

**The Journal of the Acoustical Society of America**  
**Subjective dominance as a basis for selecting frequency weightings**  
--Manuscript Draft--

<b>Manuscript Number:</b>	JASA-00366R1
<b>Full Title:</b>	Subjective dominance as a basis for selecting frequency weightings
<b>Short Title:</b>	Subjective dominance for selecting frequency weightings
<b>Article Type:</b>	Regular Article
<b>Corresponding Author:</b>	Antonio J Torija, Ph.D. University of Southampton Southampton, Hampshire UNITED KINGDOM
<b>First Author:</b>	Antonio J Torija, Ph.D.
<b>Order of Authors:</b>	Antonio J Torija, Ph.D. Ilan H Flindell, Ph.D. Rod H Self, Ph.D.
<b>Section/Category:</b>	Noise
<b>Keywords:</b>	Frequency Weightings; Noise Annoyance; Subjective Acoustics; Road Traffic Noise; Sound Spectra.
<b>Abstract:</b>	<p>The objective of this paper is to propose and illustrate a new and simple approach for the selection of frequency weightings for the assessment of environmental and transportation noise. In recent years, the A-frequency weighting has become almost universal except where existing standards and regulations mandate the use of alternative weightings and/or frequency summation procedures, but even where this has been based on extensive research, no real consensus has been achieved. The new approach is based on the concept of subjective dominance, which does not always conform to the physically dominant frequencies identified by the A- or other frequency weightings and summation procedures used in measurements and/or predictions. The new approach is illustrated by the results of a limited series of five listening tests which clearly demonstrate that no single objective frequency weighting or summation procedure is capable of providing the best-fit to subjective responses across a range of different contexts. Subjective dominance varies across different listening contexts and situations, and should therefore be considered whenever noise management and control decisions are being made. The new approach will naturally require further research because of the wide range of different contexts and situations in which it might need to be applied.</p>

# Subjective dominance as a basis for selecting frequency weightings

Antonio J. Torija<sup>a)</sup>, Ian H. Flindell and Rod H. Self

*Institute of Sound and Vibration Research, University of Southampton, Highfield, Southampton  
SO17 1BJ, UK*

<sup>a)</sup>Author to whom correspondence should be addressed.

E-mail address: [ajtorija@ugr.es](mailto:ajtorija@ugr.es) (A.J. Torija)

## **Abstract**

The objective of this paper is to propose and illustrate a new and simple approach for the selection of frequency weightings for the assessment of environmental and transportation noise. In recent years, the A-frequency weighting has become almost universal except where existing standards and regulations mandate the use of alternative weightings and/or frequency summation procedures, but even where this has been based on extensive research, no real consensus has been achieved. The new approach is based on the concept of subjective dominance, which does not always conform to the physically dominant frequencies identified by the A- or other frequency weightings and summation procedures used in measurements and/or predictions. The new approach is illustrated by the results of a limited series of five listening tests which clearly demonstrate that no single objective frequency weighting or summation procedure is capable of providing the best-fit to subjective responses across a range of different contexts. Subjective dominance varies across different listening contexts and situations, and should therefore be considered whenever noise management and control decisions are being made. The new approach will naturally require further research because of the wide range of different contexts and situations in which it might need to be applied.

PACS number(s): 43.50.Qp, 43.50.Ba, 43.50.Rq, 43.50.Lj

## I. INTRODUCTION

The optimization of acoustic metrics for predicting transportation noise impacts is an enduring research theme (Fidell *et al.*, 2011) where many laboratory studies on transportation noise annoyance have highlighted the non-negligible effects of spectral features (Hellman, 1982; Kjellbert *et al.*, 1997; Leventhall, 2003; Leventhall, 2004; Nilsson, 2007).

As a standardised method for taking different spectral features into account, the A-frequency weighting has received considerable attention in the literature (Meloni and Rosenheck, 1995), and has been adopted in many standards and regulations (Schomer *et al.*, 2001), but it has not been universally accepted. The classical justification for the A-frequency weighting is that it approximates to an inverse of the standard 40 phon equal-loudness contour (ISO, 2003a). It has therefore been criticised because it does not change with sound level (Schomer *et al.*, 2001), whereas the standardised equal-loudness contours (and the B- and C-frequency weightings) flatten out at higher sound-levels. There is experimental data that suggests that the A-frequency weighting overcompensates for the hearing system's reduced sensitivity at low frequencies (Schomer, 2004; Nilsson, 2007). Laboratory studies have shown that the A-frequency weighting can, under some circumstances, underestimate annoyance responses to noise with dominant low frequency components (Broner and Leventhall, 1980; Kjellberg and Goldstein, 1985; Persson and Björkman, 1988; Berglund *et al.*, 1996; Kjellberg *et al.*, 1997; Yifan *et al.*, 2008). Kjellberg and Goldstein (1985) found that C- and D-frequency weightings can both have higher correlations with noise annoyance when there are significant low frequency components present. On the other hand, Alayrac *et al.* (2010) (and many others) have shown the A-frequency weighted sound level to be perfectly adequate for assessing reported annoyance attributable to their particular sets of low

frequency sounds. Considered overall, the various conflicting reports suggest that the administratively convenient and widespread adoption of the A-frequency weighting in many national and international standards is not as scientifically justifiable as many users would like.

Several studies reported in the literature found that the loudness and annoyance of sounds with similar A-frequency weighted sound-levels increased with increasing (relatively) low frequency content (Watts, 1995; Leventhall, 2004; Nilsson, 2007). However, other studies found that any high frequency components present can have a greater effect on annoyance (Versfeld and Vos, 1997; Kim *et al.*, 2009). Versfeld *et al.* (1994) found that the relative difference in level between the high frequency and low frequency parts of the spectrum can have a significant effect on annoyance in addition to the effect of overall sound level measured in absolute terms.

To overcome the deficiencies of fixed frequency weighting filters that do not change with sound level, Zwicker proposed a level dependent loudness-level calculation method (Fastl and Zwicker, 1990). Zwicker's method applies a level dependent weighting to the measured level in each 1/3 octave band with adjustments for spectral masking and other features of the sound (Wang *et al.*, 2007; Lee, 2008). The frequency selectivity of Zwicker's method conforms to critical band theory which was and still is an important concept in understanding personal hearing sensation (Kook *et al.*, 2012). Steven's loudness-level (Stevens, 1956) and Kryter's Effective Perceived Noise Level (Kryter, 1968) used for aircraft noise certification follow similar principles, although the level dependent weightings and other adjustments are marginally different. All three loudness-level based methods give marginally different results, as do a number of other but broadly similar methods proposed by different authors (Blommer *et al.*, 1996; Moore and Glasberg, 1996; Moore *et al.*, 1997), which suggests that whereas one method may work 'best' under one set of circumstances, different methods often work 'better' under different circumstances.

It should also be noted that while objective measurements of transportation noise sound level using any of the different standardised procedures can be important for regulatory and contractual purposes they are not the same thing as subjective loudness which can only be determined by listening tests. For this reason, calculated or measured 'loudness-level' is not the only variable of interest for deciding between alternative options for transport infrastructure development and mitigation.

In Torija and Flindell (2014a), the effect on reported loudness and reported annoyance of altering the frequency spectrum of recorded (outdoors) urban main road-traffic sounds was analysed. In this paper, Torija and Flindell (2014a) found that the physically dominant part of the frequency spectrum is not necessarily a good guide to which part of the spectrum will be perceived as subjectively dominant. Torija and Flindell (2015) next investigated the relationship between reported loudness and reported annoyance under conditions where the low frequency content is relatively more (physically) dominant, such as in indoor conditions. In this paper, Torija and Flindell (2015) found that changes in low frequency content appeared to make smaller contributions to subjective loudness and annoyance than might be inferred from the implied objective or physical dominance of those changes. Consequently, under the indoor conditions tested, the A-frequency weighting outperformed the C-frequency weighting in the sense that higher correlations with reported annoyance were obtained. In this paper, we present further results from the above mentioned listening tests, hereinafter called study 1 (Torija and Flindell, 2014a) and study 2 (Torija and Flindell, 2015), together with the results from three further listening tests to illustrate the new approach based on subjective rather than physical dominance, i.e. selecting the frequency weighting for assessing noise annoyance in any specific context on the basis of which part of the frequency spectrum is subjectively dominant. This proposed approach might help to

explain why different frequency weighting curves and loudness-level calculation procedures appear to work 'best' in different situations. It could also inform the selection of frequency weightings for regulatory and contractual purposes wherever it is desired to take into account differences in context and subjective expectations which are often ignored by existing procedures.

The different contexts illustrated in this paper were; road-traffic-noise as heard outdoors in an open public space (study 1- Torija and Flindell, 2014a), road-traffic-noise as heard indoors (study 2 - Torija and Flindell, 2015), and road-traffic-noise as heard behind a noise barrier (study 3) with additional increases and decreases in low frequency (LF), mid-frequency (MF), and high frequency (HF) content to test the range of acoustic metrics applied. These contexts were selected on the basis of their relevance for policy and decision-making purposes in road-traffic-noise impact assessment. Two further listening tests were then carried out (studies 4 and 5) using basically similar road-traffic sounds but different comparison procedures to investigate further context dependencies. It should be noted that the separate listening tests (studies 1 to 5) are reported sequentially in this paper because the detailed design and experimental procedure for each study progressed from each study to the next. Clearly, there are many other contexts and situations where the new approach might need to be tested to investigate wider applicability.

