Journal of Environmental Psychology (1999) **19**, 295–302
0272-4944/99/030295+08 $30.00/0

Article No. jevp.1999.0133, available online at http://www.idealibrary.com on IDEAL®

Journal of ENVIRONMENTAL PSYCHOLOGY

PSYCHOMETRIC PROPERTIES OF FOUR MEASURES OF NOISE SENSITIVITY: A COMPARISON

KARIN ZIMMER AND WOLFGANG ELLERMEIER

Institut für Psychologie, Universität Regensburg, 93040 Regensburg, Germany

Abstract

Individual noise sensitivity is a stable personality trait covering attitudes towards a wide range of environmental sounds. It is a major antecendent of noise annoyance reactions, and is assessed by obtaining responses to one or several rating-scale items. The psychometric properties of four German-language noise-sensitivity measures—a translation of Weinstein's (1978) noise-sensitivity scale, a newly developed questionnaire, and two single-item questions reflecting susceptibility to *sounds* and *noise*, respectively—were evaluated, using a student sample of $n = 213$ persons. Reliability coefficients ranged from $r = 0{\cdot}70$ for the rating of susceptibility to sounds to $r = 0{\cdot}92$ for the newly constructed questionnaire. Construct validity was appraised by inter-correlating noise-sensitivity scores, and by relating noise-sensitivity scores to questionnaire measures of depression, stress, anger, and anxiety. The results indicate that, while the questionnaire measures satisfy established criteria for test evaluation, the one-item ratings do not. Further exploratory analyses on a subset of the sample found only weak relationships between self-report measures of noise sensitivity and objective performance decrements under noise.

Introduction

Individual annoyance reactions to noise have been found to depend on physical attributes of the noise, attitudes towards the noise source, and personal characteristics of respondents (e.g. Taylor, 1984; Job, 1988; Green & Fidell, 1991). According to a review of 27 studies (Job, 1988), noise exposure is the strongest determinant, accounting for an average of 17·6 per cent of variation in individual annoyance reactions. To explain the still considerable difference in noise tolerance across individuals when noise exposure is controlled for, the concept of *noise sensitivity* has been invoked. For a recent discussion of its definition and measurement, see Job (in press). Even though this concept is not unanimously defined (see Öhrström *et al.*, 1988; Lercher, 1996; Staples, 1996), its operationalization in the research literature characterizes it as a stable personality trait that captures attitudes towards a wide range of environmental noises (Moreira & Bryan, 1972; Weinstein, 1978; Guski, 1987; Stansfeld, 1992).

In an investigation of three possible groups of antecedents of annoyance by aircraft noise—noise exposure measures, personal background characteristics, and attitudes toward aircraft operation—using a path-modelling approach, Taylor (1984) found noise sensitivity to be the only personal background variable investigated to have a significant effect on annoyance; moreover, noise sensitivity proved to have the strongest single effect (direct and indirect combined) overall. Likewise, Langdon (1976*b*) reported noise sensitivity to have a stronger impact on individual annoyance than noise level and explains an additional 12 per cent in reaction variance. A set of studies investigating the relationship between individual noise sensitivity and annoyance reactions to railway, aircraft and construction noise, with noise level controlled for, found correlations ranging from 0·25 to 0·45, the mean correlation being $r = 0{\cdot}32$, that is explaining 10·24 per cent of the variance [Job, 1988; results of Langdon's (1976*a*) study that Job's review also takes into account are disregarded in the present context, as they pertain to behavioural responses towards noise which, in general, do not contribute to annoyance (Taylor, 1984)].

