CORRELATION OF OBJECTIONABILITY RATINGS OF NOISE WITH PROPOSED NOISE-ANNOYANCE MEASURES

RR-3

MAY 1964

U. S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
Public Health Service
Division of Occupational Health

CORRELATION OF OBJECTIONABILITY RATINGS OF NOISE WITH PROPOSED NOISE-ANNOYANCE MEASURES

Alexander Cohen
and
Richard F. Scherger

Occupational Health Research and Training Facility
1014 Broadway
Cincinnati, Ohio

RR-3 MAY 1964

U. S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
Public Health Service
Division of Occupational Health

TABLE OF CONTENTS

PAGE NO.

FOREWORD i

SUMMARY 1

INTRODUCTION 2

METHODOLOGY 4

RESULTS AND DISCUSSION 5

REFERENCES 8

APPENDIX 9

FOREWORD

The evergrowing problem of community noise has prompted interest in the development of methods capable of predicting annoyance magnitudes for noises typically intruding in neighborhoods. Several procedures for weighting or converting physical measurements of such noises into expressions of their perceived annoyance have been formulated and are presently undergoing empirical validation. Consistent with these efforts, the present study attempted to evaluate the ability of these various proposed procedures to predict noise-annoyance judgments of listeners as given to community noises of a widely different character.

CORRELATION OF OBJECTIONABILITY RATINGS OF NOISE WITH PROPOSED NOISE-ANNOYANCE MEASURES

SUMMARY

In a laboratory setting, recorded samples of roadway noise, aircraft flyover noise, and train noise were presented in a forced-choice, paired-comparison fashion to 100 listeners who judged which of the two noises in each pair was more objectionable. Such judgments yielded scaled objectionability ratings for the noise samples. These scaled ratings were correlated with A-scale sound level readings of the noises in decibels (dB), and with conversions of their physical measurements into subjective magnitudes of loudness in phon units as computed by Stevens' and by Zwicker's techniques, and with perceived noisiness in dB as determined by Kryter's procedure. Each of these measures has been reputed to provide quantifiable estimates of noise-annoyance and did, in fact, supply values significantly correlated with the obtained scale of subjective ratings. Loudness in phons, using Zwicker's calculations, gave values which corresponded best with the scaled listener annoyance judgments ($r = .96$).

OBJECTIONABILITY JUDGMENTS AND NOISE-ANNOYANCE MEASURES

INTRODUCTION

Holding all situational factors constant, what measures of a noise can best serve to quantify its objectionability or annoyance? At present, several different procedures have been proposed to convert physical measurements of noise into numerical expressions of their annoyance level (1, 4, 6, 8, 9). The more popular of these proposed methods involve conversions into loudness measures as developed by Stevens (11) and by Zwicker (12), and into perceived noisiness in decibels as developed by Kryter (5). Also, A-scale sound level values in decibels (dB) for given noises, as read on a conventional sound level meter, have been used to quantify their objectionability. Each of these conversion procedures and the A-scale determinations, in various ways, take account of established relationships between the physical dimensions of sound (primarily frequency and intensity), and associated auditory reactions, both psychological and physiological. For example, Stevens' loudness measures for describing noise-annoyance reflect the following relationships: 1) The changes in apparent loudness as a function of varying the frequency bandwidth of a given sound. 2) The relative loudness of different single frequencies and bands of frequencies at specifiedintensity levels. 3) The growth in loudness with increase in intensity for different single frequencies and bands of frequencies.

In Stevens' loudness calculations, spectral measurements of a noise, i.e., the distribution of sound energy across the frequencies comprising the noise in third-octave, half-octave, or full octave bands, are converted by a series of formuli into sone or phon units on a loudness scale. Loudness magnitudes described by these units are used to quantify the annoyance level of a noise under the assumption that loudness is the chief determinant in causing objectionability to sound.

Another loudness measure used for determining the amount of noise-annoyance has been formulated by Zwicker. His formulation takes into account the same relationships noted earlier for Stevens' loudness measures, but also includes a consideration of the excitation patterns along the basilar membrane, produced by critical bands of frequencies. It toodescribes loudness magnitudes and therein annoyance magnitudes in terms of sones and phons. Due to the added factor underlying Zwicker's loudness derivations, however, Zwicker's sone or phon values for a given noise may not agree with the sone or phon values as obtained by Stevens' calculations for the same noise. A further difference is that Zwicker's procedure requires third-octave band spectral analyses of the noises under evaluation, whereas Stevens' computations utilize third-octave, half-octave, or full octave band spectrum data.

