A STUDY OF THE INFLUENCES OF PHRASING, SCALING AND LOCATION OF QUESTIONS CONCERNING EMPLOYEES' DISCOMFORT RATINGS DUE TO NOISE

Olavi Manninen

The Academy of Finland, c/o Department of Public Health, Medical Faculty, University of Tampere, Lääkärinkatu 3, 33520 Tampere 52, Finland

REPORT
Tampere, Finland 1979
Research program on combined effects of environmental factors

Report was presented in Dresden at the International Symposium on The Protection of Workers against Noise, 27-30 November 1979 (organizers Central Institute of Occupational Safety and International Labour Office) and has also been published in the Proceedings of Symposium, pp 205-218
ABSTRACT

A number of methodological difficulties have been thought to occur in conjunction with discomfort or annoyance surveys. For environmental control purposes it is necessary to know the importance of different rating scale parameters for the results obtained. Taking into consideration the criticism levelled in this connection the present investigation was designed to check the magnitude of problems of this kind.

The material consisted of 60 employees in an engineering work. The sample consisted of representatives from six different stages of production (foundry, plate work, machine work, assembly, storage, and office) and types of exposure. Employees answered rating question to estimate noise and other work environmental factors at their individual work places. Altogether 30 different ratings were presented in an A4-size booklet, each page bearing one rating question formulated differently from the others. The calibration of discomfort was performed by putting two reference anchors among the first ratings. The environmental noise level was tape-recorded by setting a microphone in position at the subject's head level at the same time the rating procedure was performed. A single experimental event took about 15 minutes, and all the measurements were repeated in the same random order on week later. Discomfort ratings due to different environmental factors proved satisfactory consistent for the computed coefficients of correlation between the repeated measurements ranged from 0.66-0.68 (draught) to 0.87-0.91 (noise). Overall coefficients between individual ratings and calculated equivalent noise levels ranged from 0.71 to 0.85. The highest coefficient was obtained through a calibrated scale. A five-step scale with additive verbal indications of degrees of discomfort at each square demonstrated the next highest (0.82) association with the results of instrumental measurements of noise. The scale was presented among a check-list including rating questions on the discomfort evoked by nine other environmental factors. The lowest coefficient was obtained when an un-sectioned rating line with only verbal scale endings was used. In general, the influence of the type of scale ending, line, location or even order of ratings on the results was smaller than initially expected. When the mean standard score values of different scales were compared similarities of distribution were especially apparent. It was further found that age, hearing ability, time of exposure, or other environmental factors are of little consequence for the variations in noise ratings; these factors constitute an obvious verification of the above.
INTRODUCTION

It is apparent from present knowledge that noise provokes many detrimental changes in the human organism. A vast amount of psychological literature exists in this particular field. Briefly, noise plays an important negative role in human comfort and well-being (e.g., Schultz 1972, Miller 1974, Chalupnik 1977; May 1978).

Although a variety of techniques have been used to predict the extent of annoyance reactions, social surveys and subjective ratings are undoubtedly the most typical tools of environmental research (Lindvall and Kadford 1973). Due to procedural simplicity of obtaining quantitative measures of environmental features, these techniques have extra relevance with regard to everyday labor protection. Yet this kind of subjective measuring technique is said to cause pitfalls, some of which might affect subjective estimates. Such argumentation is especially maintained by certain statements on obvious effects of rating scale parameters, question style or e.g., location of questions within a questionnaire on rate behavior. Furthermore, some authors have pointed out the importance of the response criteria problem in annoyance surveys (Berglund et al. 1975). According to Poulton (1976), a major disadvantage of this kind is derived from the susceptibility of quantitative subjective assessments to range effects. Perhaps part of the problem here is the meaning of descriptive words given by the different groups of subjects.

Specifically, this study deals with problems related to the phrasing, scaling and location of questions in annoyance surveys. Taking into account the criticisms leveled in this connection, the present investigation was carried out to check the magnitude of problems of this kind. The questions are being considered along with the data collected in an engineering plant.
MATERIAL AND METHODS

The field study was made during the summer of 1978. The material comprised 60 employees in an engineering plant in the city of Tampere, representing six stages of production (foundry, plate work, machine work, assembly, storage, and office) and different types of exposure. Ages ranged from 18 to 62. To develop rating scales and to clarify the techniques of study an exploratory survey was performed among workers in a district central heating plant in the same town. Ratings, including the final large-scale study, were based on these pretested formats.