While these results clearly show noise sensitivity to be a major antecedent of noise annoyance reactions, current measurement of noise sensitivity remains unsatisfactory. With rare exceptions (McKennell, 1969; Weinstein, 1978), the instruments used to measure noise sensitivity have not been subjected to rigorous psychometric evaluation, and thus are of unknown quality. Often, items are arbitrarily picked from existing questionnaires or only a small number of rating-scale items are employed (e.g. Griffiths & Delauzun, 1977; Stansfeld *et al.*, 1985; Kjellberg *et al.*, 1996). While this is advantageous for economical reasons, especially in large surveys combining several questionnaires, it is well known that, as a rule, a questionnaire's brevity adversely affects its psychometric quality. Apparently, many investigators in applied noise research seem to believe that this problem may be alleviated by supplying a larger number of response alternatives. That, of course, is an issue to be settled on empirical grounds. To sum up, the assessment of noise-sensitivity is often based on *ad hoc* measures and fraught with questionnable assumptions, thus calling for more systematic study of the suitability of available instruments.

The present investigation was conducted in order to examine whether four self-report measures of noise sensitivity meet established psychometric criteria with respect to their distribution characteristics, reliability, and validity. The questionnaires evaluated were the most widely utilized and best available instruments: Weinstein's (1978) noise-sensitivity scale, and another, newly constructed questionnaire (Zimmer & Ellermeier, 1998*a*), which is designed to cover a larger variety of respondents' attitudes towards noise than does Weinstein's scale. Both questionnaires deal with a variety of sound sources encountered in everyday life. The single-item rating scales investigated are frequently used in survey research (e.g. Langdon, 1976*b*; Stansfeld *et al.*, 1985; Raw & Griffiths, 1988; Müller-Andritzky *et al.*, 1990; Lopez Barrio & Carles, 1993) and require the subject to make an overall estimate of her/his susceptibility to noise.

Furthermore, an exploratory attempt was made to link subjective noise sensitivity to objectively measured performance decrements under noise conditions similar to those in an open-plan office environment: a subset of our sample participated in an 'irrelevant speech' experiment (see Salamé & Baddeley, 1982; Jones & Morris, 1992), and individual outcomes of the experiment were related to the various noise-sensitivity measures.

Method

After a short written introduction stating the general goals of the present investigation, four German-language measures of noise sensitivity were administered: a newly constructed noise-sensitivity questionnaire (Zimmer & Ellermeier, 1998*a*), a translation of Weinstein's (1978) noise-sensitivity scale, and two one-item self-ratings capturing susceptibility to sounds and susceptibility to noise, respectively. Layout of the first three measures of noise sensitivity was identical, providing verbally labelled Likert-scale responses. For the rating of susceptibility to noise, a numbered category scale, reflecting the degree of agreement, was presented.

Materials

Weinstein's noise sensitivity scale (WNS). Weinstein's (1978) noise sensitivity scale consists of 21 items, most of which express attitudes towards noise in general and emotional reactions to a variety of environmental sounds encountered in the everyday life of students, the target population of the questionnaire (for details on the German version used here see Zimmer & Ellermeier, 1997). For every statement, six response options ranging from strong disagreement to strong agreement are presented. With 14 of the 21 items, agreement to the item indicates greater noise sensitivity of the respondent.

Noise sensitivity questionnaire (LEF). This recently developed questionaire (Zimmer & Ellermeier, 1998*a*) encompasses statements about a wide variety of environmental noises in a range of situations that affect the entire population. The material covers seven content areas: everyday life, recreation, health, sleep, communication, work, and noise in general. The 52 items presented relate to perceptual, cognitive, affective, and behavioural responses towards noise in these contexts. An almost equal number of items is scored in each direction. For every item, respondents may choose one of four response options ranging from strong disagreement to strong agreement.

One-item rating scales. Two direct ratings of individual noise sensitivity were obtained by asking for the degree of the respondent's *susceptibility to noise* (RN) and *susceptibility to sounds* (RS). The latter rating is equivalent to item 21 of Weinstein's noise sensitivity scale and features a range of six response options. The self-rating of susceptibility to noise was requested on an 11-point numerical rating scale, thus providing a finer grain of response alternatives. The endpoints of this scale were labelled with the German equivalents of 'not noise-sensitive at all' and 'very noise-sensitive', respectively.