OBJECTIONABILITY JUDGMENTS AND NOISE-ANNOYANCE MEASURES

A third conversion method, perceived noisiness in decibels, describes magnitudes of noise-annoyance on a scale which has been expressly designed for that purpose by Kryter. Underlying Kryter's scaled values are considerations and computations quite similar to those already noted with Stevens' loudness measures, except that more weight is given to the higher frequency components in the noise spectrum. This is based on findings showing higher frequency sounds to be judged more annoying, but not necessarily louder, than lower frequency sounds of the same intensity (4, 10). Noise-annoyance magnitudes on Kryter's scale are expressed in PNdB units and utilize third-octave, half-octave, or full octave band spectral values in their computation.

A more direct procedure used for gauging the amount of noise-annoyance is the A-scale measurement in decibels (dB) on the sound-level meter. The frequency response of the A-scale meter network takes account of a specific function relating sound intensity and frequency to a given amount of loudness (7). In effect, the A-scale response gives more weight to sound energy in the middle and high frequencies than in the low frequencies.

Previous tests of the ability of these various procedures to measure the magnitude of noise-annoyance have involved inter-comparisons of the apparent noisiness, or objectionability, of different noise samples drawn from a particular class. That is, different samples of jet aircraft noise or truck noise or automobile noise have been presented to listeners who have either rated, ranked, or equated them in terms of perceived objectionability. Their judgments were then evaluated for correspondence with values representing conversions of the spectra of these noises into loudness units, PNdB, and with direct A-scale readings. Under these test conditions, each of these procedures yielded values which were highly correlated with and equally capable of making predictions of the objectionability judgments for the class of noises presented (1, 4, 5, 9). More information is needed, however, with regard to the adequacy of these different procedures for predicting noise-annoyance magnitude based upon inter-comparing the annoyance caused by sounds of widely different character, e.g., between jet aircraft and truck noise, between train noise and automobile noise, between aircraft noise and train noise, etc. The identification of a method suitable for making such predictions would be extremely important, especially in considering the problem of community noise-nuisance where there may be many different types of noises intruding in a neighborhood.

The purpose of this study was to supply some preliminary information bearing upon the generalizability of the above-mentioned procedures in predicting noise-annoyance magnitudes. Specifically, samples of aircraft flyover noise, train noise, and roadway noise were presented to listeners who inter-compared and rated their objectionability in accordance with a forced-choice, paired-comparison procedure. This procedure enabled scaled values to be assigned to the objectionability ratings of the noises heard. These values, in turn, were correlated with measures of noise-annoyance, as derived from loudness, PNdB, and A-scale calculations, to identify which of the measures corresponded best with the listener scaled ratings. All of the observed correlations were also contrasted with the relationship obtained by comparing listener scaled ratings with measurements of the overall physical energy contained in the noise samples. The overall energy level of the noise is simply the decibel readings on the C or flat network of a sound-level meter.

METHODOLOGY

Tape recordings of noises previously collected in community noise surveys were reviewed for selecting the samples to be used in this study. Nine samples were chosen, three of which were aircraft flyover noises (two jet, one prop), three roadway noises (two fast-moving auto traffic, one truck) and three train noises (two slow moving, one high speed). Third-octave band spectral analyses of these noises, as measured in the listening area used for the noise rating tests, are shown in Figures 1-3. The playback levels of the noise samples were similar to their original levels as observed in the field and were measured by a condenser microphone (Bruel-Kjaer, Model 4130) coupled to a noise spectrometer (Bruel-Kjaer, Model 2112). Sizeable differences are apparent between the spectra of the noises for each class of vehicles, as well as between the vehicle classes. By re-recording, two test tapes were made, each composed of 45 pairings involving the nine different vehicular noises. The two tapes differed from one another only in that the order of noises in each pair was reversed. Each tape was presented an equal number of times in the experiment. The first nine pairs of each tape served for practice and provided two exposures to each of the nine selected noise samples. The remaining 36 pairs exhausted all possible pairing of the nine noise conditions and constituted the actual test material. Each noise sample lasted 20 seconds, there being a three-second pause between the two samples in a given pair and an eight-second interval between two successive pairs.