## **II. METHODOLOGY**

### **A. Assumptions about the experimental design**

As mentioned above, the results of five listening tests (studies 1 to 5) are presented sequentially to illustrate the development of a new approach for the selection of frequency weightings to be used in environmental noise impact assessments. Because of the need to focus

listener's attention onto the key variables of primary research interest, the master recordings were all of typical continuous urban outdoor road-traffic-noise with no significant fluctuations in instantaneous sound level or other prominent or distracting features. The recordings were all made in the city of Granada (Spain). For study 4, where comparisons were made against broadband pink noise reference sounds, the master road-traffic-noise recording was additionally selected for a reasonable subjective balance between the different (LF, MF and HF) frequency ranges. The sound-levels of the master recordings used in each listening experiment were set at typical values in the corresponding context.

For these listening tests, the audio spectrum was arbitrarily divided into three frequency regions, i.e. LF (20-250 1/3-octave bands), MF (315-2000 1/3-octave bands) and HF (2500-20000 1/3-octave bands), but as based on custom and practice widely reported in the relevant literature (e.g. Cowan, 2016). Because of different traffic dynamics and different types of road vehicles in traffic, significant variations can be observed in the relative LF, MF and HF content of urban road-traffic-noise. Also, different building envelopes or noise barriers can achieve significantly different performances in the attenuation of LF, MF and HF. In listening experiments 1, 2 and 3, the performance of standard frequency weightings is assessed for accounting for the variations in LF, MF and HF content, in the three contexts considered. In listening experiments 1 and 2, the variations in frequency content were simulated using low and high pass shelf filters, and a band pass/band stop filter in experiment 1, to cut or boost each frequency range to better represent variation which occurs in practice. To each filtered sound a relative gain setting from -9dB to +9dB was applied. These cut and boost settings were adopted to include and probably exceed typical variation in relative frequency content of urban road-traffic-noise. In listening experiment



3, frequency filters were applied to simulate the insertion loss (IL) of a set of 5 different noise barriers as based on data available in the literature (Hong *et al.*, 2012).

It should be noted that a prior motivation for listening test 1 (outdoor context) was a comparison against a field listening test carried out in Lyon, France (Torija and Flindell, 2014b) and this determined the use of a five point semantic scale of reported annoyance. After studying the feedback provided by the participants of listening test 1, it was decided to adopt the relative magnitude estimation (RME) method for subsequent listening tests. The RME method offers a series of benefits for the participants (see ISO, 1999), and encourages participants to respond on a ratio scale of measurement which then facilitates mathematical operations and parametric statistical analysis (Turpin *et al.*, 2015). Notwithstanding any differences in statistical power, it is not expected that the subjective scales used would have had any significant effect on the relativity of the results observed in each listening test considered separately.

In Torija and Flindell (2015), it was found little inter-participant variability. In all listening tests reported herein, all the participants took part in the experiments under exactly the same conditions, i.e. the same sound-proofed audiometric test room, the same experimental set-up, the same pre-experiment information, and the same stimuli. Therefore, any observed inter-participant variability may be assumed to be independent of the experimental conditions. On the other hand, environmental noise impact assessments are conducted for communities and not for individuals. For these reasons, all regression analyses between reported annoyance and frequency-weighted sound-levels were performed on the aggregated (mean) scores and not on the individual responses.

Finally, in all the listening experiments, the stimuli were randomly presented in order to avoid any order-effect on the participants' responses.

## **B. Experimental set-up**

All the listening laboratory studies described in this paper were conducted in a sound-proofed audiometric test room at the University of Southampton. Participants were recruited by advertising within the department and around the University. On arrival at the laboratory for pre-booked appointments the aims and procedures were carefully explained and voluntary consent confirmed by signing a standard consent form. Subjective impressions were recorded on standardised questionnaire forms. All audio signals (.wav files) were reproduced via a high quality sound card, and then sent to two high resolution active monitor loudspeakers via a small audio mixing console. The two loudspeakers were positioned in front of and to either side of the seated listener(s) to obtain an even sound level distribution and then calibrated/adjusted to obtain the sound-levels noted in the following sections of this paper using a Bruel & Kjaer calibrator type 4230 and a Norsonic Environmental Noise Analyser type 121, with a Norsonic free-field microphone type 1225 positioned at the listener's head position with the listener not present. Table I provides a summary of the 5 listening tests reported in this paper.

## **C. Listening test 1: urban road-traffic-noise under outdoors conditions**

Part of this study was previously reported in Torija and Flindell (2014a). All test sounds were based on a 15 second master recording of typical roadside continuous low speed busy urban road-traffic-noise ( $L_{Aeq,15\text{ sec}} = 70.3$ ). The recording comprised continuous heavy road-traffic background noise with four separately discernible individual road vehicle pass-bys within the overall 15 seconds duration but with variation in instantaneous sound level (S-time weighting)

within plus and minus 2-3 dB. The 15 second averaged frequency spectrum of the master recording is shown at Fig. 1. Three frequency filters, i.e. low shelf filter, band pass/band stop filter and high shelf filter, were applied respectively for boosting or cutting the LF, MF and HF ranges. Twelve filtered sounds were produced by applying each filter with -9dB, -3 dB, +3dB and +9 dB relative gain setting. Thirty volunteer listeners (18-65 yr of age, 14 males and 16 females) were asked to judge each sound as if it was being heard in an outdoor public space nearby to a main road using a questionnaire based on the standardized ISO/TS 15666 specification (ISO, 2003b) for noise annoyance questionnaires (five point semantic scale: “Not at all”, “Slightly”, “Moderately”, “Very”, and “Extremely”).

#### **D. Listening test 2: urban road-traffic-noise under indoor conditions**

Torija and Flindell (2015) used a 12.5 seconds master recording of typical roadside continuous low speed busy urban road-traffic-noise. The recording comprised continuous heavy road-traffic background noise with subjectively identifiable but not particularly prominent individual road vehicle pass-by events. The average sound level at the recording position was 68.9  $L_{Aeq,12.5 \text{ sec}}$  with variation in instantaneous sound level (S-time weighting) within plus or minus 2-3 dB. The master recording was filtered to simulate typical frequency dependent attenuation of double glazing sealed units made up from 3mm glass, 3 mm air gap, and 3 mm glass, according to the values reported by Quirt (1983). Also, artificial reverberation at 0.5 second reverberation time was added to increase the subjective realism of the intended indoor simulation. This laboratory study was divided into two sub-tests with reference indoor filtered sounds reproduced at 49.5  $L_{Aeq,12.5 \text{ sec}}$  and 39.5  $L_{Aeq,12.5 \text{ sec}}$  as shown in Fig. 2. Two frequency filters, i.e. low and high

frequency shelf filters were respectively applied for boosting or cutting the LF, and MHF ranges. These filters were applied separately at -9 dB, -3 dB, +3 dB, and +9 dB to the two simulated indoor filtered sounds, thereby producing  $8 \times 2 = 16$  different test sounds. Thirty three volunteer listeners (18-65 yr of age, 18 males and 15 females) were told that they should judge the sounds as if they were indoors and the road-traffic sound was coming indoors from a nearby road. Then, the listeners assessed the relative annoyance of each test sound by using the RME method (Huang and Griffin, 2014), against an arbitrary rating of 100 for each reference sound.

### **E. Listening test 3: urban road-traffic-noise with noise barriers**

Two 12.5 seconds master recordings of typical roadside continuous low speed busy urban road-traffic-noise were used in this laboratory study. The effect of LF content in the master recording on the reported annoyance of the noise barrier filtered sounds was investigated. The first master recording, hereafter called RT1, had relatively high LF content in comparison to the MF and HF content. The second master recording, hereafter called RT2, had lower LF content compared to the MF and HF content. Both master recordings had limited variation in instantaneous sound level (S-time weighting) within 2-3dB, with 3-4 individual vehicle pass-by events being subjectively identifiable but not particularly prominent. The time average sound-levels at the recording positions were  $67.1 L_{Aeq,12.5 \text{ sec}}$  (RT1) and  $71.4 L_{Aeq,12.5 \text{ sec}}$  (RT2).

The IL of 5 different noise barriers made of different materials was simulated by applying frequency filters derived from field measurements of actual noise barriers, (see Hong *et al.*, 2012),

leading to 10 different test sounds, which were all repeated with an additional 6 dB broadband attenuation making 20 test sounds in all (shown in Fig. 3).

Thirty volunteer listeners (18-65 yr of age, 19 males and 11 females) judged the relative annoyance of each test sound against the two reference sounds (RT1 and RT2) using the same RME method as in listening test 2 described above.

#### **F. Listening test 4: LF, MF and HF road-traffic-noise vs. pink noise**

Listening test 4 used an 8 second roadside master recording of a single vehicle pass by (urban bus) ( $67.6 L_{Aeq,8\text{ sec}} - S$ -time weighting), with subjectively well balanced contributions from mainly engine noise in the low frequency region (LF), from mainly rolling noise in the mid-frequency region (MF), and from mainly gas exhaust noise in the high frequency region (HF). Test sounds were produced by applying LF, MF, and HF frequency filters to the master recording (Fig. 4.A), as listed below.

(i) Low-frequency filtered road vehicle (LF\_RT\_exp1): using a high shelf filter for cutting frequencies above a cut-off frequency ( $f_c$ ) of 250 Hz.

(ii) Mid-frequency filtered road vehicle (MF\_RT\_exp1): using both low and high shelf filters for cutting frequencies below  $f_c = 315$  Hz and above  $f_c = 2000$  Hz respectively.

(iii) High-frequency filtered road vehicle (HF\_RT\_exp1): using a low shelf filter for cutting frequencies below  $f_c = 2500$  Hz.

Thirty-three volunteer listeners (18-65 yr of age, 18 males and 15 females) compared the relative annoyance of each filtered road vehicle sound against a 70 dB(Z) pink noise reference sound using a simple pair-comparison paradigm (less, equal, or more annoying) with the filtered road vehicle sound-level titrated up and down in 5 dB steps to find at which sound level the road vehicle sound was judged subjectively equivalent. The starting off points for each test sound were set at the equivalent dB(Z) sound-levels of the wide-band pink noise reference sound as if the pink noise had been filtered through the same filters as were used to prepare the LF, MF, and HF test sounds as described above (see Table II and Fig. 4.B). This was done to ensure that the starting-off point for each test sound was approximately equivalent in terms of physical sound energy, even though, due to the up-down experimental procedure, it should not (in theory) have made any difference to the final results. Table II also shows that, in terms of physical sound energy alone (of wide-band pink noise - which is generally assumed to be the closest to a generic 'flat' frequency spectrum in terms of auditory sensation), the three filter bands were approximately equivalent.