Procedure

The scales were always presented in the same order: LEF, WNS, RS, RN. Thus, the overall ratings of noise sensitivity were given the potential benefit of prior elaboration through the specific statements made in the questionnaires. The two single-item ratings were sufficiently separated in time, since demographic information, and additional validational questionnaire data (see Results) were obtained between presenting RS and RN.

The demographic information gathered included age, gender, education, number of members of, and presence of children in the participant's household, an inquiry about hearing problems and the use of hearing aids, as well as subjective estimates of the loudness of the respondent's residential area and of his or her noise exposure history. In total, data collection took 25 to 35 minutes for every participant.

Sample

The questionnaires were administered to $n = 213$ university students at Regensburg. The respondents' age ranged from 19 to 44 years, with a mean of 24·2 years (S.D. = 3·6 years). Seventy-four respondents (34·7%) were male. $n = 187$ of the participants also completed questionnaires serving construct validation, $n = 178$ participated in the retest after a 4-week interval. None of the subjects had any prior knowledge of the literature on noise effects, or of the research questions presently addressed.

Psychometric properties

For every item a score was assigned to each response option, so that the higher its numerical value the more noise-sensitive the respondent. Scores for RS and RN ranged from 0 to 5 points, and 0 to 10 points, respectively. The total score of the questionnaire measures was composed of the unweighted sum of their item scores, WNS scores ranging from 0 to 105 points, and LEF scores from 0 to 156 points.

None of the participants reported any difficulty in handling the questionnaires. For each of the noise sensitivity measures, correlations with demographic variables, the distribution of scores, and reliability as well as validity coefficients were assessed.

Distribution of scores

The range of individual scores, mean scores and standard deviations for each of the four noise-sensitivity measures are given in Table 1.

For RS, RN, and LEF, mean scores were near the midpoint of the entire possible range, while the WNS mean scored somewhat higher. Total scores of both questionnaires showed no significant deviation from a Gaussian distribution (Kolmogorov-Smirnov tests; $z = 0{\cdot}572$, N.S., for LEF), whereas scores of the one-item scales were symmetrically (Skewness $s = -0{\cdot}167$, N.S., for RN, and $s = -0{\cdot}068$, N.S., for RS), but not normally distributed ($z = 2{\cdot}255$, $p < 0{\cdot}001$, and $z = 2{\cdot}272$, $p < 0{\cdot}001$ for RN, and RS, respectively). Kurtosis of both one-item measures was negative, indicating a broader, flattened peak compared to the normal distribution ($K_{\mathrm{RN}} = -0{\cdot}998$, $p = 0{\cdot}001$; $K_{\mathrm{RS}} = -0{\cdot}669$, $p = 0{\cdot}022$); the extreme answer categories were rarely chosen.

Reliability

Internal consistency, an index of the degree to which the items of a test measure individual differences in the same way, was very good for the two questionnaires, the longer LEF scale reaching a higher consistency coefficient (Cronbach's $\alpha_{\mathrm{LEF}} = 0{\cdot}922$; $\alpha_{\mathrm{WNS}} = 0{\cdot}860$). Retest-reliabilities, characterizing the stability of responses over time, ranged from medium, for RS, to very high, for the LEF questionnaire (see Table 1). The critical differences (bottom row in Table 1) of the four noise-sensitivity scales represent the minimal difference in individual scores to reach statistical significance at $\alpha = 0{\cdot}05$ (Gulliksen, 1964, p. 22; Lienert & Raatz, 1994, eq. 15·14). They span one-third (RS), and one-fifth (RN) of the maximal range of scores for the

TABLE 1
Distribution and reliability characteristics: range of scores, mean score, standard deviation, retest-reliability, and the critical difference (α = 0·05) for each of the four noise-sensitivity measures (n = 213)

	WNS	LEF	RS	RN
Range	7–94	22–125	0–5	1–10
M	63·075	79·437	2·845	5·751
S.D.	14·703	19·382	1·232	2·114
r_{tt}	0·871	0·911	0·703	0·829
Critical difference	10·39	11·40	1·32	1·75

one-item ratings, thus turning out to be rather crude in comparison to the questionnaires, the critical differences of which amount to one-tenth (WNS), and one-thirteenth (LEF) of the maximal range, respectively.

