ON TOTAL ANNOYANCE CAUSED BY DIFFERENT ENVIRONMENTAL SOUNDS: A REVIEW AND SUGGESTIONS FOR ADDITIONAL RESEARCH

J. Vos
TNO Human Factors, P.O. Box 23, 3769 ZG, Soesterberg, Netherlands
Tel.: +31 346 356 223 / Fax: +31 346 353 977 / Email: vos@tm.tno.nl

Keywords:
TOTAL ANNOYANCE, CUMULATIVE EFFECTS, MULTIPLE SOURCES, SUMMATION MODEL

ABSTRACT
Various procedures for rating total annoyance caused by different environmental sounds have been proposed. It is explained that the adequacy of the majority of these models is very limited. A more promising weighted summation model is described in which for each single source in the combination, the sound exposure will be first expressed as the level of equally annoying road-traffic sound. Next, weighted summation of the adjusted levels results in the total rating sound level. The model successfully explained total annoyance ratings obtained in the laboratory for various combinations of different simultaneously presented sounds. To explore the general applicability of the model, suggestions for additional research are given.

1 - INTRODUCTION
A considerable part of the population is exposed to simultaneous and/or successive environmental sounds from different sources. To prevent or reduce the annoyance in areas with two or more different sound sources, insight is needed into the population’s total annoyance caused by all sounds together. Any procedure for rating combined environmental sounds should meet at least a few requirements. Firstly, the procedure must be able to cope with the systematic differences among the (source-specific) dose-response relationships for the constituting sources. At the same A-weighted equivalent sound pressure levels (ALEQ), aircraft sounds are more annoying, and railway sounds are less annoying than road-traffic sounds (e.g., see [1]). Similarly, shooting sounds produced by small firearms are more annoying than road-traffic sounds (e.g., see [2]). The comparability among the community response to environmental sounds can be increased by adding adjustments to the noise dose, relative to the dose-response relationship for road-traffic sounds (e.g., see the new ANSI method described in [3]). If single event level and number of events are exchangeable, as it is assumed to be the case for traffic sounds and the impulse sounds produced by small firearms, the adjustments may be directly applied to ALEQ. For shooting sounds produced by medium-large and large firearms, and for other high-energy impulsive sounds, the adjustments should be directly applied to the sound exposure level of the single events (e.g., see [4-8]). It is assumed here that total annoyance is related to a weighted sum of the adjusted sound levels.

Secondly, the total rating sound level for multiple sound sources should never be lower than the rating sound level for the more annoying source. Stated differently, a procedure that predicts total annoyance from multiple sources to be lower than the maximum of the source-specific annoyance of the separate sources, is regarded as being invalid and must therefore be rejected.

Examples of models that do not meet both requirements just described (also see [9]), are the response summation model [10], the summation and inhibition model [11], and the energy summation or energy difference models [12].

Both requirements are satisfied in a weighted summation model with the general formula

\[ L_t = k \log \left\{ \frac{\sum_{i=1}^{n} 10^{(\text{adjusted } L_{Aeq} \text{ of source } i)/k}}{k} \right\} \]
in which $L_t$ is the total rating sound level for the combined sources, and $k$ is a constant. The adequate value of $k$ has to be determined on the basis of experimental data.

**2 - ILLUSTRATION OF THE MODEL**

The promise of the weighted summation model defined in Eq. (1) is illustrated with the help of the results from laboratory studies reported in Vos [13].

**2.1 - Method**

In one of the experiments (Experiment 2), gunfire (G), road traffic (R), and aircraft (A) served as distinct sources, which were presented in isolation and in various combinations. Briefly, G consisted of about 20 pistol shots reproduced at variable time intervals. T was based on recordings of free-flowing road traffic. A consisted of one passage of a trident (3 jet engines). Remote-traffic sound at $L_{Aeq} = 30$ dB served as a soft continuously present background noise. Sixteen subjects were individually tested. They were told that they were going to be presented with various environmental sounds which could consist of impulse, road-traffic, and/or aircraft sounds, each sound or sound combination lasting for 45 s. After each period they had to respond to the question "How annoying would you find the sound in the preceding period if you heard it at home on a regular basis?" A rating scale with values from 0 ("not annoying at all") to 9 ("extremely annoying") was used.