**G. Listening test 5: LF, MF and HF road-traffic-noise vs. LF, MF and HF filtered pink noise**

Listening test 5 used a 10 seconds master recording of typical roadside continuous low speed busy urban road-traffic-noise with subjectively audible but not prominent individual road vehicle pass-by events. The average sound-level at the recording position was 67.2  $L_{Aeq,10 \text{ sec}}$  with variation in instantaneous sound-level (S-time weighting) within plus or minus 2-3 dB. Test sounds were produced by applying frequency filters to the master recording (Fig. 5.A) and to wide

band pink noise, following the same generic procedures as for listening test 4 described above, with the exception that the master recording was continuous road-traffic in test 5 and a single vehicle pass-by in test 4. In test 5, the listeners judged the relative annoyance of each filtered road vehicle sound against every other filtered pink noise reference sound. The filtered pink noise reference sounds (Fig. 5.C) were all pre-set at 70 dB(Z) to ensure consistency in terms of physical sound-levels with the 70 dB(Z) wide band pink noise reference sound used in test 4, and for test 5, the same 70 dB(Z) starting off point was used for all the filtered test sounds (see Table II and Fig. 5.B). In all other respects the same simple pair-comparison paradigm (less, equal, or more annoying) was used in test 5 as was used in test 4. The same thirty-three volunteer listeners (18-65 yr of age, 18 males and 15 females) took part in test 5 as in test 4 but on different days.

### **III. RESULTS**

#### **A. Listening test 1 – urban road-traffic-noise as heard outdoors (Torija and Flindell, 2014a)**

Fig. 6 shows the linear relationship between the different frequency weightings and the reported annoyance for the urban road-traffic filtered stimuli. Square, asterisk and triangle symbols correspond to LF, MF and HF filter gain respectively. The highest correlations with reported annoyance were found for D-weighted sound-levels (Fig. 6.D) and Zwicker's loudness-level (Fig. 6.E) (see also Table III). More detailed examination of the data shown in Fig. 6 suggested that the Z-frequency weighting (effectively unweighted within the audio frequency

range) and C-frequency weighting (almost unweighted within the audio frequency range) over-estimated the contribution made by LF content to reported annoyance, while the A-frequency weighting underestimated the contributions made by both LF and HF content to reported annoyance.

#### **B. Listening test 2 – urban road-traffic-noise as heard indoors (Torija and Flindell, 2015)**

In this listening test, Torija and Flindell (2015) found that even under conditions where LF content was physically dominant (indoor conditions), changes in LF content made smaller contributions to reported annoyance than might be inferred from such physical dominance. Fig. 7 and Fig. 8 show the regression analyses between reported annoyance and the different frequency weightings for the 49.5  $L_{Aeq}$  indoor filtered and 39.5  $L_{Aeq}$  indoor filtered stimuli respectively. Square and asterisk symbols correspond to LF and MHF filter gain respectively. For indoor sounds, the highest correlations with reported annoyance were found for A-frequency weighted sound-levels and Zwicker's loudness-level, with the C-, D- and Z-frequency weightings performing worst (see Table IV). More detailed examination of the data in Fig. 7 and Fig. 8 suggested that in this case, the extent to which the A-frequency weighting and Zwicker's loudness-level downweighted the relative importance of the physically dominant LF content to reported annoyance was about right, whereas the Z-, C- and D-frequency weightings over-estimated the relative contribution of the physically dominant LF content to reported annoyance.



### **C. Listening test 3 – effect of roadside noise barriers**

As can be seen in Table V and Fig. 9, notwithstanding the relatively high LF content in filtered road-traffic-noise simulating the effects of different roadside noise barriers (see Fig. 3), the A- and D-frequency weightings and Zwicker's loudness-levels show higher correlations with reported annoyance than the C- and Z-frequency weightings. In more detail, for master recording RT1, with relatively more LF content than master recording RT2, higher correlations are observed between reported annoyance and D-frequency weighted sound-levels and Zwicker's loudness-level than between reported annoyance and A-frequency weighted sound-levels. These results appear to be generally consistent with the results obtained in listening tests 1 and 2, i.e. that for continuous urban road-traffic-noise, the LF region is not subjectively as important for reported annoyance as the MF and HF regions unless it has been significantly relatively enhanced through being heard indoors or on the shielded side of a roadside noise barrier, when it can become subjectively more important. Moreover, the difference in the relative contribution to subjective annoyance made by the LF and MHF regions seems to be important for the selection of frequency weightings. Thus, under the range of test conditions investigated, the sound-levels weighted with the frequency filters that better accounted for the subjective contribution of each frequency region achieved the best correlation to reported annoyance.

### **D. Listening tests 4 and 5 – subjective equivalence between LF, MF and HF frequency regions**

In listening tests 4 and 5, LF, MF and HF filtered road-traffic test sounds were compared against broadband and similarly filtered pink noise reference sounds to determine points of subjective equivalence. It should be noted that the relative performance of Zwicker loudness-level was not investigated in tests 4 and 5 because the output of Zwicker's loudness-level calculation procedure is expressed in 'sone' units. Sone units are not directly comparable against dB units (i.e. the A-, C-, D-, and Z-frequency weightings) except by making assumptions which, for the type of experimental design adopted herein, would have invalidated any direct comparison of the relative performance of the two types of metric. Thus, the analysis of tests 4 and 5 only considered the relative performance of the A-, C-, D- and Z-frequency weightings.

The mean differences in dB units for each frequency filter from the starting point of physical equivalence against broadband and equivalently filtered pink noise reported by the participants are shown in Table VI (listening test 4) and Table VII (listening test 5). For instance, the average difference in sound level found between LF and MF road-traffic test sounds to be reported as equally annoying is 10.9 dB in listening test 4 and 11.3 dB in listening test 5 when measured unweighted (Z-frequency weighting), but only -1.0 dB and -0.1 dB respectively when measured using the A-frequency weighting. This finding suggests that for these particular sounds, the A-frequency weighting gives a better description of the observed differences in reported annoyance between LF and MF content. The data shown in Tables VI and VII suggest that, among the evaluated frequency filters, the A-frequency weighting best describes the differences in reported annoyance between LF and MF road-traffic test sounds, while the D-frequency weighting best describes the differences in reported annoyance between HF road-traffic test sounds and both LF and MF road-traffic test sounds. These results are consistent with the findings of listening tests 1

and 3 (outdoors), where D-weighted sound-levels had shown the highest correlation with reported annoyance.

On the other hand, Tables VI and VII also show subjective differences between the different frequency weightings for listening tests 4 and 5. In listening test 4 the broadband pink noise reference sound level was set at 70 dB(Z), but the starting points for each of LF, MF, and HF filtered road-traffic test sounds were set at the corresponding sound-levels to which the same filters applied to the broadband pink noise would have produced (as shown in Table II). In listening test 5, the overall sound-level of each filtered pink noise reference sounds was pre-set at 70 dB(Z), and the same 70 dB(Z) starting off point was used for all the filtered test sounds. For instance, a difference in physical magnitude between LF and HF content is observed in Table II for listening test 4, while no difference is observed for listening test 5. These results demonstrate that subjective differences do not always correspond to physical differences measured objectively using standard acoustic metrics. Between the single vehicle pass-by test sound (test 4) and the continuous road-traffic test sound (test 5), there were no substantive differences in the pattern of results, i.e. the A-frequency weighting appeared to provide the best description of subjective equivalence between LF and MF, and the D-frequency weighting appeared to provide the best description of subjective equivalence between MF and HF and LF and HF.

#### **IV. DISCUSSION**

The four listening tests dealing with urban road-traffic-noise heard outdoors (1, 3, 4 and 5) pointed towards the HF and to a lesser extent the MF frequency ranges as being, for these types of

sound, subjectively dominant. This finding is consistent with previous research suggesting that the HF frequency range generally makes the greatest contribution to reported annoyance attributed to road-traffic-noise (Versfeld *et al.*, 1994; Versfeld and Vos, 1997; Kim *et al.*, 2009). Zwicker's loudness-level, and both the A- and D-frequency weightings achieved the highest correlations with reported annoyance under outdoors conditions. However, and notwithstanding the subjective dominance of the MF and HF frequency ranges under the general case, it appears that the LF region should also be taken into account whenever present in sufficient amounts (Watts, 1995; Leventhall, 2004; Nilsson, 2007). Thus, whenever there is a strong LF presence (see listening test 3) Zwicker's loudness-level and the D-frequency weighting achieved better performance in assessing reported annoyance than the A-frequency weighting. Moreover, in listening tests 4 and 5 it was found that subjective differences between the different frequency ranges could be considerably affected by physical frequency spectrum differences between the different sounds tested.

In general and under typical indoor listening conditions, the MF and HF regions are significantly attenuated by the building envelope, and the LF region becomes relatively more dominant. However, this physical dominance does not appear to lead to subjective dominance, at least not without considerable MF and HF attenuation. The data from listening test 2 suggest that under indoor conditions where the LF frequency region approaches subjective equivalence to the more usually dominant MF and HF frequency regions, the A-frequency weighting becomes more appropriate, and the C-, D-, and Z-frequency weightings become less so. Zwicker loudness-level appears to work better at the higher than at the lower overall sound level tested.

