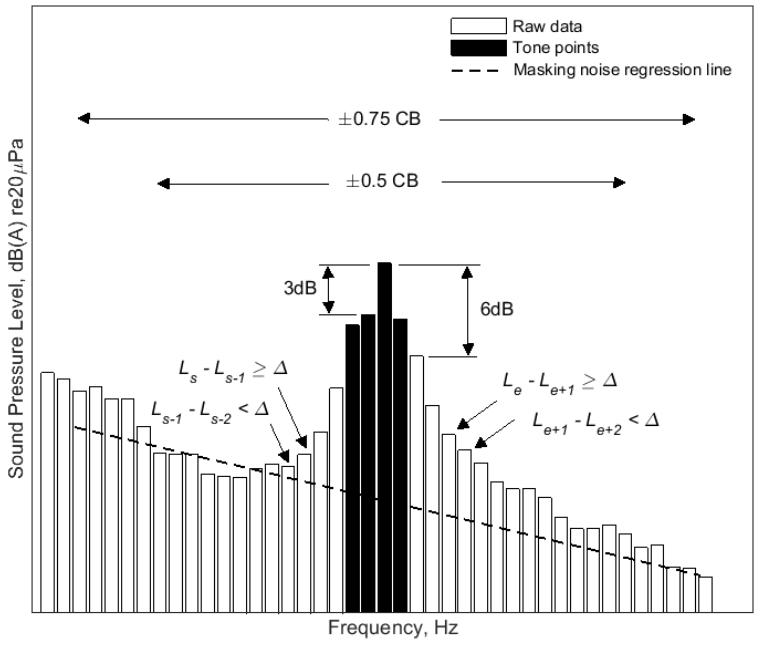

Figure 1: Graphical illustration of the important features from a sound pressure level autospectrum required to calculate the tone audibility using the reference method

The data is first sorted to identify 'noise pauses', these are points in the autospectrum that could contain a tone. The start of the noise pause is the point where the spectrum increases by a certain level, Δ, beyond the previous frequency bin. The end of the noise pause is the next point following the noise pause start where the spectrum decreases by Δ. For the data evaluated in this study Δ was set to 1dB. A tone exists in this noise pause if one of the frequency bins is 3dB higher than the adjacent ones. Where this is the case, the energy sum of the 6dB bandwidth is the tone sound pressure level, L_{pt}.

The masking noise is determined by calculating a linear regression line for the frequency bins $\pm 0.75CB$, centred on f_c, with the noise pauses in the data excluded. The masking noise sound pressure level, L_{pn}, is the energy sum of the linear regression line in the range $\pm 0.5CB$.

The tone audibility, ΔL_{ta}, is then given using

$$\Delta L_{ta} = (L_{pt} - L_{pn}) + 2 + \log_{10}\left[1 + \left(\frac{f_c}{502}\right)^{2.5}\right]$$

Equation 1

and the graded tonality correction, K_T, applied in BS 4142 is as follows:

10dB < ΔL_{ta}	→	K_t = 6dB
4dB ≤ ΔL_{ta} ≤ 10dB	→	K_t = ΔL_{ta} - 4dB
ΔL_{ta} < 4dB	→	K_t = 0dB

2. Subjective method vs. reference method

The equivalence of the subjective and reference methods implemented in BS 4142 is an important consideration. The standard is used extensively to aid planning decisions in the UK where residential developments are proposed near existing industrial or commercial premises. It is also widely used to determine the acceptability of placing new commercial or industrial activities close to existing residential receivers. The range of these activities included in the scope is varied: from a new air conditioning condenser unit to hand car washes to large manufacturing plants. The misapplication of a tonality correction can make up to 6dB difference to the assessment outcome, which could make a significant difference to the outcome of a BS 4142 type assessment. Many local environmental health departments will not recommend planning permission for a development where the rating level exceeds the background sound level, regardless of the context of the assessment. This means that making the system for applying corrections for tonality robust, repeatable and reproducible is of high importance.

2.1 Subjective tests

This study seeks to determine the extent to which the subjective and reference methods are providing the same assessment outcomes in practice. This has been done by considering one common method for evaluating specific sound subjectively, making recordings of the sound in the field and reviewing them in an office environment via headphones.

36 audio files were prepared, each having an objectively determined tonality using the reference method. 19 of the sounds were based on a real audio recording made during an environmental noise survey at the boundary between a factory and a nearby residential receiver. One of these sounds remained unaltered, with two tones that would attract tonality corrections at 443 and 1770Hz. Another recording was prepared with these tones suppressed and a further 16 with each tone presented individually with varying degrees of audibility, ensuring that the complete range of graded tonality corrections from 0dB to +6dB was covered. The remaining 17 of these sounds were entirely artificial, containing broadband noise and one of two tones at the same 443 and 1770Hz as the authentic recording. The tones were varied in amplitude to cover the range of graded corrections.