With reliability coefficients exceeding 0·9, LEF meets the psychometric standards required for personality and achievement tests, and thus may be employed to differentiate between individual respondents. Taken together, these results suggest that all four noise-sensitivity scales measure a homogenous construct stable over time.

Correlations with demographic variables

Previous research found no relationship between noise sensitivity and demographic characteristics other than *age*, with which noise sensitivity increases (e.g. Moreira & Bryan, 1972; Weinstein, 1978; Taylor, 1984; Stansfeld, 1992). The present finding is consistent with these results in that none of the noise-sensitivity measures correlated with any of the demographic data collected. (In order to account for the multiple correlations computed on the data set, the test-wise significance level was Bonferroni-corrected to $\alpha_{adj}=0{\cdot}0102$.) Contrary to previous findings, however, there was also no correlation with age, and while WNS and LEF scores showed a tendency in the expected direction, product–moment correlations failed to reach statistical significance ($r_{WNS}=0{\cdot}130$, $p=0{\cdot}058$, and $r_{LEF}=0{\cdot}163$, $p=0{\cdot}018$, respectively, for the questionnaires; $r_{RS}=0{\cdot}042$, $p=0{\cdot}547$ and $r_{RN}=0{\cdot}085$, $p=0{\cdot}215$ for the one-item scales). The failure to find a significant correlation between age and noise sensitivity may well be due to the limited age range in the present student sample: data collected during LEF-scale development from a representative sample ($n=117$) with a much wider age span (range: 18–83; $M=43{\cdot}1$ years of age) exhibited a highly significant correlation between the sum of scores of the 52 items subsequently selected for the final version and age ($r_{LEF}=0{\cdot}459$, $p<0{\cdot}001$). Correlation of age with the one-item rating RN, on the other hand, remained insignificant ($r_{RN}=0{\cdot}148$, $p=0{\cdot}110$, $\alpha=0{\cdot}0102$) in this sample as well.

Furthermore, the estimated loudness of the respondent's residential area correlated significantly with both of the one-item rating scales (Spearman's $\rho_{RS}=0{\cdot}177$, $p=0{\cdot}010$, and $\rho_{RN}=0{\cdot}191$, $p=0{\cdot}005$), but not with the questionnaires, the correlation with WNS indicating a tendency, but remaining insignificant ($\rho_{WNS}=0{\cdot}161$, $p=0{\cdot}018$, and $\rho_{LEF}=0{\cdot}033$, $p=0{\cdot}629$).

Cues to validity

To decide whether the four noise-sensitivity measures are based on a single construct, a principal-component factor analysis was performed on the total scores of these measures. It confirmed the unidimensional nature of the underlying construct: with an eigenvalue of 3·00, this factor explains 75 per cent of variation. No other factor reaches an eigenvalue larger than 1. Pearson product–moment inter-correlations of the four measures are given in Table 2.

Interestingly, while both the two questionnaire measures, and the two single-item responses, correlate highly with each other, the correlations between the two types of measures are somewhat lower.

As there is no universally accepted performance criterion for noise sensitivity, further analyses focused on the construct-related validation of the four noise-sensitivity measures. Previous studies demonstrated noise sensitivity to correlate with emotional stress (Dornic *et al.*, 1990) and with depression (Öhrström *et al.*, 1988; Stansfeld, 1992). The relationship between noise sensitivity and anxiety as well as anger-expression has also been discussed in the literature, however, it does not present a clear picture (Stansfeld, 1992). In the present study, the discriminative construct-related validities of the four noise-sensitivity measures were obtained by correlating their scores with the following German-language questionnaires: (1) ADS, a depression scale designed for subclinical populations (Hautzinger & Bailer, 1993), (2) STAI, an inventory assessing state and trait anxiety (Laux *et al.*, 1981), (3) the state and trait scales of the STAXI inventory of anger expression (Schwenkmezger *et al.*, 1992), and (4) KFB, 'a daily-hassles scale', measuring stress (Flor, 1991).