The test tapes were reproduced in a sound-treated laboratory area, using a high fidelity tape recorder (Ampex 351) with the signal being amplified

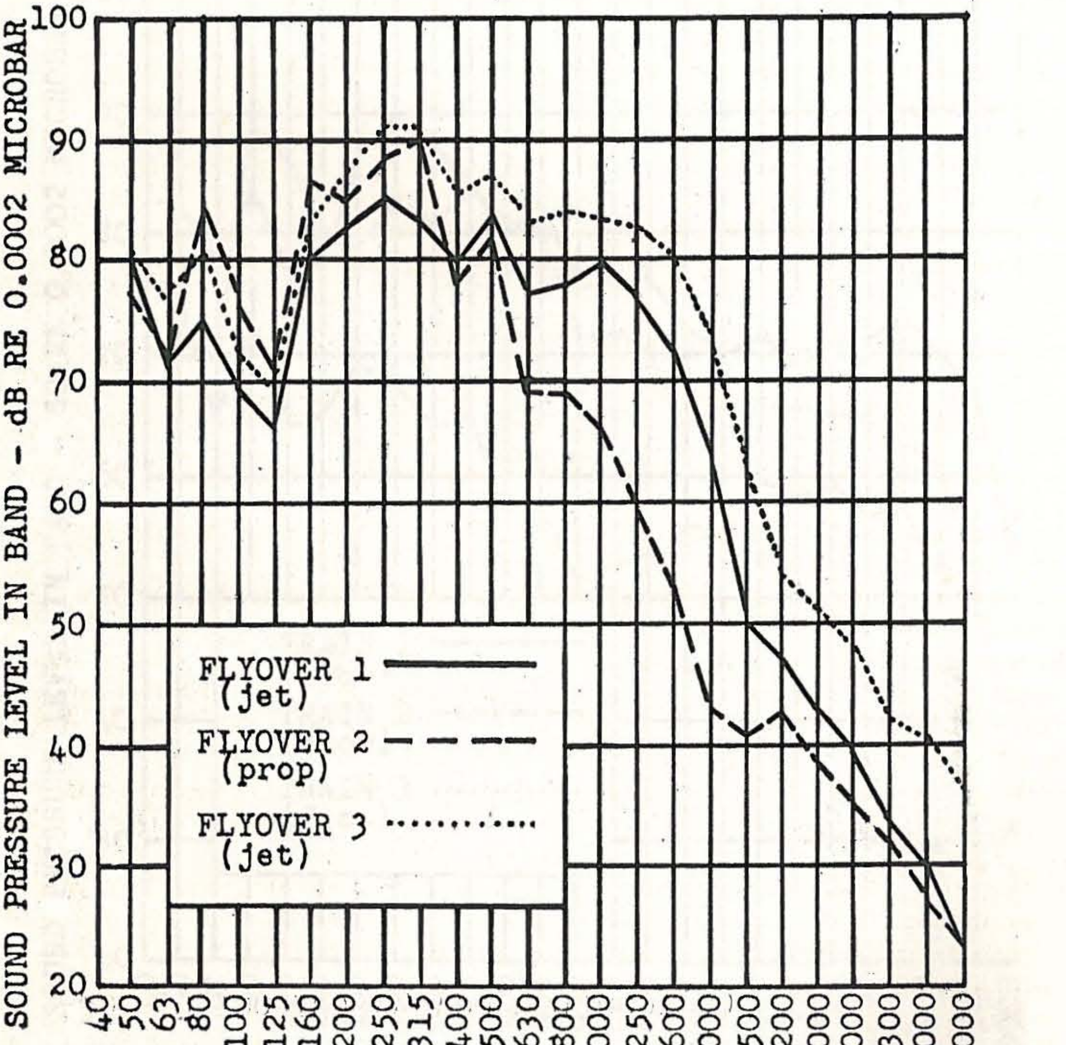

Figure 1. Third octave-band spectrum analyses of three aircraft flyover noises.