In light of these results, it seems quite unlikely that a unique frequency weighting filter could be defined for assessing the reported annoyance associated with different transportation noise sources under the wide range of different contexts and situations that can occur in practice. Instead,

it would seem more appropriate to select frequency weightings according to whichever frequency region is subjectively dominant in any particular case, and thereby encouraging noise control action (if any such is justified) aimed at the frequency region of greatest subjective importance. Consideration should also be given to taking the changing contribution to subjective annoyance made by each frequency region under different loudness regimes into account. Of course, any study of this kind has limitations when extrapolating to real-life conditions, not the least being the unavoidably restricted range of road-traffic conditions covered by the master audio recordings, and the difficulty of obtaining more than a relatively small number of loudness and annoyance comparisons from each volunteer listener without exhausting either their patience or goodwill. Further work including a much wider range of road-traffic conditions would be required in order to justify any changes to existing standards and regulations, and the direct relevance to real life of listening tests carried out under laboratory conditions would also need to be established.

## **V. CONCLUSIONS**

This paper presents the results of a series of 5 listening tests aimed at investigating the main reasons why different frequency weighting and loudness-level calculation schemes appear to perform differently under different sets of input conditions. The results presented in this paper point towards the suggestion that the optimum frequency weighting for use in any particular context may depend to some considerable extent on whichever frequency region is subjectively dominant, possibly taking into account both relative loudness and the relative contribution towards reported annoyance.

Of course, before any recommendations can be made with respect to existing standards and regulations, further research work covering a much wider range of different contexts and conditions would be required.

## **ACKNOWLEDGMENTS**

This work is funded by the University of Malaga and the European Commission under the Agreement Grant no. 246550 of the seventh Framework Programme for R & D of the EU, granted within the People Programme, «Co-funding of Regional, National and International Programmes» (COFUND), and Ministerio de Economía y Competitividad (COFUND2013-40259).

Alayrac, M., Marquis-Favre, C., Viollon, S., Morel, J., and Le Nost, G. (2010). "Annoyance from industrial noise: Indicators for a wide variety of industrial sources," *J. Acoust. Soc. Am.* 128(3), 1128-1139.

Berglund, B., Hassmén, P., and Job, R. F. S. (1996). "Sources and effects of low-frequency noise," *J. Acoust. Soc. Am.* 99(5), 2985-3002.

Blommer, M., Otto, N., Wakefield, G., Feng, B. J., and Jones, C. (1996). "Calculating the loudness of impulsive sounds," *SAE Transactions* 104(6), 2302-2308.

Broner, N., and Leventhall, H. G. (1980). "A modified PNdB for assessment of low frequency noise," *J. Sound Vib.* 73(2), 271-277.

Cowan, J. P. (2016). *The effects of sound on people*, first ed. (Wiley, Chichester), pp. 81-85.

Fastl, H., and Zwicker, E. (1990). *Psychoacoustics: Facts and Models*, second ed. (Springer, Heidelberg), pp. 203-238.

Fidell, S., Mestre, V., Schomer, P., Berry, B., Gjestland, T., Vallet, M., and Reid, T. (2011). "A first-principles model for estimating the prevalence of annoyance with aircraft noise exposure," *J. Acoust. Soc. Am.* 130(2), 791-806.

Hellman, R. P. (1982). "Loudness, annoyance, and noisiness produced by single-tone-noise," *J. Acoust. Soc. Am.* 72(1), 62-73.

Hong, J. Y., Jang, H. S., and Jeon, J. Y. (2012). "Evaluation of noise barriers for soundscape perception through laboratory experiments," in *Proceedings of the Acoustics 2012 Conference*, Nantes, France, pp. 2159-2162.

Huang, Y., Griffin, M. J. (2014). "Comparison of absolute magnitude estimation and relative magnitude estimation for judging the subjective intensity of noise and vibration," *Appl. Acoust.* 77, 82-88.



ISO (1999). Sensory analysis – Methodology – Magnitude estimation method. ISO 11056:1999, International Organization for Standardization (ISO), Geneva, Switzerland.

ISO (2003a). Normal equal-loudness level contours. ISO 226:2003 Acoustics, International Organization for Standardization (ISO) 2nd edition, Geneva, Switzerland.

ISO (2003b). Acoustics-Assessment of noise annoyance by means of social and socio-acoustic surveys. ISO/TS15666:2003(E), International Organization for Standardization (ISO), Geneva, Switzerland.

Kim, J., Lim, C., Hong, J., and Lee, S. (2009). “Noise-induced annoyance from transportation noise: Short-term responses to a single noise source in a laboratory,” *J. Acoust. Soc. Am.* 127(2), 804-814.

Kjellbert, A., and Goldstein, M. (1985). “Loudness assessment of band noise of varying bandwidth and spectral shape,” *J. Low Freq. Noise Vib.* 4, 12-26.

Kjellberg, A., Tesarz, M., Holmberg, K., and Landström, U. (1997). “Evaluation of frequency-weighted sound level measurements for prediction of low-frequency noise annoyance,” *Environ. Int.* 23(4), 519-527.

Kook, J., Koo, K., Hyun, J., Jensen, J. S., and Wang, S. (2012). “Acoustical topology optimization for Zwicker’s loudness model – Application to noise barriers,” *Comput. Method Appl. M.* 237-240, 130-151.

Kryter, K. D. (1968). “Concepts of perceived noisiness, their implementation and application,” *J. Acoust. Soc. Am.* 43(2), 344-361.

Lee, S. –K. (2008). “Objective evaluation of interior sound quality in passenger cars during acceleration,” *J. Sound Vib.* 310(1-2), 149-168.

Leventhall, G. (2003). A review of published research on low frequency noise and its effect. Report for Department for Environment, Food and Rural Affairs (DEFRA), pp. 1-88. Available at [http://westminsterresearch.wmin.ac.uk/4141/1/Benton\\_2003.pdf](http://westminsterresearch.wmin.ac.uk/4141/1/Benton_2003.pdf). (date last viewed 21/01/16).

Leventhall, H. G. (2004). "Low frequency noise and annoyance," *Noise Health* 6, 59-72.

Meloni, T., and Rosenheck, A. (1995). "Choice of frequency weighting for the evaluation of weapon noise," *J. Acoust. Soc. Am.* 97(6), 3636-3641.

Moore, B. C. J., and Galsberg, B. R. (1996). "A revision of Zwicker's loudness models," *Acustica* 82(2), 333-345.

Moore, B. C. J., Glasberg, B. R., and Baer, T. (1997). "A model for the prediction of thresholds, loudness and partial loudness," *J. Audio Eng. Soc.* 45(4), 224-240.

Nilsson, M. E. (2007). "A-weighted sound pressure level as an indicator of perceived loudness and annoyance of road-traffic sound," *J. Sound Vib.* 302(1-2), 197-207.

Persson, K., and Björkman, M. (1988). "Annoyance due to low frequency noise and the use of the dB(A) scale," *J. Sound Vib.* 127(3), 491-497.

Quirt, J. D. (1983). "Sound transmission through windows II. Double and triple glazing," *J. Acoust. Soc. Am.* 74(2), 534-542.

Schomer, P. D., Suzuki, Y., and Saito, F. (2001). "Evaluation of loudness-level weightings for assessing the annoyance of environmental noise," *J. Acoust. Soc. Am.* 110(5), 2390-2397.

Schomer, P. D. (2004). "The importance of proper integration of and emphasis on the low-frequency sound energies for environmental assessment," *Noise Control Eng.* 52(1), 26-39.

Stevens, S. S. (1956). "Calculation of the loudness of complex noise," *J. Acoust. Soc. Am.* 28(5), 807-832.

Versfeld, N. J., Vos, J., and Geurtsen, F. W. M. (1994). "Annoyance of individual vehicles pass-by noise for light and heavy vehicles," *J. Acoust. Soc. Am.* 96(5), 3335.

Versfeld, N. J., and Vos, J. (1997). "Annoyance caused by sound of wheeled and tracked vehicle," *J. Acoust. Soc. Am.* 101(5), 2677-2685.

Torija, A. J., and Flindell, I. H. (2014a). "Differences in subjective loudness and annoyance depending on the road traffic noise spectrum," *J. Acoust. Soc. Am.* 135(1), 1-4.

Torija, A. J., and Flindell, I. H. (2014b). "Listening laboratory study of low height roadside noise barrier performance compared against in-situ field data", "Build. Environ. 81, 216-255.

Torija, A. J., and Flindell, I. H. (2015). "The subjective effect of low frequency content in road traffic noise," *J. Acoust. Soc. Am.* 137(1), 189-198.

Turpin, A., Scholer, F., Mizzaro, S., and Maddalena, E. (2015). "The benefits of magnitude estimation relevance assessments for information retrieval evaluation", in *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, New York, USA, pp. 565-574.

Wang, Y. S., Lee, C. -M., Kim, D. -G., and Xu, Y. (2007). "Sound-quality prediction for nonstationary vehicle interior noise based on wavelet pre-processing neural network model," *J. Sound Vib.* 2007(4-5), 933-947.

Watts, G. R. (1995). "A comparison of noise measures for assessing vehicle noisiness," *J. Sound Vib.* 180(3), 493-512.

Yifan, H., Guoqing, D., Yiting, Z., Youpeng, H., and Bangjun, Z. (2008). "Pair-wise comparison experiment on subjective annoyance rating of noise samples with different frequency spectrums but same A-weighted level," *Appl. Acoust.* 69(12), 1205-1211.

## Tables/figure captions

TABLE I. Summary of the experimental procedures used in the five listening tests.