**Figure 1:** Annoyance as a function of ALEQ for three separately presented sounds.

**2.2 - Results for sounds in isolation**

Mean annoyance ratings for G, T, and A are shown in Fig. 1 as a function of ALEQ of the sounds. The linear regression lines inserted into Fig. 1 explain 98% or 99% of the variance in the mean annoyance ratings ($y$). For G, T, and A, the respective equations are $y = -6.65 + 0.241L_{Aeq}$, $y = -1.34 + 0.179L_{Aeq}$, and $y = -5.75 + 0.210L_{Aeq}$. The ratings were subjected to an analysis of variance (ANOVA). Overall, G was more annoying than T and A. The most powerful effect on the ratings was caused by ALEQ. The increase of the ratings with ALEQ was greater for T than for G and A. The level-dependent penalties for G and A relative to T are given by $P_g = 22.0 - 0.26L_{Aeq}$ and by $P_a = 3.7 - 0.13L_{Aeq}$, respectively. The significance of these level-dependent penalties was confirmed in separate ANOVAs.

**2.3 - Results for combinations with two or three different sounds**

$G+A$

The mean ratings for G combined with A49 are given in Fig. 2a as a function of ALEQ of G. Fig. 2a clearly shows that especially in the conditions in which the differences between the annoyance ratings for the separate sounds were relatively small (which was the case in G28+A49, G33+A49, and G38+A49),
the total annoyance rating was higher than the rating of the more annoying sound. A t-test showed that the annoyance caused by G33+A49 was significantly higher than (1) the annoyance for G33, and (2) the annoyance for A49 in isolation.

A+T

The ratings for A combined with T40 are shown in Fig. 2b. In the conditions in which the differences between the annoyance ratings for A and T in isolation were relatively small, the total annoyance rating was higher than the rating of the more annoying sound. T-tests showed that (1) the annoyance caused by A40+T40 was significantly higher than the annoyance for T40 in isolation, and (2) the annoyance caused by A45+T40 was significantly higher for the interpolated annoyance for A45.

T+A+G33

The ratings for T and A combined with G33 are shown in Fig. 2c. In all conditions the total annoyance rating was significantly higher than the rating of the more annoying sound in the combination. T-tests showed that (1) the annoyance caused by T42+A45+G33 was significantly higher than the annoyance for G33, and (2) the annoyance caused by T47+A50+G33 was significantly higher than the annoyance caused by A50.

![Figure 2: Total annoyance for various sound combinations, with the annoyance caused by the separate sounds as references; the solid lines represent model predictions.](image)

2.4 - Applicability of the model

The model is considered successful if the relation between the total annoyance rating and the predictions of the total rating sound level $L_t$ is equal to the relation between annoyance and ALEQ for road-traffic sound T in isolation. Here, the discrepancies between the predictions and the experimental results are expressed as decibel values.
The data shown in Fig. 2a are used as an example. First, the ALEQs of G and A (columns 1 and 2 in Table I) are converted into ALEQ of equally annoying T by adding level-dependent penalties to ALEQ of the respective sources (columns 3 and 4 in Table I). Second, the model with $k = 10$ is applied, yielding the $L_t$-values shown in column 7. Third, the total annoyance ratings are converted into the ALEQ of equally annoying T in isolation (columns 5 and 6). The discrepancies between the model predictions and the experimental results are given in column 8. Averaged across the combinations of G and A, the model underestimates $L_t$ by 2.3 dB; the root mean square of the differences (rms) equals 2.5 dB. The underestimation is reduced by application of $k$-values greater than 10. With $k = 19.5$ (iterative procedure) the underestimation was completely cancelled (see columns 9 and 10). In Fig. 2a, the model predictions with $k = 19.5$ are represented by the solid line. For the data shown in Fig. 2b and Fig. 2c, unbiased predictions ($M = 0$ dB) with low rms-values were obtained with $k = 13$ (rms = 0.8 dB) and $k = 14.9$ (rms = 0.7 dB), respectively.

Based on all 49 conditions investigated in Experiment 1 (Vos, [13]), optimal predictions of $L_t$ were obtained with $k = 16$. For the 48 combinations included in Experiment 2, the adequate value for $k$ was equal to 14. Considering this remarkably consistent result, obtained in two separate experiments with different subjects and partly different stimuli, it seems obvious to set $k$ to a mean value of 15 when applying the model.