The rating tasks were printed on one side of each page. They formed an A4-size booklet consisting of three major sections. The first section dealt with background data such as age, occupation, exposure time, hearing ability, job satisfaction, mentally or physically heavy work, wearing of personal ear protectors. The second section contained two reference anchors to achieve a common criterion for estimations, and the third section contained 30 different rating tasks to estimate the degree of annoyance evoked by work environmental factors. The first page was used to provide a brief introduction to the research. Through three extra red interleaves employees were reminded to think carefully about the rating scale and to make the assessment in terms of this scale. Before commencing the rating procedure the study researcher made sure that each employee understood his task.

For reasons of economy, ratings were asked for mainly about environmental noise. During the period of each subjective rating, noise measurements were made with a tape recorder (Nagra TV SJ) over the measuring periods of a full day at each selected personal working place. Afterwards, the availability of a NOVA laboratory computer prompted the techniques used for analysis of the noise measurement data. Analysis of these recordings provided values of equivalent noise level (Leq). A microphone was set up in position at the subjects’ head level. A single experimental event took about 15 minutes, and all the measurements were repeated in the same way and in random order one week later.
A researcher handed the rating booklet to the employee and was present continuously at the employee's personal work place. Immediately after the completion of each rating task the researcher collected the rating format. Only fully completed rating sheets were accepted. Of the employees who took part in the study, half made their ratings in random order whilst the other half did them in a planned order.

Prior to the survey, a tour was made of the entire factory base to locate the employees' personal work places. On the basis of this tour, the sample was divided into groups located in six different sub-areas, each of which included 10 employees. Only one employee from each sub-area was part of the measuring timetable of a day. In administering the survey, the researcher entered the employee's personal work place in randomized order on each measuring day.

Standardization of primary score values were made for data treatment. In addition to computed mean values, correlation and regression analyses were performed between the results obtained from employees' ratings and instrumental measurements. Reliability coefficients were expressed in terms of the Pearson product - moment correlation between repeated measurements.

RESULTS

In the study of the validity of ratings the criterion was the degree of correspondence between rating results and the equivalent noise levels. The most interesting rating scales used in this study are shown in Table 1. The table also summarizes the values of correlation coefficients involved.

The results are of the type expected from consideration of previous work. Assessments of the degrees of discomfort obtained through rating scales R1 and R7 corresponded best to the equivalent noise levels and were quite superior to all other rating questions. Comparing these figures to earlier findings from a survey in metal industry (Manninen 1977a), for example, the individual correlation coefficients increased from 0.67 to 0.82 - 0.85, i.e. about three-quarters of the variance in
TABLE 1. Correlation coefficients (r) between rating scale values and simultaneously determined equivalent noise levels (Leq) at identical work places. For presentation only some of the most interesting types of rating scales have been tabulated. The letter R and sequential figures have been used as identification symbols of rating scales (n=60).