	Objective	Stimuli	Experimental procedure	Perceptual assessment method
Listening test 1	Road-traffic-noise (outdoors)	15-s recording of continuous urban road traffic noise ( $L_{Aeq,15 \text{ sec}} = 70.3 \text{ dB(A)}$ )	Low shelf filter, band pass/band stop filter and high shelf filter applied for boosting/cutting the LF, MF and HF ranges. 12 filtered sounds synthesised by applying each filter with -9dB, -3 dB, +3dB and +9 dB relative gain setting.	Standardized ISO/TS 15666 specification for noise annoyance questionnaires. 5 point scale: “Not at all”, “Slightly”, “Moderately”, “Very”, and “Extremely”.
Listening test 2	Road-traffic-noise (indoor)	2 x 12.5-s recording of continuous urban road traffic noise filtered to simulate indoor conditions.  Part 1: $L_{Aeq,12.5 \text{ sec}} = 49.5 \text{ dB(A)}$  Part 2: $L_{Aeq,12.5 \text{ sec}} = 39.5 \text{ dB(A)}$	Low shelf filter and high shelf filter applied for boosting/cutting the LF and MHF ranges. 8 + 8 filtered sounds synthesized (Part1 + Part 2) by applying each filter with -9dB, -3 dB, +3dB and +9 dB relative gain setting.	Relative magnitude estimation (RME) method.
Listening test 3	Road-traffic-noise (noise barriers)	2 x 12.5-s recording of continuous urban road traffic noise.  RT1: $L_{Aeq,12.5 \text{ sec}} = 67.1 \text{ dB(A)}$ (significant content in LF)	For each master recording (RT1 and RT2), simulation of the IL of 5 noise barrier (made by different materials) plus the same 5 noise barriers with additional 6 dB IL.	Relative magnitude estimation (RME) method.

---

		RT2: $L_{Aeq,12.5\text{ sec}} = 71.4$ dB(A) (low content in LF)	20 stimuli altogether.	
Listening test 4	Subjective equivalence between LF, MF and HF road-traffic-noise (ref: pink-noise)	8-s second master recording of a road vehicle passing-by ( $L_{Aeq,8\text{ sec}} = 67.6$ dB(A)).	Each test sound was compared against a 70 dB(Z) pink noise reference sound, with the filtered road vehicle sound level varied up and down in 5 dB.	The comparison was repeated until it was found the sound level at which the test sound was perceived as equally annoying as the reference sound.
Listening test 5	Subjective equivalence between LF, MF and HF road-traffic-noise (ref: LF, MF and HF filtered pink-noise)	10-s recording of continuous urban road traffic noise ( $L_{Aeq,10\text{ sec}} = 67.2$ dB(A)).	Each test sound was compared against a 70 dB(Z) frequency filtered pink noise reference sound, with the filtered road vehicle sound level varied up and down in 5 dB.	The comparison was repeated until it was found the sound level at which the test sound was perceived as equally annoying as the reference sound.

---

TABLE II. Starting off sound-level (dB(Z)) for each test sound in listening tests 4 and 5.

		Sound-level (dB(Z))
Listening test 4	LF	65.9
	MF	64.6
	HF	65.1
Listening test 5	LF	70.0
	MF	70.0
	HF	70.0

TABLE III. Results of the linear regression analysis (N=12) for estimating reported annoyance from A-, C-, D- and Z-weighted sound-levels, and Zwicker's loudness-level for the stimuli used in listening test 1.  $p \leq 0.05$ .

		A-weighting	C-weighting	D-weighting	Z-weighting	Zwicker's Loudness
RT1	R <sup>2</sup>	0.50	0.38	0.78	0.34	0.81
	Standard Error of the Estimate	0.42	0.47	0.28	0.48	0.25
	F	10.08	6.04	34.77	5.21	43.59

TABLE IV. Results of the linear regression analysis (N=8) for estimating reported annoyance from A-, C-, D- and Z-weighted sound-levels, and Zwicker's loudness-level for the stimuli used in listening test 2.  $p \leq 0.05$ .

		A- weighting	C- weighting	D- weighting	Z- weighting	Zwicker's Loudness
Part 1: $L_{Aeq,12.5 \text{ sec}} =$ 49.5dB(A)	$R^2$	0.97	0.77	0.84	0.76	0.91
	Standard Error of the Estimate	4.16	11.25	9.46	11.32	6.94
	F	182.26	19.80	30.47	19.44	61.76
Part 2: $L_{Aeq,12.5 \text{ sec}} =$ 39.5dB(A)	$R^2$	0.96	0.73	0.81	0.73	0.87
	Standard Error of the Estimate	4.08	10.76	9.19	10.83	7.42
	F	150.82	16.54	24.87	16.23	41.37



TABLE V. Results of the linear regression analysis (N=10) for estimating reported annoyance from A-, C-, D- and Z-weighted sound-levels, and Zwicker's loudness-level for the stimuli used in listening test 3.  $p \leq 0.05$ .

		A-weighting	C-weighting	D-weighting	Z-weighting	Zwicker's Loudness
RT1	$R^2$	0.87	0.79	0.91	0.80	0.92
	Standard Error of the Estimate	2.88	3.60	2.34	3.56	2.26
	F	52.75	30.87	83.96	31.86	90.34
RT2	$R^2$	0.97	0.82	0.97	0.82	0.98
	Standard Error of the Estimate	2.28	5.72	2.29	5.70	1.78
	F	265.72	35.38	262.75	35.80	442.352

TABLE VI. Subjective equivalence between LF, MF and HF test sounds in listening test 4, for each frequency weighting evaluated. In brackets there are shown the confidence intervals for the mean ( $p \leq 0.05$ ).

	A-weighting	C-weighting	D-weighting	Z-weighting
LF-MF	-1.0 (-3.7 – 1.7)	10.5 (7.8 – 13.2)	3.0 (0.3 – 5.7)	10.8 (8.1 – 13.5)
MF-HF	3.1 (1.0 – 5.2)	5.5 (3.4 – 7.5)	-1.9 (-4.0 – 0.2)	4.6 (2.4 – 6.6)
LF-HF	2.2 (-1.3 – 5.6)	15.9 (12.4 – 19.4)	1.1 (-2.4 – 4.6)	15.3 (11.8 – 18.8)

TABLE VII. Subjective equivalence between LF, MF and HF test sounds in listening test 5, for each frequency weighting evaluated.

In brackets there are shown the confidence intervals for the mean ( $p \leq 0.05$ ).

	A-weighting	C-weighting	D-weighting	Z-weighting
LF-MF	-0.1 (-2.4 – 2.1)	11.0 (8.7 – 13.3)	3.9 (1.6 – 6.2)	11.3 (9.1 – 13.6)
MF-HF	4.5 (2.6 – 6.5)	6.9 (5.0 – 8.9)	-1.1 (-3.1 – 0.8)	6.0 (4.0 – 7.9)
LF-HF	5.9 (2.7 – 9.1)	19.4 (16.2 – 22.7)	4.3 (1.1 – 7.5)	18.8 (15.6 – 22.0)

## Figure captions

FIG. 1. Frequency spectra of master recording urban road-traffic-noise used in listening test 1 (outdoors).

FIG. 2. Frequency spectra of the master recording and indoors-filtered urban road-traffic sounds used in listening test 2 (indoors).

FIG 3. Frequency spectra of the original and frequency filtered road-traffic sounds to simulate the erection of the different noise barrier, for master urban road-traffic sounds RT1 (A) and RT2 (B) in listening test 3.

FIG. 4. Frequency spectra of the frequency filtered road vehicle test sounds (A) used in listening test 4. In (B) it is shown the frequency spectra of the frequency filtered road vehicle test sounds with sound-levels at starting off points, LF = 65.9 dB(Z), MF = 64.6 dB(Z) and HF = 65.1 dB(Z).

FIG. 5. Frequency spectra of the frequency filtered road traffic test sounds (A) used in listening test 5. (B) and (C) shows the frequency spectra of the frequency filtered road traffic test sounds and filtered pink noise reference sounds respectively, with sound-levels at 70 dB(Z).

FIG. 6. Linear relationship between the Z-weighted (A), A-weighted (B), C-weighted (C) and D-weighted (D) sound-levels, and Zwicker's loudness-level (E) and reported annoyance in listening test 1 (outdoors). Square, asterisk and triangle symbols correspond to LF, MF and HF filter gain respectively.

FIG. 7. Linear relationship between the Z-weighted (A), A-weighted (B), C-weighted (C) and D-weighted (D) sound-levels, and Zwicker's loudness-level (E) and reported annoyance in listening

test 2 (indoors), road-traffic-noise stimulus = 49.5  $L_{Aeq}$ . Square and asterisk symbols correspond to LF and MHF filter gain respectively.

FIG. 8. Linear relationship between the Z-weighted (A), A-weighted (B), C-weighted (C) and D-weighted (D) sound-levels, and Zwicker's loudness-level (E) and reported annoyance in listening test 2 (indoors), road-traffic-noise stimulus = 39.5  $L_{Aeq}$ . Square and asterisk symbols correspond to LF and MHF filter gain respectively.