Correlations close to zero with the state scales administered, and moderate correlations with depression, stress, and the trait scales were expected;

TABLE 2

Results of convergent validity analyses: product-moment inter-correlations of the four noise-sensitivity measures ($n=213$)

	WNS	RS	RN
WNS	0·790	0·572	0·642
LEF		0·582	0·639
RS			0·717

Note: attached probabilities are below $p<0{\cdot}001$ in all cases. As RS is included in WNS, their correlation coefficient is part-whole corrected.

substantial correlations, on the other hand, would have questionned a noise-sensitivity measure's validity, suggesting it captures a different construct.

As may be seen in Table 3, the predicted pattern of outcomes is most fully met by WNS. As the investigations generating our predictions were based on Swedish and British versions of Weinstein's scale, this result comes as no surprise. In contrast, LEF, RN, and RS deviate from the ideal pattern. Both single-item ratings lack the expected correlation with the trait anger-expression. Furthermore, the RN score correlates significantly with state anxiety. LEF scores only relate to the trait scales; however, they show no significant correlation with either depression or stress.

As regards to the magnitude of the correlations, irrespective of statistical significance, WNS scores exhibit higher correlations overall and LEF scores somewhat lower correlations to the other personality and mood scales, than have been previously reported in the literature. Scores of the one-item ratings correlate more strongly with depression and less with stress, than has been found in previous research (Öhrström *et al.*, 1988; Dornic *et al.*, 1990; Stansfeld, 1992).

Relationship between objective and subjective impairment by noise

How do subjective estimates of noise susceptibility relate to objective performance decrements under noise? To address this question, a subset of $n = 72$ of our sample, in addition to completing the questionnaires, participated in a standard 'irrelevant speech' experiment (for details see Ellermeier & Zimmer, 1997), in which three types of auditory backgrounds were presented via headphones while the subject was performing a serial-recall task: (1) quiet, (2) pink noise at 76 dB (A), and (3) Japanese speech (an arbitrary 15-s segment from a lecture) presented at 76 dB (A). In order to be able to measure individual differences unconfounded with practice effects, these conditions were randomly mixed on a trial-by-trial basis. The subject's task was to report a series of nine digits displayed on a computer monitor at a rate of one per s after a 5-s retention interval while ignoring the auditory input. Whenever the subject failed to report the correct digit in the correct position, an error was scored. The sum of errors in 20 trials computed separately for each auditory condition served as the dependent variable.

Consistent with the literature (e.g. Salamé & Baddeley, 1982), a large and highly significant [$F(2, 142) = 97{\cdot}16$; $p < 0{\cdot}001$] overall 'irrelevant speech effect' was obtained, in that subjects made considerably more errors under speech ($M_S = 74$) than in quiet or with uniform pink noise, with the latter two conditions yielding almost identical results ($M_Q = 50$ and $M_P = 52$, respectively).

For each subject, the difference between errors under speech and errors in the quiet control condition served as a measure of the magnitude of the individual irrelevant-speech effect. Its correlation with the scores of the four noise-sensitivity scales, corrected for the irrelevant-speech effect's low reliability (see Ellermeier & Zimmer, 1997), accounted for only a small portion of the variance—between 5 per cent and less than 1 per cent ($r_{LEF} = 0{\cdot}232$; $r_{WNS} = 0{\cdot}167$; $r_{RS} = 0{\cdot}116$; $r_{RN} = 0{\cdot}084$). Thus, noise sensitivity is only weakly associated with a performance criterion of impaired recall in the 'irrelevant speech' paradigm.