### Table 1: Application of the model for gunfire (G) in combination with aircraft (A) sounds.

<table>
<thead>
<tr>
<th>$L_{Aeq}$</th>
<th>adjusted $L_{Aeq}$</th>
<th>rating</th>
<th>corrresp. $L_{Aeq}$ of T</th>
<th>$k = 10$</th>
<th>$k = 19.5$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>pred. $L_t$</td>
<td>dif.</td>
</tr>
<tr>
<td>G</td>
<td>A</td>
<td>G</td>
<td>A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>49</td>
<td>39.1</td>
<td>46.4</td>
<td>4.8</td>
<td>47.7</td>
</tr>
<tr>
<td>28</td>
<td>49</td>
<td>42.8</td>
<td>46.4</td>
<td>5.6</td>
<td>50.8</td>
</tr>
<tr>
<td>33</td>
<td>49</td>
<td>46.5</td>
<td>46.4</td>
<td>5.7</td>
<td>51.2</td>
</tr>
<tr>
<td>38</td>
<td>49</td>
<td>50.3</td>
<td>46.4</td>
<td>6.8</td>
<td>55.6</td>
</tr>
<tr>
<td>43</td>
<td>49</td>
<td>54.0</td>
<td>46.4</td>
<td>7.2</td>
<td>57.6</td>
</tr>
<tr>
<td>48</td>
<td>49</td>
<td>57.7</td>
<td>46.4</td>
<td>8.0</td>
<td>60.7</td>
</tr>
<tr>
<td>53</td>
<td>49</td>
<td>61.4</td>
<td>46.4</td>
<td>8.5</td>
<td>62.9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>dif.</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>M=</td>
<td>-2.3</td>
<td>M=</td>
<td>0.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rms=</td>
<td>2.5</td>
<td>rms=</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The results from the studies reported in Vos [13] support a weighted summation model which accurately predicts the total rating sound level for combinations of several types of sounds, both in conditions in which two or more sounds are about equally annoying and in conditions in which one of the sounds is much more annoying than the sounds of the remaining sources. Only for the latter conditions are the predictions with the present model similar to the predictions given by a dominant source model. Since weighted summation of the ALEQs of the separate sources is performed after addition of the level-dependent penalties, the total annoyance model represents a mental integration of annoyance judgments rather than an integration of the ALEQs from separate sources.

### 3 - SUGGESTIONS FOR ADDITIONAL RESEARCH

The present model successfully explained total annoyance ratings obtained in the laboratory for various combinations of different simultaneously presented sounds. In the future, it should be investigated whether the model also holds for successive sounds presented within a relevant part of the day or even within a 24 hours’ day. Moreover, to explore the general applicability of the model, the laboratory results have to be supplemented with responses from representative samples of residents who are exposed to multiple sound sources in their living environments for sufficiently long periods of time.

### 3.1 - Questionnaires

From the various counterintuitive results obtained in previous field surveys (e.g., see [13]), it must be concluded that questions in the questionnaire need to be designed very carefully. Actually, the development of questionnaires to measure how people integrate annoying events is still in its infancy. This holds true for total annoyance caused by multiple sound sources, as well as for total annoyance resulting from the sound exposure during the day, evening and night, and for total annoyance resulting from the sounds heard in indoor and in outdoor conditions.
3.2 - Method of paired comparison

For pragmatic reasons, it is strongly suggested to determine the total rating sound level for the combined sources directly. This can be achieved by inviting the residents to participate in specially designed listening experiments carried out in a (mobile) laboratory. The sound fragments to be used should closely match the sounds heard in the residential areas with respect to source type, sound level, and sound spectrum. Since one common reference source will be used, road traffic must be one of the relevant sources. The listening tests should continue only if the resident under investigation feels that the various sources sound realistically. Next, the level of the reference sound source is determined at which a) this sound is as annoying as the combination of the various sounds, and b) this sound is as annoying as the non-reference sounds in isolation, yielding the adjusted levels. This information may be obtained by means of the method of paired comparison. As a result of the possible occurrence of biases [14], the method of adjustment is not recommended. For a successful determination of the required data, various versions of the reference sound have to be prepared. The main difference among these versions is the overall sound level. Since higher overall sound levels are associated with smaller source-receiver distances, at least a portion of the versions should also be different with respect to spectral content. With the help of these data and the weighted summation model, the optimal $k$-value and the overall precision of the predictions can be determined.

REFERENCES