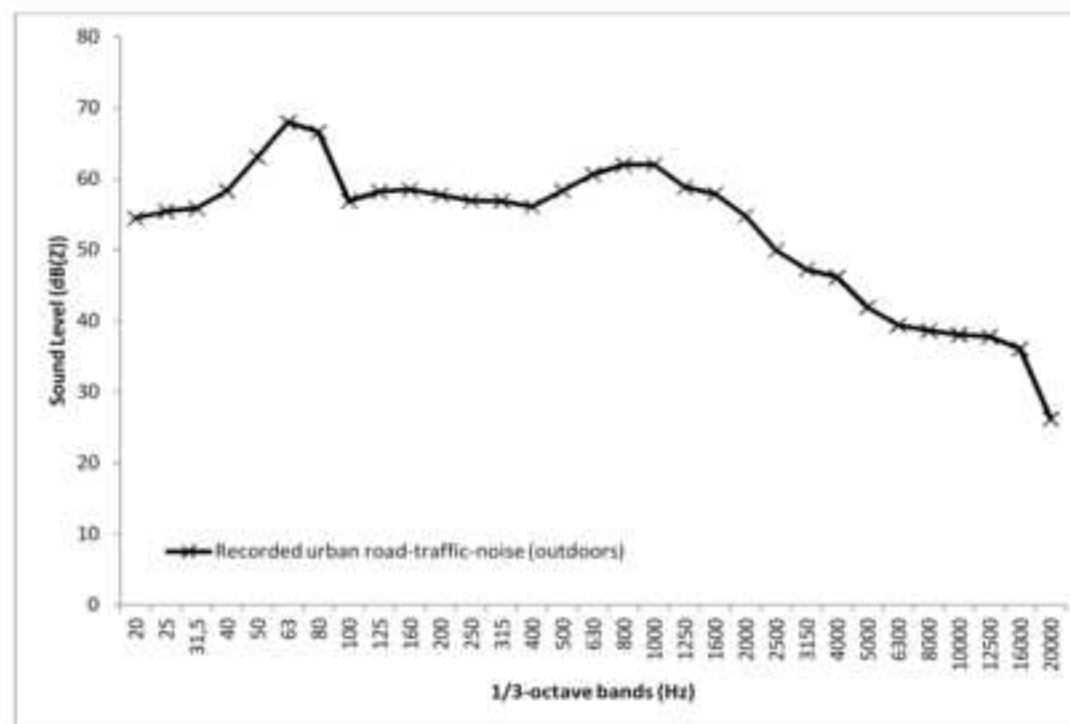
FIG. 9. Linear relationship between the Z-weighted (A), A-weighted (B), C-weighted (C) and D-weighted (D) sound-levels, and Zwicker's loudness-level (E) and reported annoyance in listening test 3 (noise barriers). Diamond and square symbols correspond to RT1 and RT2 road-traffic-noise stimuli.

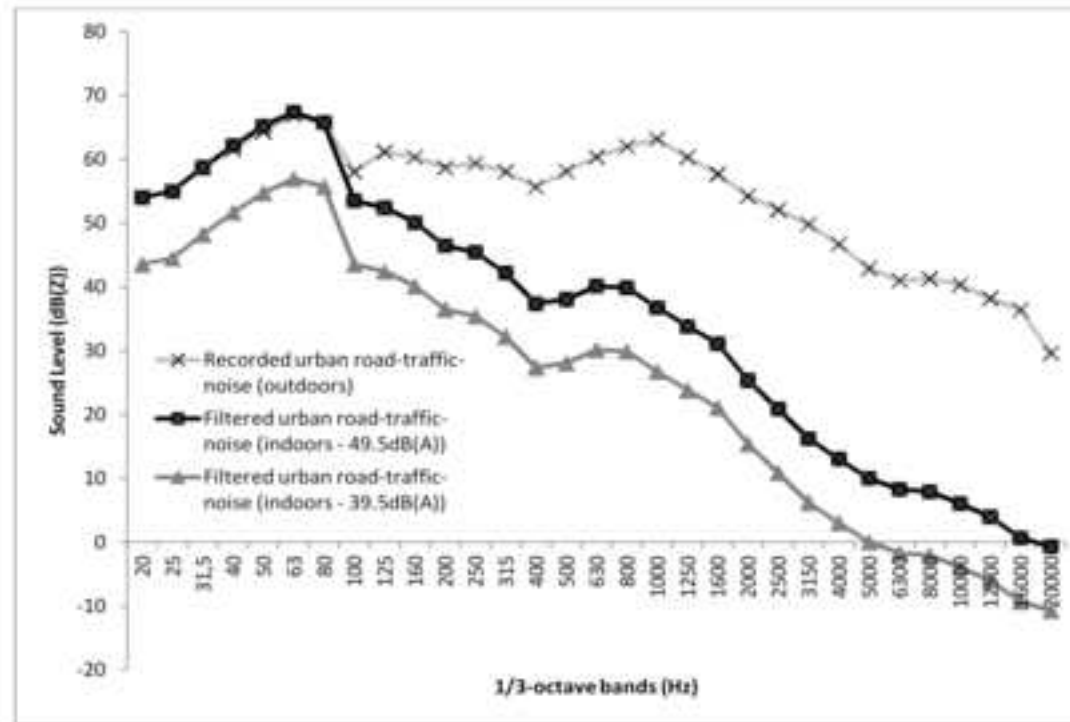


Click here to access/download

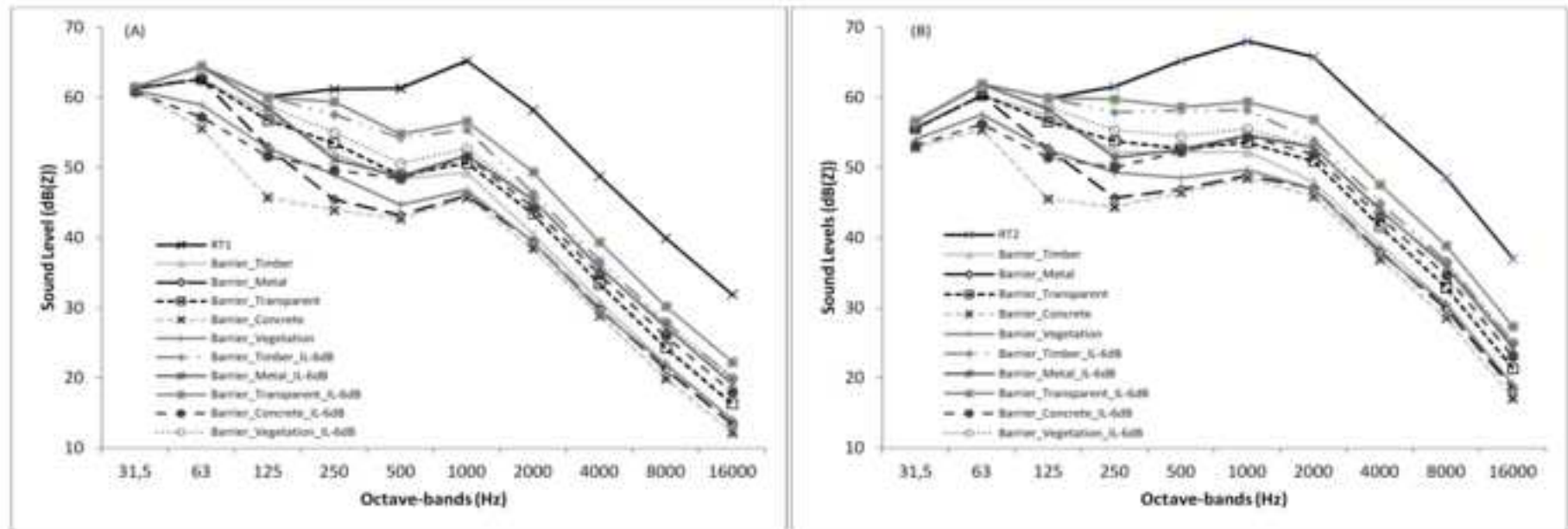
**Reviewer PDF with line numbers, inline figures and captions**