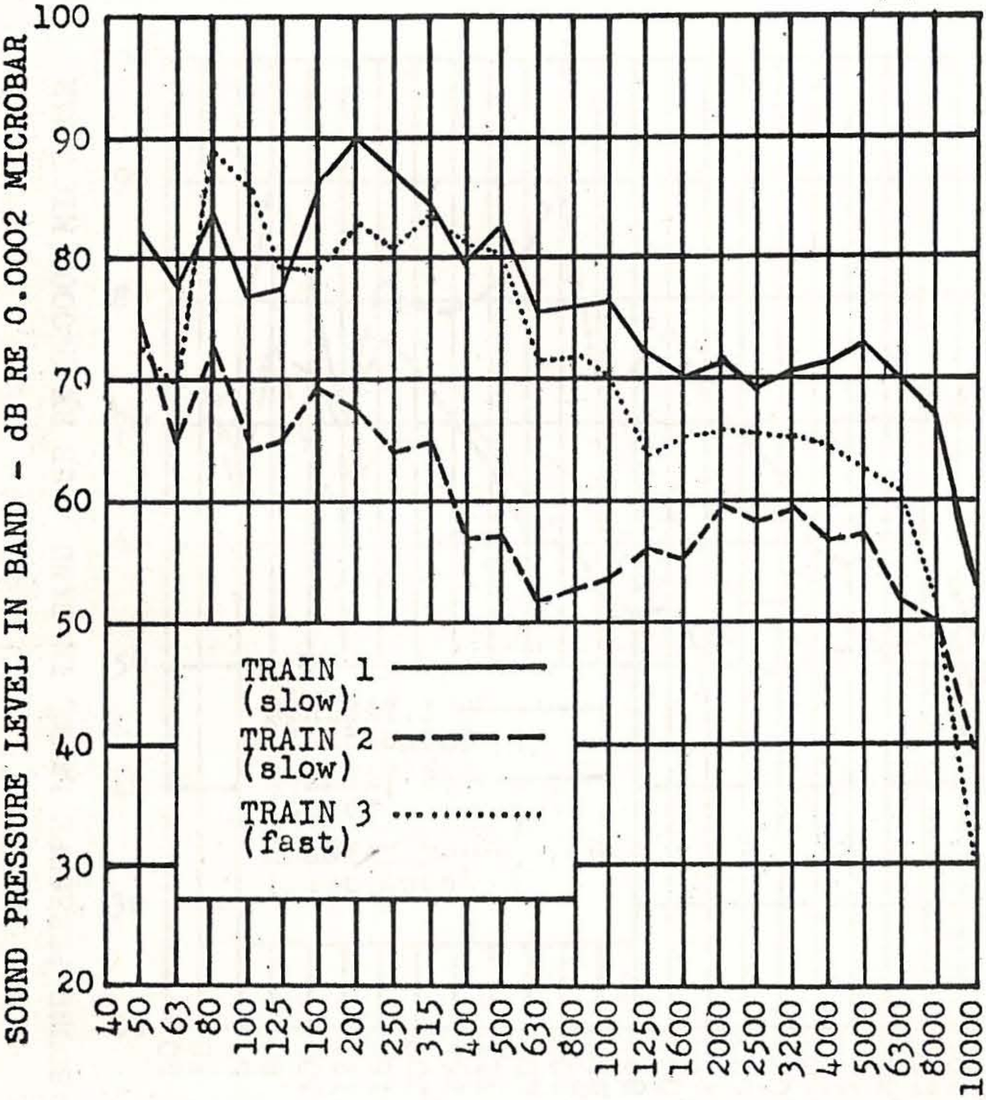

Figure 2. Third octave-band spectrum analyses of three train noises.

