



Rating the Dieselness of engine-sounds

H. Fastl^a, B. Priewasser^a, M. Fruhmann^a and H. Finsterhölzl^b

^aAG Technische Akustik, MMK, TU München, Arcisstr. 21, 80333 München, Germany

^bBMW Group, Knorrstr. 147, 80788 München, Germany

fastl@mmk.ei.tum.de

Sounds of idling Diesel engines were rated according to their Dieselness with two different psychoacoustic methods. On the one hand, a yes/no procedure was used. Subjects were presented sounds of 1 500 ms duration and had to answer by “yes” or “no” to the following question: “Does the sound which you just heard stem from a Diesel engine?” Histograms were calculated and a hypothesis was put forward as follows: The more positive answers are given to a specific sound, the larger is the Dieselness of this sound. On the other hand, the method of paired comparison was used, applying a Bradley-Terry-Luce (BTL) procedure. Both procedures yield rather similar results with an advantage of the yes/no-procedure in efficiency. Comparisons of psychoacoustic evaluations with predictions of Dieselness by four different algorithms show much room for improvement of the instrumental evaluations.

1 Introduction

Traditionally, Diesel engines had an image to be weak, noisy, and sooting. Although modern Diesel engines have overcome most of these preconceptions, the typical sound character of Diesel engines, called Dieselness, is still an issue. Even today it is not “expected” that luxury sedans are driven by a Diesel engine, despite significant torque and quite acceptable exterior and interior noise.

In this situation it is desirable to rate the Dieselness of engine sounds. Therefore, psychoacoustic experiments on the rating of Dieselness of engine sounds were performed. In addition, psychoacoustic data were compared to instrumental data of four algorithms for the evaluation of Dieselness.

2 Experiments

The BMW-Group provided a CD with a calibration signal and 40 sounds with not too much detail about their features. Sounds came from idling engines with around 700 rpm and to our knowledge had been recorded by a dummy head in a semi-anechoic chamber of BMW (cf. Finsterhölzl 2006, 2008). The dummy head had been situated at a distance of 2 m from the front wheels and 1.65 m above ground. Listening to the 40 sounds, the first author guessed that they included sounds from Diesel and Gasoline engines as well as about five stochastic sounds, which did not stem from an engine of which type soever. In order to keep the details secret and to minimize the danger of influencing the subjects by prior knowledge, the information of BMW about the sounds was pretty skimpy. In addition the BMW-Group provided instrumental Dieselness ratings of the 40 sounds by four different algorithms, the details of which again were not disclosed.

Two different groups of subjects participated in the experiments: One group of nine persons (one female, eight male) encompassed experts from the BMW-Group. The other group of male non-experts included eight persons who for the first time participated in psychoacoustic experiments, three experienced members of the lab, and eight students of the faculty of electrical engineering and information technology of the Technische Universität München. The age of the subjects ranged from 24 to 60 years (median 26 years).

Subjects were seated in a sound attenuating booth. Sounds of 1 500 ms duration with 40 ms Gaussian-shaped gating signals were presented dichotically via electrodynamic headphones (Beyer DT48) with free-field equalizer (Fastl

and Zwicker, 2007, p. 7). Two sets of experiments were performed: In a first set, a yes/no-procedure was applied. Subjects were presented a sound and had to answer by “yes” or “no” to the question: “Does the sound which you just heard stem from a Diesel engine?” No time limit was given for the response, and after the subject’s hitting the “y” or “n” key of the keyboard, the next sound was presented. As a rule the response interval was less than 1 000 ms. Each sound was presented four times in different order to get more stable results and to minimize order effects. For each sound, from each of the 26 subjects four responses were collected, leading to 104 data points from which medians and interquartiles were calculated. As a whole, in this set of experiments 4.160 data points were collected.

In a second set of experiments, sounds were arranged in pairs with 1 500 ms signal duration and 300 ms interval. Without limit for response time, subjects had to answer the following question: “Does the first sound or the second sound produce more Dieselness?” Accordingly, they had to enter “1” or “2” into the keyboard. To limit the number of comparisons necessary, a BTL-procedure was used (e.g. Thurstone, 1927, Bradley and Terry, 1952, Burton, 2003) to arrive at 390 pairs. In this set of experiments participated 22 male non-expert subjects .