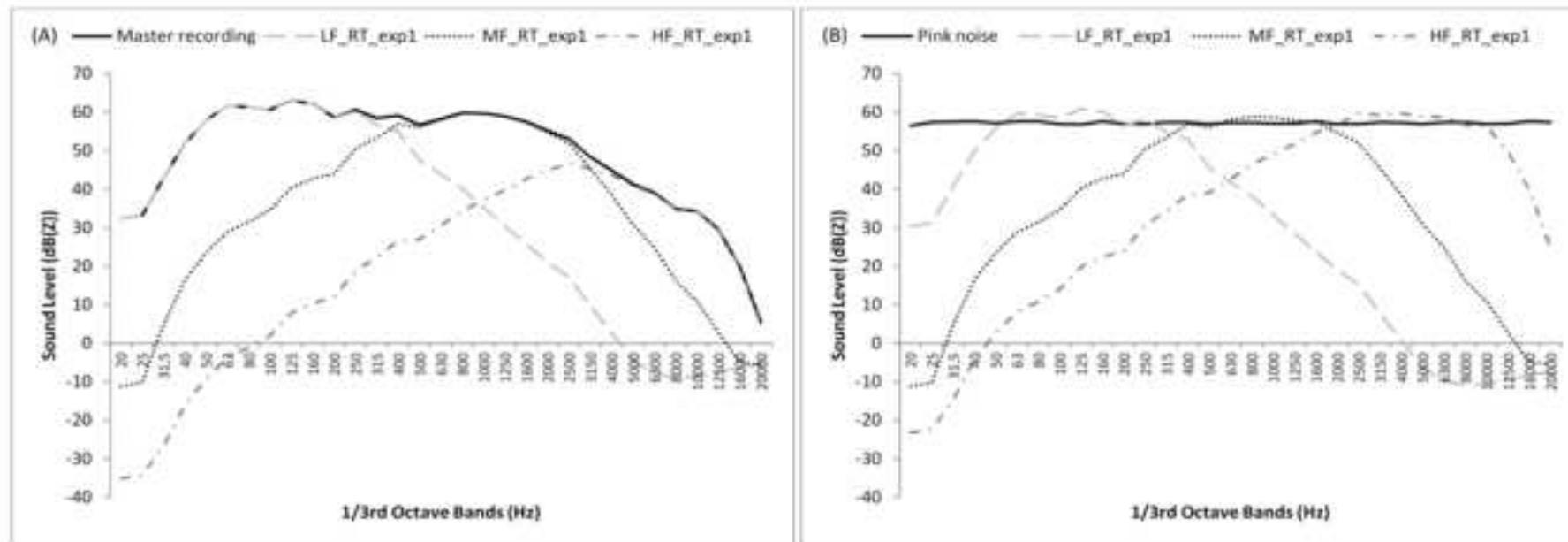
Revised\_pdf\_for\_reviewers.pdf

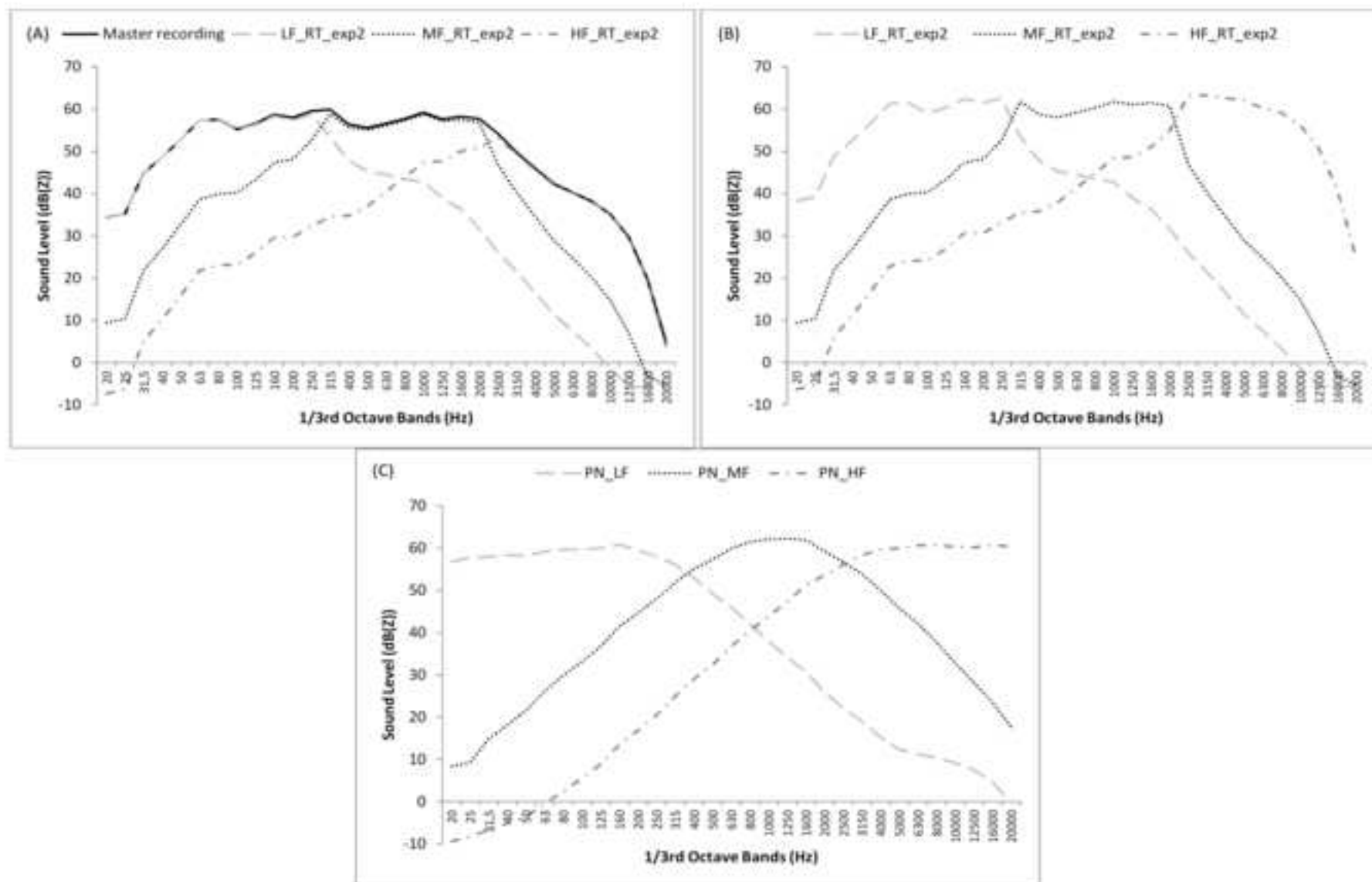


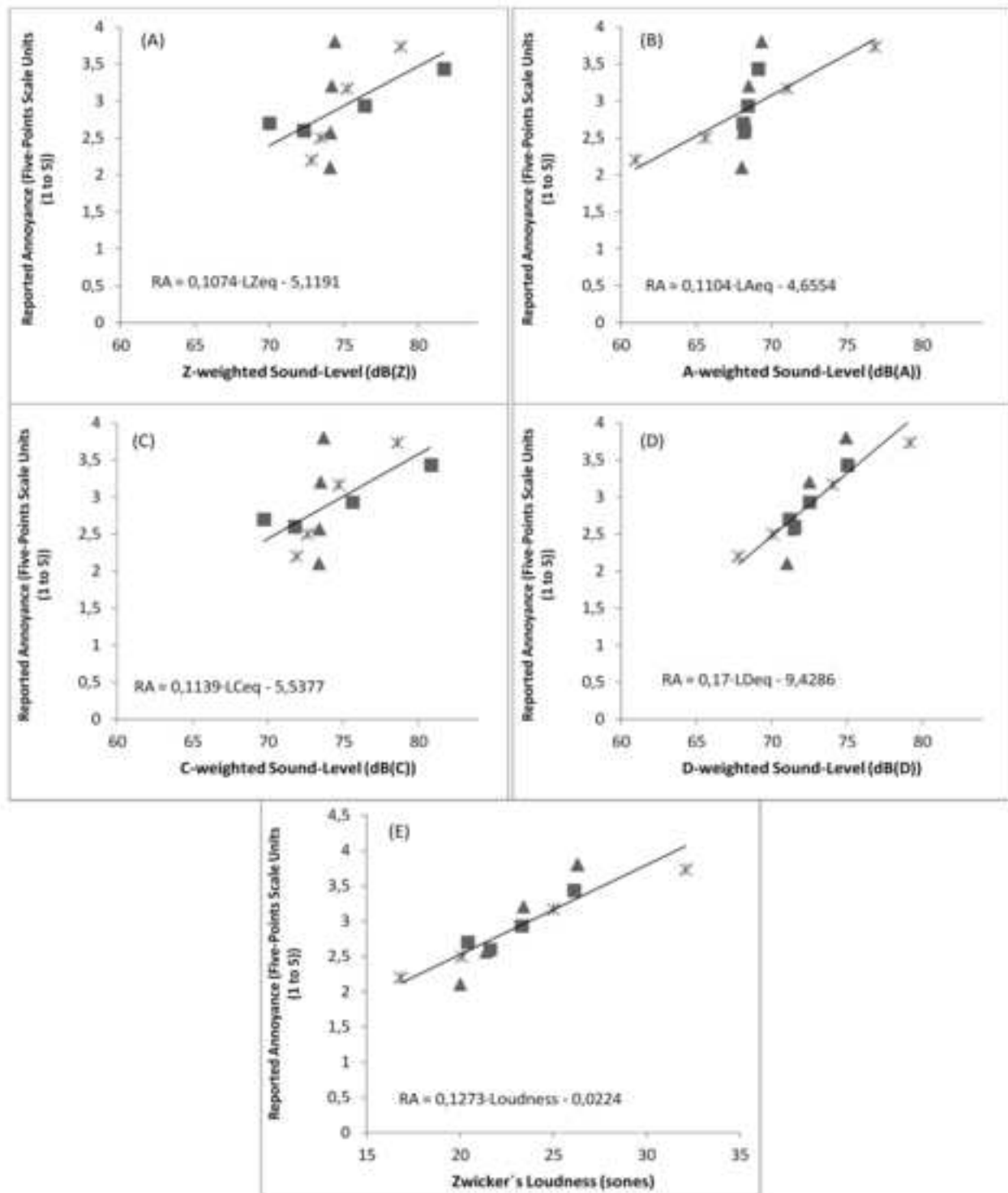


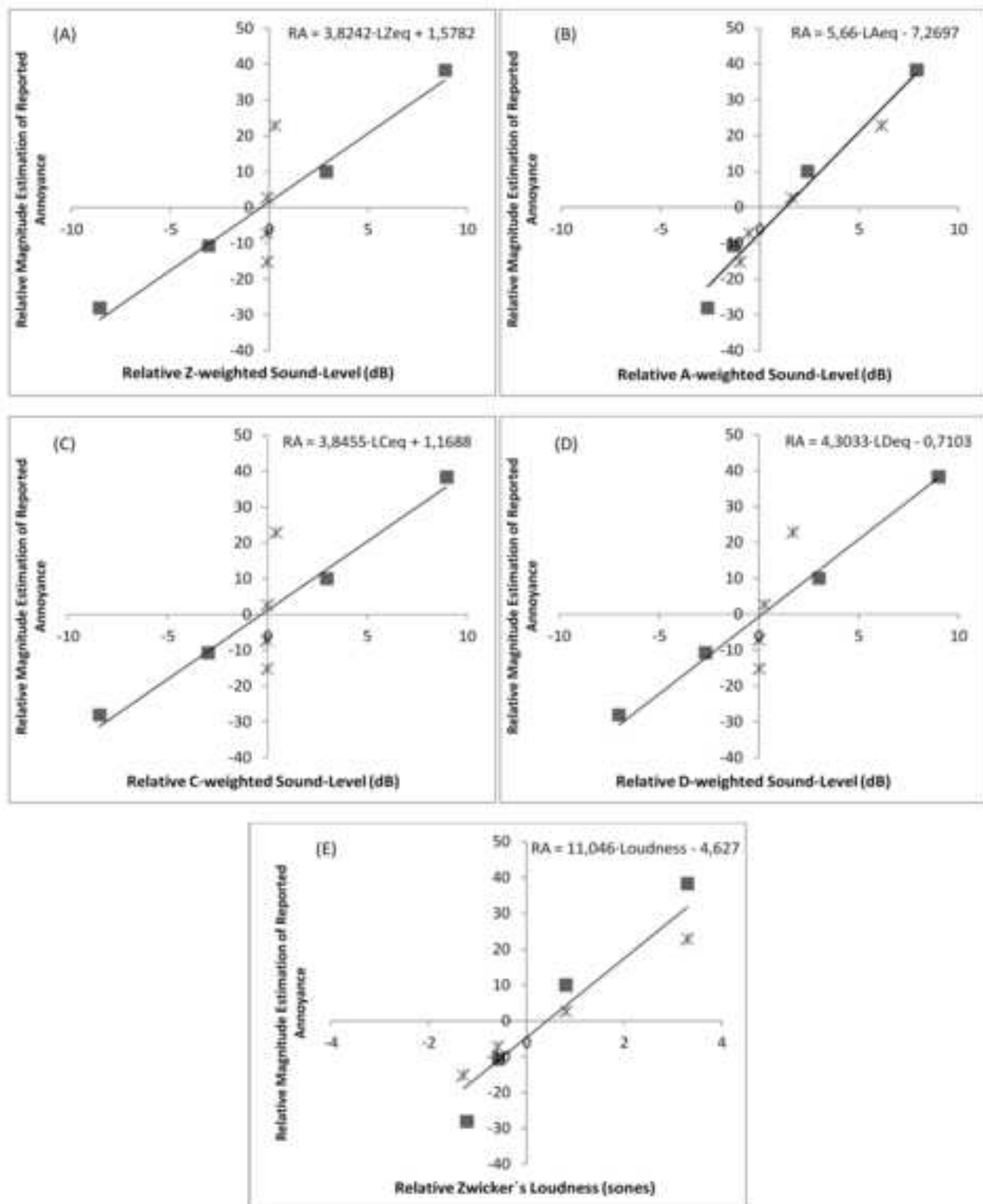


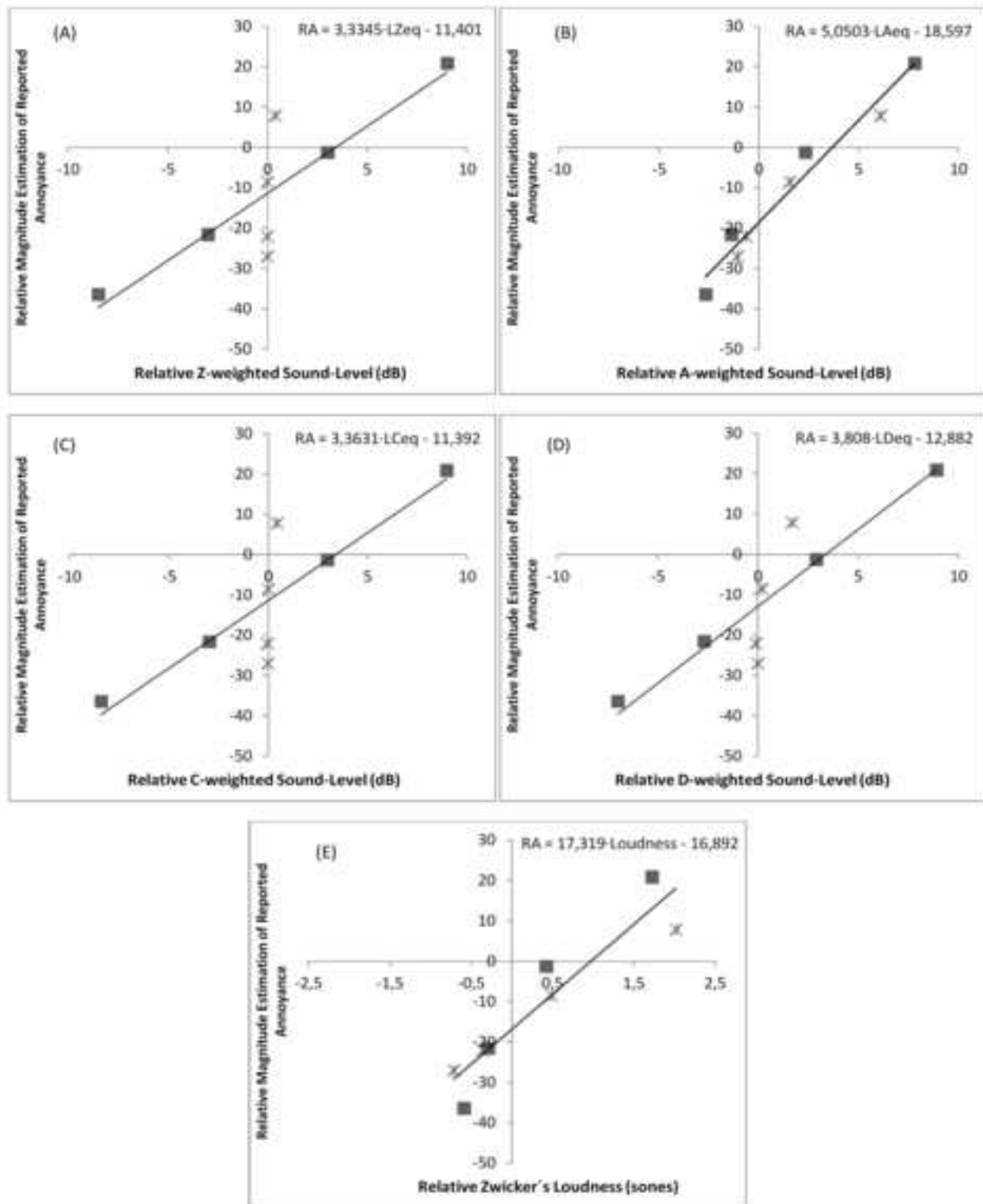