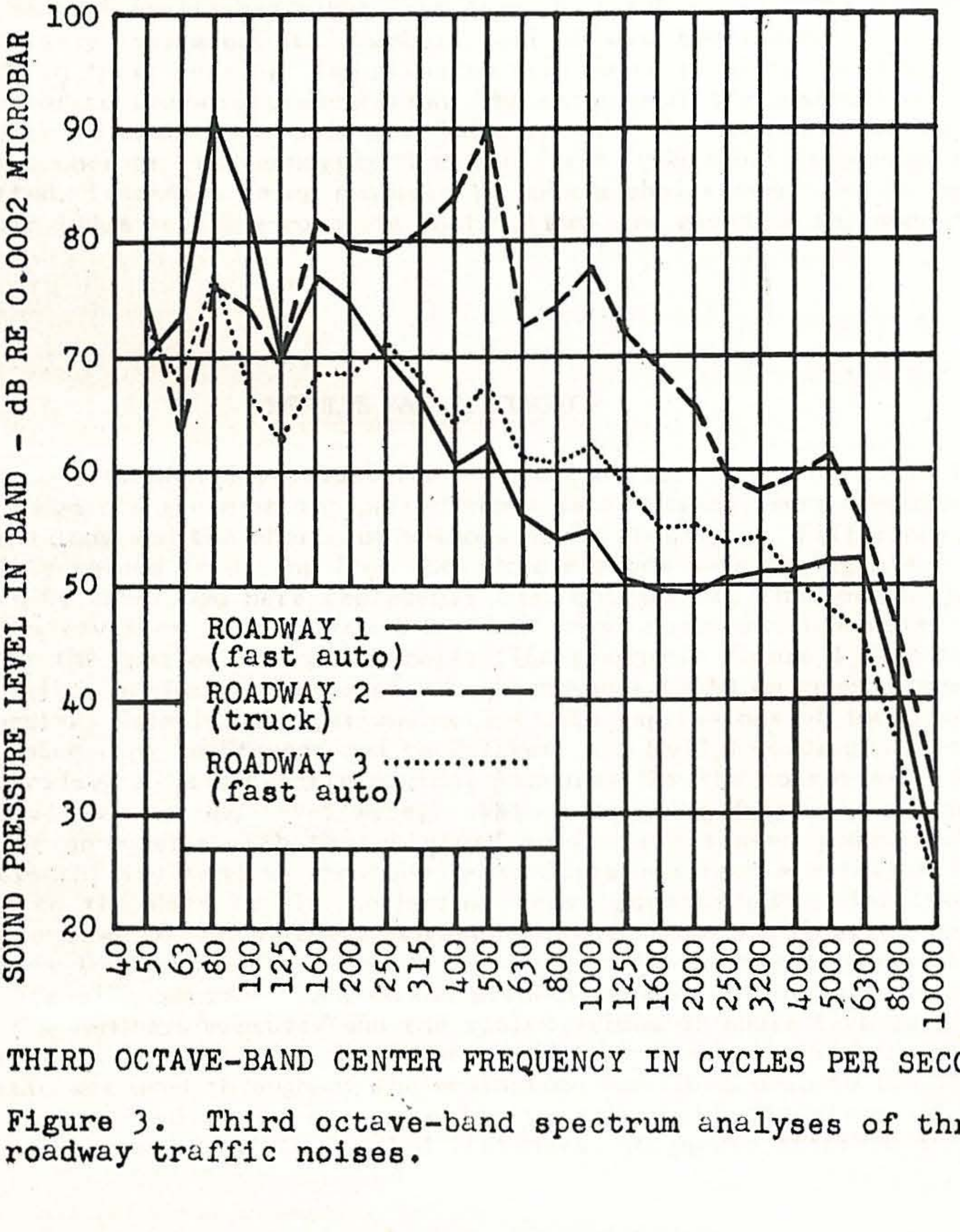

Figure 3. Third octave-band spectrum analyses of three roadway traffic noises.

by a 50 watt power amplifier (Eico, Model HF-50) and supplied to a bank of 6 loudspeakers (four AR-2A and two University horns, Model BLC). One hundred listeners, tested two and three at a time, sat in sound reverberating booths[1] positioned eight feet from the speaker array. The listeners were largely laboratory staff members, all of whom had indicated no known defects in their hearing. Taped instructions were given to the listeners just prior to the noise presentation. The essence of the instructions was to listen to each noise pair and judge whether the second noise in the pair was more or less annoying than the first. No equal judgments were permitted, listeners being required to make a choice even when in doubt of their judgment. The complete instructions are given in the appendix.