The original recording was made using a Norsonic 140 sound level meter for 1-minute and was monaural.

2.1.1 Presentation of the sounds

The subjective test was presented via a computer user interface and headphones designed to make the acquisition of the subject's responses quick. The use of the computer interface also allowed the audio files to be presented in a randomised order to each subject, which helped to mitigate some of the effects of subject training that are discussed in later sections. The audio files were presented at the same volume to each test subject.

A screenshot of the main test screen is presented in Figure 2.

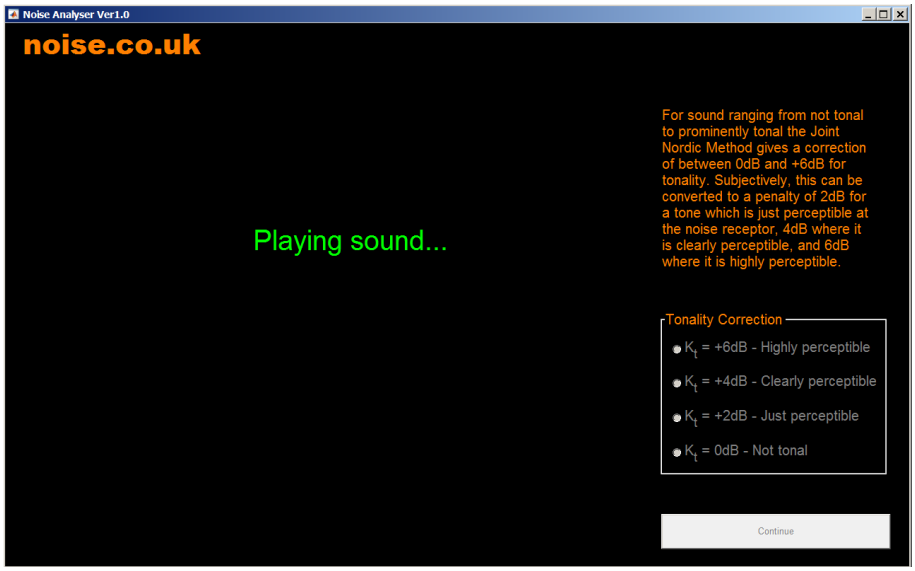

Figure 2 – Screenshot of the test screen presented to the users. On each slide the user is given the choice of the four tonality corrections from BS 4142 and asked to select which correction applies to the sound they have been presented with

Each audio file automatically started after the continue button from the previous audio file was pressed. The subject was offered the four possible character corrections for not tonal, 'just perceptible', 'clearly perceptible' and 'highly perceptible' and could not advance until one had been selected. The subject was offered the opportunity to replay the audio file as many times as they wished before selecting their chosen tonality correction. The options remained greyed out until they had listened to the 5-second sample at least once. The subjects were not able to return to samples once they had made their selection.

2.1.2 Subjects

The subjects that took part in the study were people that would normally be expected to offer their subjective opinion of tonality in the context of a BS 4142 assessment. This included 17 environmental health officers, 30 noise consultants and 10 academic staff or students.

The subjects were not screened in any other way and they were not asked about the state of their hearing. The decision not to do this was taken based on the time available to carry out the user tests but also because the subjects make these judgements about tonality in the course of their job, regardless of the state of their hearing.

2.1.3 Test environment

The tests were delivered at the places of work of the test subjects. In most instances a meeting area was provided but in two cases dedicated listening rooms were provided. In the context of carrying out a BS 4142 type assessment it would be expected that a desktop review of audio recordings would take place in a variety of environments and this variability is considered acceptable for the purposes of this study.

3. Results and discussion

3.1 Subjective vs. reference

The subjective method allows for one of four discrete tonality corrections of 0dB, +2dB, +4dB and +6dB, whereas the reference method allows for a graded correction that is permitted to be reported to one decimal place. The audio files were designed so that only integer values of the character corrections would result. However, this could still leave a discrepancy of ±1dB between the subjective and reference method results even if there were a 100% agreement between the two methods. A summary of the subjects' responses are given in Table 1.