TABLE 3
Results of divergent validity analyses

	WNS	LEF	RS	RN
ADS (depression)	**0·165**	0·112	**0·196**	**0·198**
KFB (stress)	**0·243**	0·119	**0·191**	**0·169**
STAI-state (anxiety)	0·085	0·028	0·083	**0·148**
STAI-trait (anxiety)	**0·247**	**0·144**	**0·289**	**0·247**
STAXI-state (anger)	0·041	0·011	0·050	0·068
STAXI-trait (anger)	**0·241**	**0·153**	0·109	0·125

Note: Pearson product-moment correlations of the four noise-sensitivity measures with a depression scale (ADS), a daily-hassles scale (KFB), state and trait anxiety scales (STAI-state and STAI-trait, respectively), and state and trait anger expression scales (STAXI-state and STAXI-trait, respectively) ($n = 187$). Significant correlations ($\alpha = 0{\cdot}05$) are highlighted.

Discussion

The most important outcome of the present study is that the four measures of noise sensitivity investigated did indeed exhibit systematic differences with respect to psychometric quality and indicators of validity. The pattern of results emerging from this data set shall be discussed with four questions in mind: (1) are full-length noise-sensitivity questionnaires preferable over one-item self ratings? (2) in what way do the two questionnaires investigated differ from each other? (3) what is the relationship between noise sensitivity and objectively measured noise effects? and (4) what is lacking in the prevalent conceptualization of noise sensitivity?

Questionnaires vs one-item responses

The present investigation revealed a number of systematic differences between questionnaires laboriously constructed according to the standards of test theory, and *ad hoc* rating scales often preferred in large surveys. Specifically, except for the obvious advantage of economy, single-item ratings suffered from the following shortcomings:

(1) Precision turned out to be crude, as reflected in critical differences covering as much as one-third (RS) or one-fifth (RN) of the possible range of scores.
(2) Retest reliability was lower than that of the questionnaires, which approximated or exceeded 0·90.
(3) The single-item ratings failed to capture the increasing noise sensitivity with age while showing unwanted correlations with the respondent's (estimated) noise exposure. In the prevalent conceptualization, noise exposure might affect annoyance reactions, but should not modify the pre-existing trait of noise sensitivity (see Taylor, 1984).
(4) The one-item scales did somewhat worse in matching the pattern of correlations expected with related psychological concepts (discriminative validity). Namely, they did not capture the anger component inherent in increased noise sensitivity.

The evidence thus suggests that the single-item scales investigated do not meet established psychometric criteria. Furthermore, differences in the layout of the ratings presented to subjects do not seem to matter as much as intuition would suggest. In the present investigation, the finely gained numerical scale did only marginally better than a cruder set of verbal response options. Note that, if anything, the present investigation *over-estimates* the psychometric quality of the single-item scales, since their presentation after 20 to 50 questionnaire items elaborating the subject most likely serves to increase the validity of a global self-rating regarding noise-sensitivity.

Differences between the two questionnaires

Reliability coefficients for both questionnaire measures were very good, the longer LEF scale achieving slightly better results. The expected pattern of discriminative validity was better matched by the questionnaire measures than by the one-item ratings, with WNS's coefficients somewhat higher, and LEF's coefficients somewhat lower than expected.

Weinstein's scale (WNS) features items relevant for a student population, and therefore may be of limited value in field research. LEF, while more promising from a conceptual point of view, is twice as long as WNS. In order to address the need for an economic instrument while striving to maintain minimal psychometric standards, we have recently proposed a selection of nine items from LEF to serve as a short form (Zimmer & Ellermeier, 1998*b*) suitable for investigating differences between groups of respondents.

It must be emphasized, however, that the present investigation provides information only on the psychometric properties of self-report noise-sensitivity measures when applied to a student sample; that is, primarily for academic or laboratory settings. Transfer of the conclusions to samples with different demographic characteristics, particularly samples representative of the population at large, may not be straightforward, as became evident when the relationship between noise sensitivity and age was considered.

Noise sensitivity and objective noise effects

For three of the four noise-sensitivity measures investigated, the present study failed to show a correlation with objective performance decrements under noise; the association with the fourth measure, LEF, however, was weak. The application of the broadly-defined concept of noise sensitivity to a task requiring highly specific memory resources may well have constrained the strength of the correlation. At least with regard to the LEF scale, exploring the relationship to objectively measured noise effects in a *variety* of tasks may be more promising.