RESULTS AND DISCUSSION

Scaled objectionability values for the nine noises, based upon the listener judgments given in the paired-comparison testing, were computed by both the long and the short-cut methods of Guilford (3). Differences in the scale values resulting from these two methods were negligible. The data to be described here represents that obtained by the long method. Such scale values ranged from 0.00, the least objectionable noise, to 4.08 for the most objectionable noise. The graphs in Figure 4 plot these scale values against measures of noise-annoyance based on conversions of the spectral data for the various noises into expressions of loudness in phons, according to Stevens and to Zwicker, and PNdB according to Kryter. The specified A- and C-scale decibel measures for the noises were read directly from a sound level meter. Values relating Zwicker's measures of noise-annoyance with the obtained subjective scaled judgments of objectionability tend to show the least deviation from a straight line fitted to the data by the method of least squares (3). This would indicate that of the noise-annoyance indices under evaluation here, Zwicker's loudness measures most closely corresponded with subjective judgments of annoyance. A Pearson product-moment correlation between Zwicker's loudness measures and the scaled values of subjective judgment yielded an r of .959. It should be noted that the Pearson correlation statistic was used throughout the evaluation described, despite the small number of paired measures, in order to take full advantage of the interval properties of the scaled listeners' judgments obtained in the

[1]The booths were plywood enclosures, 30" x 31" x 79", with a slanted roof and completely opened toward the speakers.

paired-comparison testing. A set of Spearman rank-order correlations were also computed in all cases and gave results comparable to those obtained with the Pearson r, but slightly higher for the loudness and PNdB and slightly lower for the A- and C-scale measures. The Spearman rho values are indicated in parentheses for each relationship in Figure 4.

As expected, values relating C-scale overall energy measurements to the scaled values of objectionability judgments show the greatest deviations from the least squares fitted regression line, thereby indicating the poorest degree of correspondence of all of the relationships noted. The computed Pearson r value utilizing C-scale determinations of noise-annoyance was .753. The Pearson r's obtained between the other proposed measures of annoyance and objectionability judgments were .910 for Stevens' loudness measures in phons, .896 for Kryter's PNdB measures and .831 for A-scale measure in dB. All of the obtained correlations, including that with the C-scale overall energy values, were significantly different from zero ($p < .05$, 7df). According to Guilford, however, a correlation based upon a small number of samples can be only regarded as significant if the correlation value is .900 or higher. Therefore, only the Zwicker and Stevens loudness measures,[2] yielding correlations of .959 and .910 respectively with the subjective annoyance ratings, can be regarded as being truly significant.

More detailed inspection of the data indicated several instances when the correspondence between the proposed measures of annoyance and the listener's annoyance judgments was lacking. For example, two of the roadway noise samples had identical Zwicker loudness values. Hence, when directly compared to each other on the basis of annoyance, it would be predicted that judgments of the more annoying noise would be evenly split between the two samples. The results obtained for such comparisons indicated, however, that one of the two noises was judged more annoying than the other 80 per cent of the time. In another instance, all of the proposed measures of noise-annoyance indicated a flyover noise sample to be more objectionable than a particular train noise sample. When these two noises were directly compared against one another, over 60 per cent of the listener group judged the train noise to be more annoying than the aircraft flyover noise. This was a reversal of the expected outcome in view of the noise-annoyance measures calculated by the various procedures.

---

[2]Kryter's PNdB measure yielded a Spearman rho value of .917 which would also be considered to be truly significant.