3 Psychoacoustic results

In figure 1, the positive answers of all subjects to the question “Does the sound which you just heard stem from a Diesel engine?” are displayed for all 40 sounds presented. Medians and interquartiles calculated from the respective 104 data points for each sound are given. The synthetic sounds 5, 14, 30, 36 and 38 are clearly identified, i.e. nobody mistook them as Diesel sounds. Pretty clear Diesel sounds like 11, 33 or 35 are easily identified with rather small interquartiles. On the other hand, in the mid range, larger interquartiles suggest somewhat larger inter-individual differences.

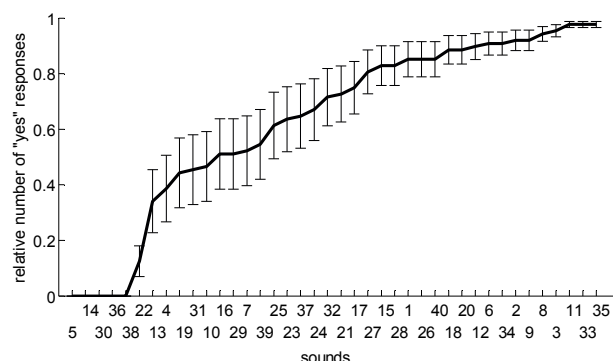


Fig. 1 Rating of sounds as stemming from a Diesel engine by all subjects.

Figure 2 gives the results of the yes/no-procedure for the nine experts. For each sound, 36 data are given as medians with interquartiles. As for the non-experts, synthetic sounds are easily uncovered and for the other sounds, interquartiles usually are smaller. Again, in the mid range, larger interquartiles due to inter-individual differences show up.

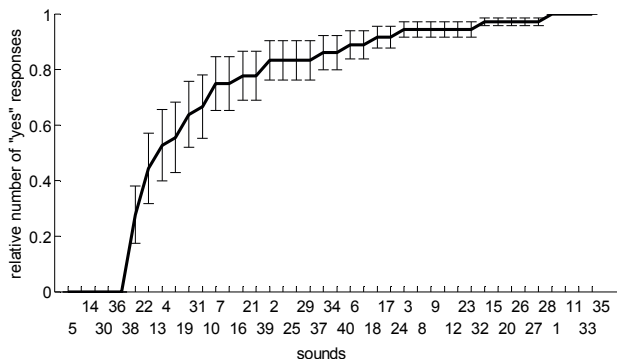


Fig. 2 Rating of sounds as stemming from a Diesel engine by experts.

The data displayed in figure 3 enable a direct comparison of expert versus non-expert rating. By and large, there is fair agreement between the rating by experts and non-experts. However, for the experts, the transition is steeper. Obviously, the experts are more familiar with modern Diesel sounds, whereas non-experts sometimes have difficulty to distinguish sounds from modern Diesel engines and sounds from Gasoline engines.

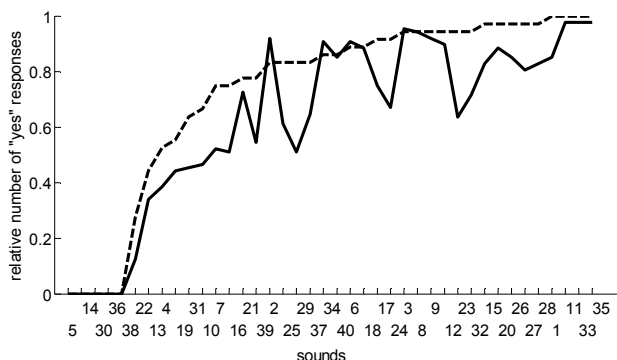


Fig. 3 Comparison of ratings of sounds as stemming from a Diesel engine by experts (dashed) versus non-experts (solid).

In figure 4, the results from the paired comparisons are displayed. The relative Dieselness is given for the sounds presented. For some sounds (e.g. 25, 32, 12, 24) a larger variance shows up. Since this set of experiments was performed by non-experts, these sounds were not so easy to rate. The first author guesses that the sounds could stem from modern Diesel engines with advanced motor management or from Gasoline engines with direct injection. The synthetic sounds 38, 30, 5 again were identified; however, some subjects felt that synthetic sound 36 and in particular synthetic sound 14 would contain some Dieselness.