<table>
<thead>
<tr>
<th>Type of rating scale</th>
<th>r</th>
<th>Type of rating scale</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>(R1) First reference rating</td>
<td></td>
<td>(R16)</td>
<td></td>
</tr>
<tr>
<td>&quot;If noise caused by a heavy lorry going uphill is ascribed a value of five and noise in the morning around the main gate of the factory is ascribed the value of one, which of the values 1-5 would you give to characterize the present level of noise that you are experiencing at your personal workplace?&quot;</td>
<td>.85</td>
<td>&quot;Is noise causing you discomfort at your work place?&quot; (Please put a cross (x) on the scale line below. Give as accurate an assessment as possible.)</td>
<td>Very low</td>
</tr>
<tr>
<td>(R2) Second reference rating</td>
<td></td>
<td>(R17)</td>
<td></td>
</tr>
<tr>
<td>&quot;If noise level in a well-soundproofed room is ascribed a value of zero and the average noise level caused by a pneumatic hammer in operation is ascribed a value of five, what value would you give to the present level of noise you are experiencing at your personal workplace?&quot;</td>
<td>.75</td>
<td>&quot;Is noise causing you discomfort at your work place?&quot; (Please put a cross (x) in the scale square below. Give as accurate an assessment as possible.)</td>
<td>Very low</td>
</tr>
<tr>
<td>(R22) &quot;Are some of the following factors causing you discomfort at your work place?&quot; (Please put a cross (x) against each environmental factor.)</td>
<td></td>
<td>(R27)</td>
<td></td>
</tr>
<tr>
<td>Lighting, Vibration, Dust, Solvents, Noise</td>
<td>.82</td>
<td>&quot;How much there are following factors at your work place?&quot; (Please put a cross (x) against each environmental factor.)</td>
<td>Not at all</td>
</tr>
<tr>
<td>Draught, Smell, Temperature, Humidity, Dry air</td>
<td></td>
<td>(R28)</td>
<td></td>
</tr>
<tr>
<td>&quot;Is noise causing you discomfort at your work place?&quot; (Please put a cross (x) on the scale line below. Give as accurate an assessment as possible.)</td>
<td>.78</td>
<td>&quot;How much there are noise at your work place?&quot; (Please put a cross (x) in the scale square below. Give as accurate an assessment as possible.)</td>
<td>Not at all</td>
</tr>
</tbody>
</table>

Each line was 10 cm long. Unsectioned line was plain whilst sectioned lines were divided into five equal sections. Only the question style was differently formulated in both check-lists.

Note: The correlation coefficients (r) represent the strength of the linear relationship between the rating scale values and the simultaneously determined equivalent noise levels (Leq) at identical workplaces. Higher values of r indicate a stronger correlation.
individual responses could be attributed to physical noise exposure as measured here. This includes the improved correlation between annoyance and physical measurement when the former was made using a calibrated technique. On the other hand, it can be shown that rating scales formulated differently from each other are not absolutely alternative tools of the annoyance measure.

The purpose of calibration questions (R1, R2) was to offer the employee a familiar and common reference frame of ratings. To characterize the superiority of either one of the two reference anchors, means of standardized score values were first computed by environmental noise strata. These results are shown in Figure 1. The mean ratings for these parallel forms fall at nearly the same position along the equivalent noise level. Nevertheless, mean score values obtained through the first reference ratings seem to have a somewhat higher association between noise levels of $L_{eq}$ - especially in the range of under 60 and over 90 $L_{eq}$ - than the second reference rating.

![Comparison of the reference ratings (R1, R2) and corresponding equivalent noise levels ($L_{eq}$) according to work environmental noise categories. Plots describe arithmetic means of standardized scores.](image-url)
Even to the extent that employees' responses are directly related to the environmental stimulus, there is the question, which of the possible direct effects is of most concern: for example, rapidly increasing noise of a repetitive, impulsive character might produce a greater subjective response than the same level of steady noise. In this connection, it was enlightening, therefore, to explore the presence of any marked differences between personal ratings and different types of exposure. For the description of the findings, arithmetical means of the scores have been plotted according to the production stage. As Figure 2A demonstrates, there was good association between the five-step scales presented among check-lists (R7, R27), the first reference rating (R1) and noise level at each production stage. The biggest differences were revealed at the second production stage (i.e. plate work; see Fig. 2B) when employees had made their ratings by five-step scales laid out singly on the page (R15, R17).

![Figure 2. Comparison of rating results and equivalent noise levels (Leq) by stage of production. Ratings were made using scales formulated differently from each other (for further information, see Table 1). Plots describe arithmetic means of standard scores. The sample was grouped into categories which represent six stages of production: foundry (1), plate work (2), machine work (3), assembly (4), storage (5), and office (6).](image-url)
In each case, however, all the scales produced fairly similar profiles of mean standardized values. On the basis of this study the employees did not seem inclined to exaggerate their assessment, at least due to the impulsive character of noise.

Due to the function of the research being undertaken certain typical background variables were included among the first questions. A further function of the questions was to introduce employees to rating procedure in order to clarify the influence of these variables on employees' rating results. Consequently, questions about, e.g., employees' age, exposure time, hearing ability or regularity of wearing ear protectors, were presented at the beginning of the booklet. The means of standard score values of the ratings and instrumental noise measurements are presented in Figure 3.