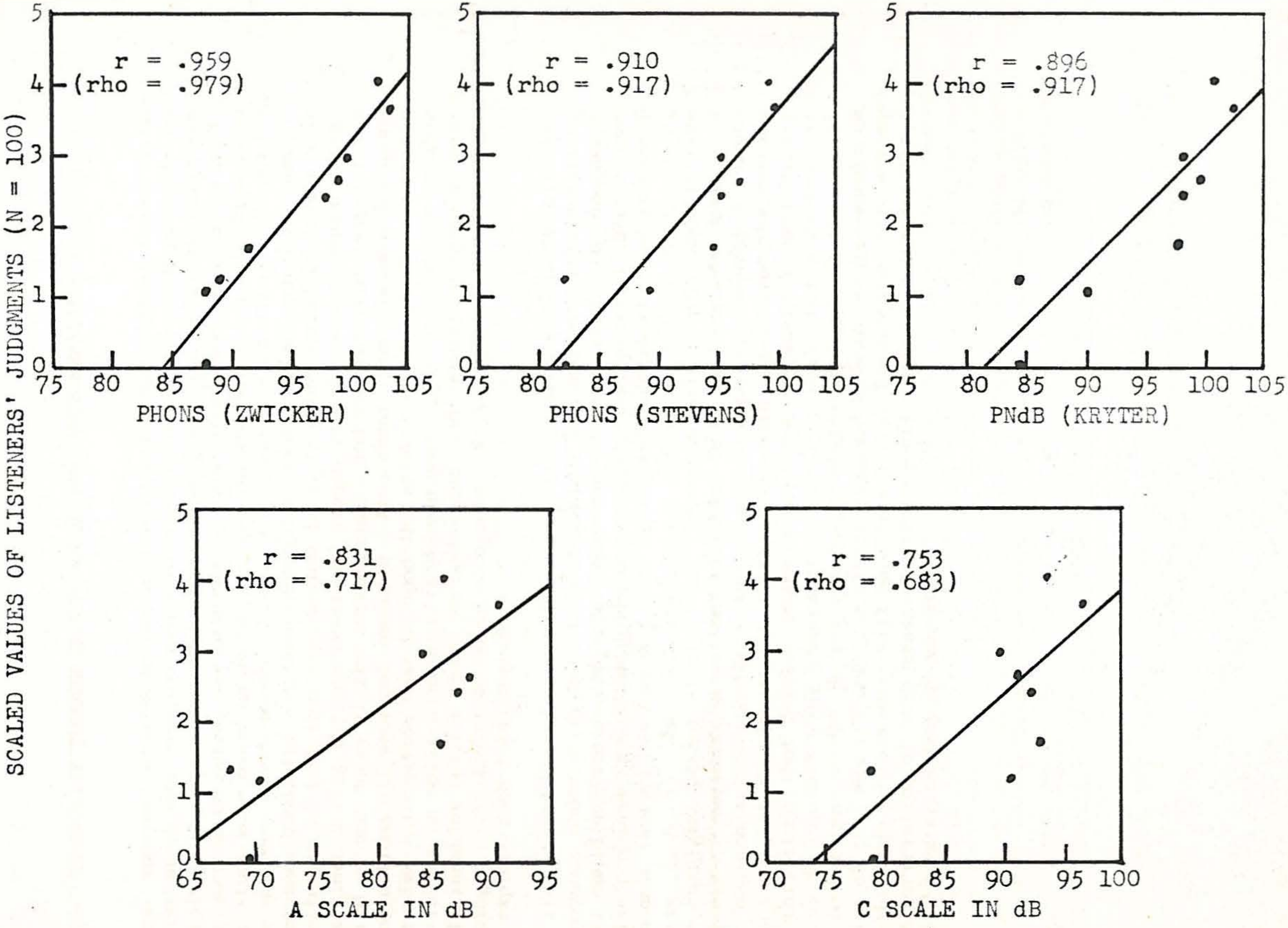

Figure 4. Scaled values of listener objectionability judgments plotted against measures of Loudness (Stevens and Zwicker), Perceived Noisiness measures (Kryter), and A- and C-scale readings for the nine noises presented.

OBJECTIONABILITY JUDGMENTS AND NOISE-ANNOYANCE MEASURES

These apparent discrepancies in the relationship between measures of noise-annoyance and judgments of objectionability could probably be resolved if more acoustical information about the noise stimulus was included in computing the proposed measures. Such measures, in the present study were based on noise spectrum information which only considered the maximum sound pressure level readings developed in third-octave frequency bands for the twenty-second presentation of each noise sample. The length of time that the maximum level or other specified levels were sustained during the twenty-second exposures, the rate of sound energy build-up and decay in the noise sample, and the rate and range of its fluctuations all were not included. At present, there is need for quantitative information bearing upon the effects of any of these acoustic variables on noise-annoyance, and even more important, on the trading relationships between these variables and still others in terms of producing a given amount of noise-annoyance. Kryter (5) has recently begun to explore such relationships.