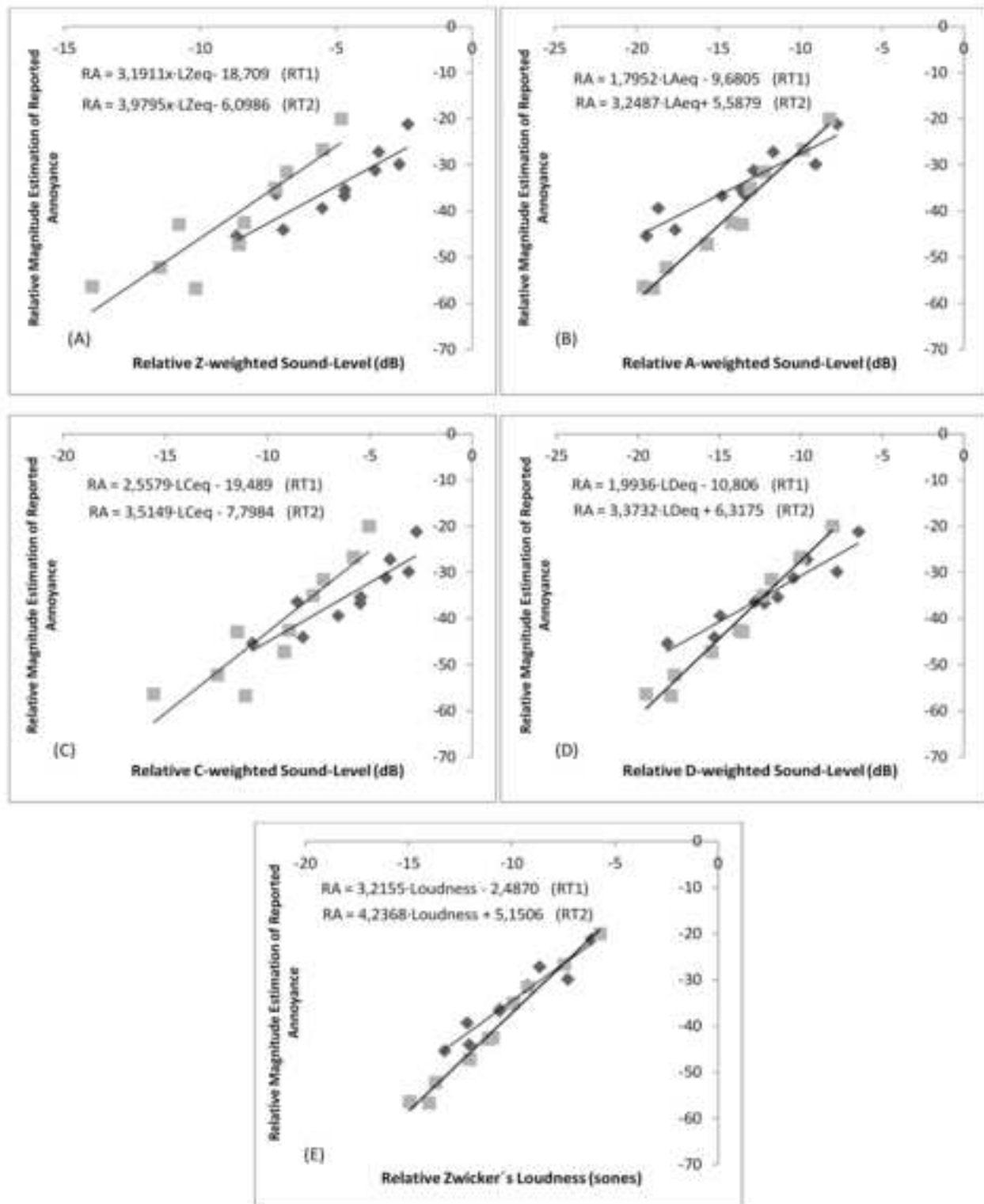


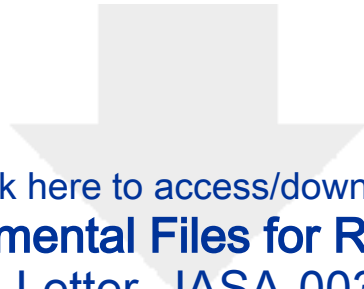












[Click here to access/download](#)

**Supplemental Files for Reviewer**  
Transmittal\_Letter\_JASA-00366\_v2.docx