Audio file group	No. audio files in group	Percentage response (%)				Range
		Not tonal (0dB)	Just perceptible (+2dB)	Clearly Perceptible (+4dB)	Highly perceptible (+6dB)	
No tones	1	93.0	3.5	3.5	0.0	0-4dB
0dB (not tonal)	7	52.4	34.6	11.3	1.8	0-6dB
+1dB correction	4	21.1	53.9	21.1	3.9	0-6dB
+2dB correction	4	18.9	53.1	24.6	3.5	0-6dB
+3dB correction	4	17.5	43.4	35.5	3.5	0-6dB
+4dB correction	4	12.7	43.9	36.8	6.6	0-6dB
+5dB correction	4	7.9	35.5	49.6	7.0	0-6dB
+6dB correction	9	3.5	16.4	41.5	38.6	0-6dB

Table 1 – Summary of subjective evaluation of tonality

Superficially, the results immediately illustrate that more than one subject detected tones in a recording where there were none. The 'No tones' audio file contained artificial broadband sound only but two subjects felt that tones were 'just perceptible' and a further two subjects felt that tones were 'clearly perceptible'. Furthermore, two subjects believed that audio files that were designed to be 'highly tonal' contained no tones.

A simple regression analysis has been carried out between the results of the subjective tests and the results when analysed using the reference method. The results have been plotted in Figure 3. An attempt to account for the 2dB steps in the subjective method has been made in Figure 4 by plotting the % match against tolerance in the subjective result.

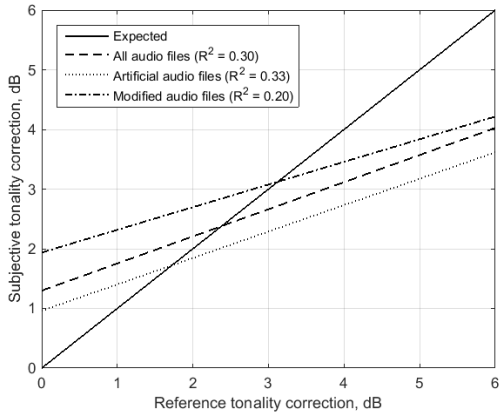

Figure 3 – Plot of the linear regression between the reference and subjective evaluation of tonality for the sample audio files

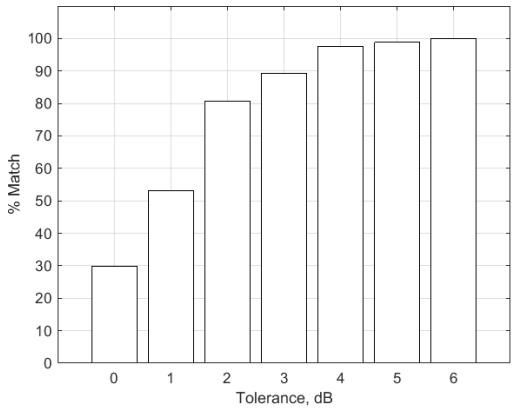

Figure 4 – Plot showing the % match of subjective responses with the reference method results sorted by tolerance

The regression analysis shows that there was a tendency amongst the population not to rate the audio files as having no tones or as the tones being 'highly perceptible', with a greater prevalence of responses being 'just perceptible' and 'highly perceptible'.

Taking into account the potential tolerance of the subjective method, approximately 30% of all responses matched that of the reference method. Given the 2dB steps in the subjective method's scale it might have been expected that a significant majority of responses would be ±1dB of the reference method's results but only 53% matched at this tolerance level. 81% matched with a tolerance of ±2dB, 89.4% matched with a tolerance of ±3dB and there was nearly 100% agreement beyond this tolerance.

The results show that generally, taking into account the entire population, the greater the tone audibility determined by the reference method the greater the tone perceived by the subjects. However, there was significant variability amongst the individual subjects. The most obvious example of this was that all of the recordings that had some tonal content were rated by one subject as 'not tonal' and another as having 'highly perceptible' tones. In practice this could result in a 6dB change in a BS 4142 assessment which could easily change the assessment outcome.

3.2 Subject feedback

The author was keen to canvas opinion about the test and each subject was asked for feedback. Many of the subjects thought that the test represented training in some way, often commenting that the first few audio files where treated differently to the remaining ones. It is unlikely that this had a significant effect on the overall population because the order that audio files were presented was randomised. However, this feedback itself indicates that the range of tones presented to the subjects informed the subjects' view of what a perceptible tone was and wasn't rather than judging them based purely on the 'just perceptible', 'clearly perceptible' and 'highly perceptible' descriptions from the standard. The question remains after these tests whether the subjects would subsequently rate any one audio file the same if it was presented in isolation, i.e. if they were carrying out a real BS 4142 assessment and didn't have a library of sounds with which to compare it to.