It also should be emphasized that none of the proposed measures of noise-annoyance under study here take account of non-auditory considerations that could influence annoyance judgments. Indeed, studies of community noise problems have shown quite dramatically that personal, social, and economic factors may condition the residents' reactions to a neighborhood noise (2, 6, 8). For this reason, any purposeful method for measuring the complaint potential of a wide variety of noises in an equally wide variety of community situations will have to be based on both acoustical and non-acoustical considerations. In limiting noise-annoyance evaluations to just the acoustical factors, it would appear from the results of this study that Zwicker loudness measures can serve as an adequate means for quantifying annoyance levels, at least for a variety of vehicular-type noises. A more conclusive statement regarding the superiority of this measure over the others proposed to measure noise-annoyance in this context, however, will have to wait for further studies involving the correlation of such measures with objectionability judgments for a still greater variety of sounds.

REFERENCES

1. Anon. Objective Limits for Motor Vehicle Noise. Report No. 824. Cambridge, Mass.: Bolt, Beranek & Newman, Inc., 1962.

2. Borsky, P. N. Some of the Human Factors Underlying Community Reactions to Air Force Noise. Paper presented at Sixth Annual Meeting, Armed Forces--National Res. Counc. Comm. on Hearing and Bio-Acoustics, Wash., D. C., 1958.

3. Guilford, J. P. Psychometric Methods. New York: McGraw-Hill, 1936.

4. Kryter, K. D. The Meaning and Measurement of Perceived Noise Level. Noise Control, 6, No. 5, 1960.

5. Kryter, K. D. & Pearsons, K. S. Some Effects of Spectral Content and Duration of Perceived Noise Level. J. Acoust. Soc. Amer., 35, (6), 866-883, 1963.

6. Parrack, H. O. Community Reaction to Noise. In C. M. Harris (Ed.). Handbook of Noise Control. New York: McGraw-Hill, Ch. 36, 1957.

7. Peterson, A. P. G. & Gross, E. E., Jr. Handbook of Noise Measurement. (5th Ed.) New York: General Radio Co., 1963.

8. Preliminary Survey Predicted Changes in Neighborhood Noise in West Orange, New Jersey. Lewis S. Goodfriend & Associates. Little Falls, New Jersey, Aug. 1963.

9. Robinson, D. W. Subjective Scales and Meter Readings. Paper presented at a conference held at the National Physical Laboratory, 1961. Cited in The Control of Noise. London: Her Majesty's Stationery Office, 1962.

10. Spieth, W. Annoyance Threshold Judgments of Bands of Noise. J. Acoust. Soc. Amer., 28, No. 5, 872-877, 1956.

11. Stevens, S. S. Procedure for Calculation Loudness: Mark VI. J. Acoust. Soc. Amer., 33, 1577-1685, 1961.

12. Zwicker, E., Flottrop, G. & Stevens, S. S. The Critical Bandwidth in Loudness Summation. J. Acoust. Soc. Amer., 29, 548, 1957. Cited from Bruel & Kjaer, Technical Review, B. & K. Instruments, Inc., 1962.

APPENDIX

APPENDIX

INSTRUCTIONS

The purpose of the study is to determine the relative acceptability of different community noises and sounds. You will hear, on the recording to follow, samples of community noise created by aircraft flyovers, by trains and by roadway traffic. These various noises will be presented to you in a paired-comparison fashion. That is, one community noise sample will be followed immediately by a second noise sample. You are to judge which of the two noise samples would be more disturbing to you if you heard them outside or inside your home many times during the day. If you think the second of the two noises presented would be more disturbing to you, put a plus on your answer sheet after the number announced for each pair of sounds. If you think that the second of the two noises would be less disturbing, put a minus after the proper number. If you think the two noises would be equally disturbing, please make a choice anyway even though you feel you are guessing.

Remember, your job is to judge whether the second noise in the pair is more or less disturbing than the first noise in that pair. Base your answers on how the noises affect you. There are no right or wrong answers and it is important that we find out how people differ, if they do, in their judgments of these noises. It does not matter whether your answers agree or disagree with others taking the test, as long as you make the best judgment you can for each pair of noises.

Again, record a plus if the second of the two noises would be more disturbing to you than the first noise of that pair and a minus if the second of the two noises would be less disturbing than the first noise of that pair. In case of doubt, make the best guess you can. Make a judgment for all pairs.

Also, please write on the appropriate spaces provided on your answer sheet, the date, your age, and your sex. You need not record your name. All answer sheets will remain unidentified with respect to the names of the persons taking this test.