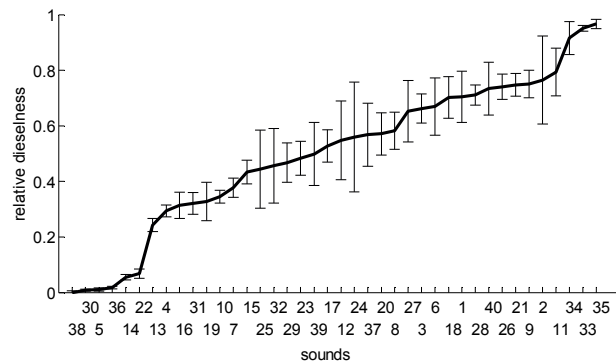


Fig. 4 Rating of Dieselness by non-experts with the paired comparison-procedure.

The results displayed in figure 5 enable a comparison of ratings obtained by the yes/no-procedure versus the paired comparison-procedure. The relative Dieselness is plotted for the sounds evaluated. To ease the comparison, results for the yes/no-procedure (dashed) were normalized relative to the data for sound 18 of the paired comparison-procedure (solid). As a rule, both procedures lead to rather similar ratings. However, for sounds 15, 12, 20, and 8 the yes/no-procedure yields substantially higher ratings, for sound 21, the opposite is true.

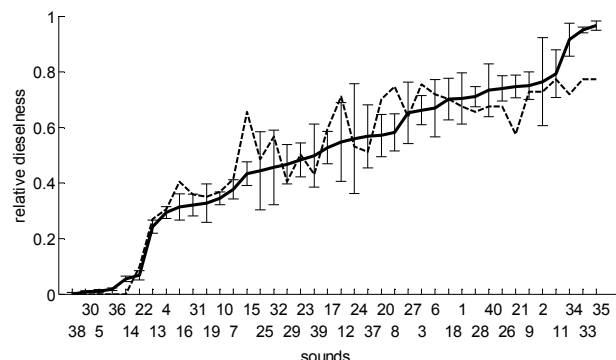


Fig. 5 Comparison of Dieselness ratings by non-experts using a yes/no-procedure (dashed) versus a paired comparison-procedure (solid).

4 Comparison of psychoacoustic and instrumental evaluations

The results displayed in figure 6 enable a comparison of the psychoacoustic evaluations of the Dieselness for the 40 sounds used with the predictions by four different instrumental algorithms to calculate Dieselness. The psychoacoustic ratings of Dieselness as obtained by the paired comparison-procedure are given by the solid curve, the instrumental predictions by the dashed curves.

As a rule, instrumental ratings of Dieselness (dashed) deviate substantially from the psychoacoustic ratings (solid). In most cases, for the synthetic sounds, large values of Dieselness were predicted by the algorithms, whereas results from psychoacoustic evaluations suggested essentially no Dieselness. It should be mentioned that in the third panel, data from an algorithm are displayed, for which the synthetic sounds were not measured.