![Diagram](image-url)

**Fig. 3.** Comparison of rating results and equivalent noise level (Leq) with regard to employees' age, hearing ability, exposure time and wearing of ear protectors. Plots describe arithmetic means of standard scores in each category.
As can be seen, a certain similarity of distribution of mean values occurs in keeping with the categories of the "background variables". There is little doubt, therefore, that the contribution of this set of variables to the variance in annoyance is more important than that for noise exposure itself.

In order to clarify the influence of other environmental factors or the order of ratings on the rating results, the data was treated according to these aspects. The results may be best illustrated by reference to Figures 4 and 5. As a generalization from the findings, one can state that noise ratings are quite independent, and no appreciable difference was found between the ratings made in planned or random rating order.

Fig 4. Comparison of rating results and equivalent noise levels (Leq) according to an computed environmental index of hazards.

Fig 5. Comparison of rating results and equivalent noise levels (Leq) when ratings were made in considered and random order.
Although a high correlation coefficient is necessary for a satisfactory dose - response relationship, it is in itself insufficient for prediction of the response. Subsequently, authorities ought to have a suitable criterion of acceptability to determine the noise exposure that should not be exceeded if excessive probability of discomfort is to be avoided.

Figure 6 shows the regression equations obtained at environmental noise levels for two rating questions. The slope of rating scale R1 is somewhat steeper than that of rating scale R7. This comparison gives a variance of estimations which, translated into equivalent range of noise exposure, has a maximal fluctuation of the order of 5 dB between the mean rating results. For example, when the expression "very high" (5) on the rating scale (R7) is equivalent to a mean equivalent noise level of 95, the corresponding environmental noise level is about 90 when assessed through the severest expression "noise caused by a heavy lorry going uphill" (5) on the first reference rating scale (R1).
From a practical point of view and for corrective purposes or transformations from one scale to another, the established association of this kind is obviously meaningful.

In order for employees to become accurate and reliable human meters of their work environment, it is necessary for them to achieve reasonable repetitive accuracy in making the discomfort ratings. Double assessments are an important part of this procedure. Therefore all the measurements were made in a systematic way in the same random order one week later. Table 2 records these individual ratings. It can be seen that the information obtained using this rating procedure is reasonably stable. In most cases the results of discomfort ratings seemed to be a little more repeatable than the results of magnitude ratings of the same work environmental factors. For comparison, the correlation coefficients between the repeated measurements made by the reference ratings were 0.76 (R1) and 0.85 (R2).

### Table 2. Pearson product-moment correlation coefficients (r) calculated between two measurements made by check-lists including discomfort (R1) and magnitude (R27) estimation of work environmental noise.

<table>
<thead>
<tr>
<th>Environmental factors</th>
<th>( r ) (discomfort)</th>
<th>( r ) (magnitude)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lighting</td>
<td>.70</td>
<td>.63</td>
</tr>
<tr>
<td>Vibration</td>
<td>.75</td>
<td>.67</td>
</tr>
<tr>
<td>Dusts</td>
<td>.82</td>
<td>.73</td>
</tr>
<tr>
<td>Solvents</td>
<td>.81</td>
<td>.82</td>
</tr>
<tr>
<td>Noise</td>
<td>.91 (R1)</td>
<td>.87 (R27)</td>
</tr>
<tr>
<td>Draught</td>
<td>.66</td>
<td>.68</td>
</tr>
<tr>
<td>Smell</td>
<td>.80</td>
<td>.69</td>
</tr>
<tr>
<td>Temperature</td>
<td>.73</td>
<td>.70</td>
</tr>
<tr>
<td>Humidity</td>
<td>.72</td>
<td>.75</td>
</tr>
<tr>
<td>Dry air</td>
<td>.85</td>
<td>.82</td>
</tr>
</tbody>
</table>
DISCUSSION

The study was designed to throw light on the importance of different rating scale parameters for the results of annoyance surveys. The material was such that a relatively thorough analysis could be made. The procedure itself was based on an analysis of systematically conducted instrumental measurements and simultaneous personal ratings as described and evaluated above. Findings suggest that except for slight differences between results obtained by different rating scales, the influence of the rating line, location, order or wording of ratings on the assessments was smaller than initially expected. It was further found that "background variables" did not determine the variation of results of the discomfort ratings. These findings, of course, are not unique, but are encouraging for further studies. If one utilizes the same rating scales and similar subjects, one will obviously get a reasonable correlation between ratings and noise level for a variety of noise sources. First of all, a decisive prerequisite is the timing of evaluations. When comparisons are to be made, instrumental measurements and employees' personal ratings should be conducted in a fully comparable way with respect to the time dimension. Accordingly, in comparing present results to this author's previous findings, the individual correlation coefficient involved increased from 0.67 to an excellent 0.82-0.85; by training ratee an association of this kind can be increased up to 0.90 at least (see Hanninen 1977a).