Audible tones were only presented at one of two frequencies and many of the subjects believed that this meant they were listening for tones in a specific range. Some more experienced listeners suggested that having a longer rest between recordings might alleviate this. This could be one reason for the number of false positives for the audio recordings that were not designed to contain perceptible tones. Subjects may have been focussing on a particular frequency range and tone that would not normally be perceptible.

Some of the subjects did not know what a tone was and understood 'tone' to be 'sound'. The data for the subjects where this was identified has been removed from the dataset but it is possible that some of this data remains and the false positives are a result of this misunderstanding. It should be stressed that all of the participants were either acousticians or environmental health professionals and that further explanation of what tonality is may need communicating within the industry.

On a more technical basis, some subjects fed back that the presentation of the audio files via headphones had its limitations. It is argued that, in the context of reviewing audio samples collected in the field, this might be how audio files are listened to in practice; the limitations of the sound reproduction are part of this and were identical between subjects. Other recommendations were made regarding the presentation of the audio files, that they should be faded in and out to avoid a sudden onset of sound and that a user changeable volume control should have been available. This is valuable feedback and will be incorporated into any future work.

3.3 Limitations of the reference method

As much as there was variability in the subjective responses, it is possible that the reference method is not adequate in quantifying the tones in the particular audio files that were presented. One of the positive features of the JNM compared to other means of evaluating tonality was its ability to account for more than one tone in a given critical band. This is done by summing the tones on

an energy basis. However, this does not account for other psychoacoustic phenomena such as the perception of difference tones or beating.[11] For the modified recordings that were presented, there were many other tones present in the audio file that were not objectively determined to be perceptible. This does not necessarily mean they weren't interfering with the perceptible tones in some other way.

4. Conclusions and recommendations

The use of the subjective method for evaluating tonality in BS 4142 has been shown to result in significant inter subject variability when used to assess 36 audio files. The audio files contained a range of perceptible tones which were objectively identified using the reference method in the same standard.

It is recommended that some further guidance be issued to help users of the standard gauge what constitutes a tone that is subjectively 'just', 'clearly' and 'highly' perceptible. It is also recommended that a greater reliance on the reference method be encouraged, even if only to help users of the standard gain experience of assessing sound subjectively.

5. Acknowledgements

The author would like to thank Mike Lotinga of WSP Acoustics for his constant encouragement and technical support, Kirsten Hinks-Knight and James Boyle for their help delivering the tests, the subjects who participated in the tests and Bill Whitfield for providing the time to carry out the study.

REFERENCES

1 British Standards Institute, BS 4142:2014 *Methods for rating and assessing industrial and commercial sound*, BSI, (2014)

2 British Standards Institute, BS 4142:1990 *Method for rating industrial noise affecting mixed residential and industrial areas*, BSI, (1990)

3 Porter, N. D., *Study of the Application of British Standard BS 4142:1990 "Method for Rating Industrial Noise Affecting Mixed Residential and Industrial Areas" (The Data Sheet Study)*, National Physical Laboratory, (1995)

4 British Standards Institute, BS 4142:1997 *Method for rating industrial noise affecting mixed residential and industrial areas*, BSI, (1997)

5 Porter, N. D., Flindell, I. H., Berry, B. F., *An acoustic feature model for the assessment of environmental noise, Institute of Acoustics Bulletin*, 18(6), (1993)

6 Porter, N. D., *The Assessment of Industrial Noise – Subjective Listening Tests and Objective Assessment Procedures"*, National Physical Laboratory, (1995)

7 Berry, B. F., Porter, N. D., National Measurement System Acoustics Programme: NMS Project 2.2: Environmental Noise, Progress Report 5, Hoare Lea Acoustics (2004)

8 Berry, B. F., Porter, N. D., A review of methods for quantifying tonal and impulsive features in environmental noise, *Proceedings of Inter-Noise 2005*, Rio de Janeiro, Brazil, 07-10 August, (2005)

9 International Standards Organisation, ISO 1996-2:2007 Acoustics – Description, measurement and assessment of environmental noise – Part2: Determination of environmental noise levels, ISO, 2007

10 Pederson, T. H., Sondergaard, M., Anderson, B., *Objective Method for Assessing the Audibility of Tones in Noise – Joint Nordic Method v2*, DELTA Technical Report AV 1952/99, (1999)

11 Fastl, H., Zwicker, E., Psychoacoustics – Facts and Models, 3rd Edition, Spriner-Verlag Berlin Heidelberg, New York, (2007)