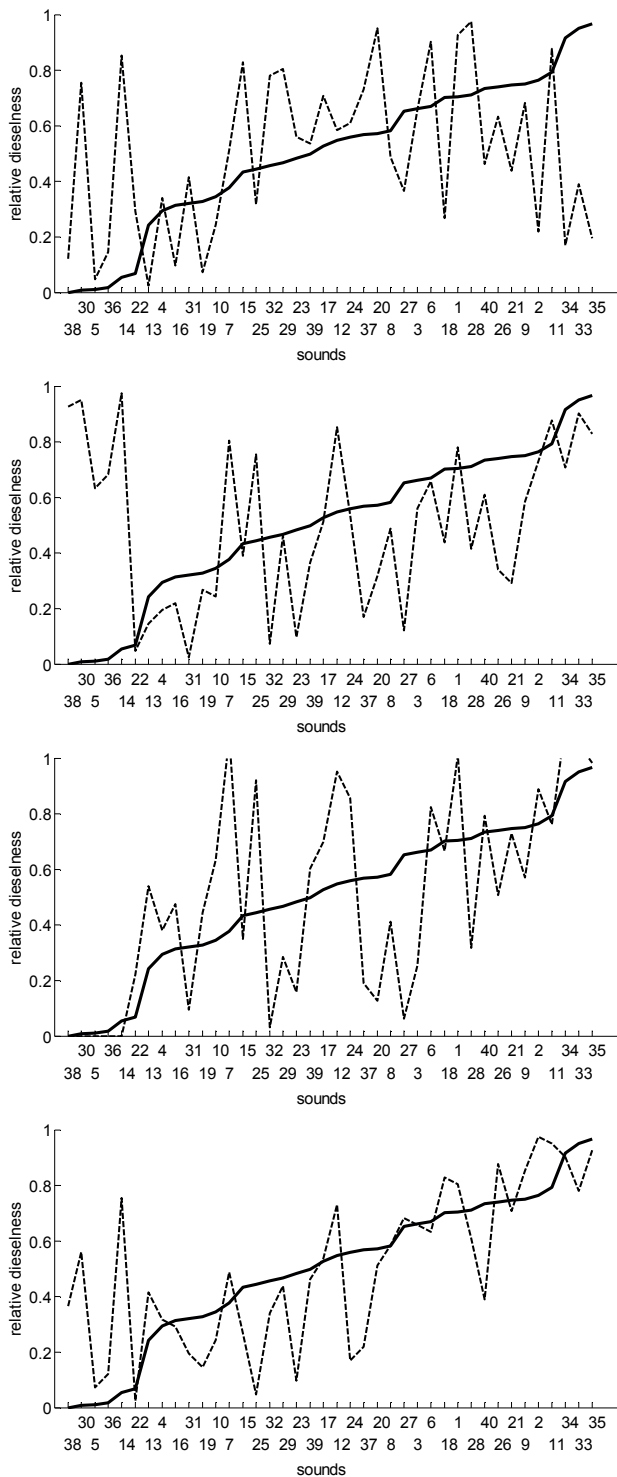


Fig. 6 Rating of Dieselness by non-experts (solid) in comparison to predictions by four different algorithms (dashed).

5 Discussion

Dieselness of engine sounds can be rated with great precision by experts and with considerable precision by non-experts. Although a paired comparison-procedure gives somewhat more precise results, a simple yes/no-procedure is preferred because of its efficiency with respect to the ratio of precision versus measurement time. Psychoacoustic evaluations of Dieselness differed considerably from predictions by four different algorithms.

In particular, synthetic sounds with no Dieselness were easily detected by subjects, but frequently mistaken as Diesel sounds by the algorithms. On the other hand, concerning the typical "nocking" or "nailing" of Diesel sounds, an algorithm was reported (e.g. Bodden et al. 2008) which can describe the impulsiveness of the sounds in line with subjective evaluations.

Acknowledgments

The authors wish to thank Dipl.-Ing. Daniel Menzel for editorial support.

References

- [1] R. A. Bradley, M. E. Terry (1952), Rank analysis of incomplete block designs, I. the method of paired comparisons. *Biometrika*, 39, 324-345.
- [2] M. Bodden, R. Heinrichs, M. R uth (2008), Diesel-Impulshaltigkeit Teil II: Technischer Prozess, Fortschritte der Akustik, DAGA 08, DEGA Berlin
- [3] M. L. Burton (2003), Too many questions? The use of incomplete cyclic designs for paired comparisons, *Field Methods* 15, 115-130
- [4] H. Fastl, E. Zwicker, (2007) *Psychoacoustics-Facts and Models*, 3rdEd., Springer Heidelberg, New York
- [5] H. Finsterh lzl, V. Caldiero, J. Hobelsberger, W. Baumann, F. Daiber (2006), Neuer Au enger uschpr ufstand im Entwicklungsprozess bei BMW, *ATZ* 108, 256-267
- [6] H. Finsterh lzl (2008), Pr ftechnik in der Au enger uschentwicklung, Fortschritte der Akustik, DAGA 08, DEGA Berlin
- [7] L. L. Thurstone (1927). A law of comparative judgement. *Psychological Review*, 34, 278-286.