Illustrative of all efforts at noise control measures is the need to relate personal response to the noise characteristics of that environment. This is undoubtedly an area of research from which most interesting and useful information may be gained, but it is also an area which is complicated by many factors. From time to time, in the planning stages of large-scale surveys, the particular question has arisen as to which of many available noise ratings does, in fact, offer the best tool for subjective assessment. It is in the choice of suitable scales that a major difficulty causes.
Firstly, the design of a questionnaire or rating format is very important. In a comparative study of magnitude and annoyance estimations of noise, there was a tendency for the verbal expression "causing you discomfort" in the formulation of the question to give slightly more precise and more consistent results than the questions where the expression "how much" was used. The disadvantages of the sectioned or un-sectioned lines were also evident in annoyance assessments in the field, although lines did not affect subsequent ratings in a significant way. The use of five response categories really did permit a high level of accuracy, particularly when a five-step scale is provided with additional verbal indications of degrees of discomfort ("not at all" - "very high"). Reliability seemed to improve when the rating scale was presented in a check-list comprising a number of other annoying environmental factors.

Discussion here is limited to a summary of the results in assessing the annoyance caused by industrial-type noise. Support for the attempt to broaden the concept of validity of rating techniques to the field of other environmental factors is given by the following results of ratings of discomfort due to odor and dust. As Figure 7 shows, the mean degree of annoyance among metal industry workers correlated very well with mean ratings of an external panel of raters. In this connection it is worthwhile mentioning that certain types of atmospheric gases and particulates will have their first impact through complaints of workers in workshops. Merely from practical standpoint, the function of the sense of smell has an important place with regard to the prevention of accidents (e.g. Maus 1975). As further indication of the validity of personal assessments, Osborne and Clarke (1974) obtained product-moment correlation coefficient of up to 0.99 between ratings and intensity of vibration stimuli.
Fig. 7. Comparison of discomfort ratings due to odor and dust by stage of production. The arithmetic means of odor and dust are presented in terms of categories which represent the nature and stage-wise progress in the production of metal product, and the production stages of management functions based on the line organization (see further information from Hanninen 1977a, b).

Recent findings display striking evidence for a proper tool of annoyance scaling based on self-rating questions. However, efforts to refine these measurements will and should continue. Basically, certain common reference anchors are a prerequisite for work of this kind. As previously demonstrated, offering the employees familiar and common reference frames of ratings provided data with good sensitivity and value scaling. For obvious reasons, further research is needed in order to recommend the fittest reference frame for multipurpose annoyance survey.
ACKNOWLEDGEMENTS

The research was supported by a grant from the Medical Commission of the Finnish Academy. The author wishes to thank Mr. Aimo Tikka and Mr. Matti Pasanen for their valuable help with this study.

REFERENCES

Berglund B, Berglund U and Lindvall T (1975)

Chalupnik JD (1977)
Transportation noises. A symposium on acceptability criteria. Ann Arbor Science Publishers, Michigan, USA.

Lindvall T and Radford EP (1973)

Männinen O (1977a)
Ratings of discomfort due to noise and inadequate illumination at work. Reports of Public Health Sciences H11. Department of Public Health, University of Tampere, Tampere, Finland.

Männinen O (1977b)
Environmental factors and employees discomfort in three machine industry plants. Reports of Public Health Sciences H12. Department of Public Health, University of Tampere, Tampere, Finland.

May DF (1978)

Miller JD (1974)

Neus A (1975)

Oborne DJ and Clarke MJ (1974)
The determination of equal comfort zones for whole-body vibration. Ergonomics 17:769-782.

Poulton EC (1976)

Schulte TJ (1972)