Data collected on a subset of our sample ($n = 25$) indicate, however, that even if queried specifically to estimate the amount of disruption produced in the given task, the association between objective performance decrements and the *estimated* disruption is not much stronger than the association with general noise sensitivity: $r = 0{\cdot}29$ (for details see Ellermeier & Zimmer, 1997). This finding is in line with results by Mabe and West (1982), who showed that performance and its subjective evaluation are only weakly related in a variety of task domains.

Conceptualization of noise sensitivity

The stated goal of the present study was to make a *methodological* contribution to the measurement of noise sensitivity; that is, to identify the best available instruments on the basis of accepted psychometric standards. Nevertheless, a number of conceptual issues that have been continuously plagueing research on noise sensitivity re-surfaced, especially in the context of our validational strategy. Clearly, replicating a previously obtained pattern of correlations with other personality measures is a preliminary solution due to the absence of a unifying theory of noise sensitivity and its interrelations with other constructs. This deficiency has been recognized (Jones & Davies, 1984; Öhrström *et al.*, 1988), and proposals have been outlined to integrate the concept into environmental stress theory (see Staples, 1966) or other interactionist approaches (Lercher, 1996).

Obviously, positioning noise sensitivity in a larger web of constructs calls for more data to be collected to investigate specific hypotheses. In our own laboratory, we have recently addressed the question of whether individual differences in noise sensitivity are related to differences in auditory functioning at all (Ellermeier *et al.*, in preparation). It turned out that groups of participants exhibiting 'low' vs 'high' noise sensitivity (LEF) were indistinguishable on the basis of absolute thresholds, intensity discrimination, simple auditory reaction time, or power–function exponents for loudness. Small but systematic differences emerged only when judgmental aspects entered into the psychoacoustic task, such as in verbal loudness estimates, or in ratings of the unpleasantness of sounds, suggesting that noise sensitivity reflects attitudinal or evaluative components of the response to noise, rather than a sensory predisposition.

Further laboratory studies might specify which component of the annoyance reaction (e.g. intensive, emotional, or evaluative) noise sensitivity measures predict, while field studies might clarify the role of exposure (or exposure history) in modifying individual noise sensitivity.

Conclusions

To conclude, the results of the present study show systematic differences between available measures of noise sensitivity, to the effect that, in contrast to full-fledged questionnaires, which produce precise, reliable, and valid results, one-item ratings cannot be recommended for the measurement of noise sensitivity. More fundamentally, the present results also caution strongly against the use of *ad hoc* scales not subjected to thorough psychometric evaluation.

Acknowledgement

We would like to thank Anders Kjellberg and two anonymous reviewers who made detailed comments that helped to improve the clarity of the manuscript. We are also grateful to Maria Klatte for providing us with a computer program to conduct the 'irrelevant speech' experiment, and to the students in two courses on 'test construction' for putting in overtime to collect data. Portions of this research were presented at the International Congress on Noise Control Engineering held in Budapest, Hungary, August 1997.

Note

Reprint requests and correspondance should be addressed to Karin Zimmer, Institut für Kognitionsforschung, FB5/ A6, Carl-von-Ossietzky-Universtät Oldenburg, Postfach 2503, 26111 Oldenburg, Germany.

References

Dornic, S., Laaksonen, T. & Ekehammar, B. (1990). *Noise sensitivity: general self-reports vs noise effect in laboratory situations.* Reports from the Department of Psychology, University of Stockholm, No. 716.

Ellermeier, W. & Zimmer, K. (1997). Individual differences in susceptibility to the 'irrelevant speech effect'. *Journal of the Acoustical Society of America* **102,** 2191–2199.

Ellermeier, W., Eigenstetter, M. & Zimmer, K. (in preparation). Sensory and judgmental aspects of individual noise sensitivity.

Flor, H. (1991). *Psychobiologie des Schmerzes.* Bern: Huber.

Green, D. M. & Fidell, S. (1991). Variability in the criterion for reporting annoyance in community noise suryeys. *Journal of the Acoustical Society of America* **89,** 234–243.

Griffiths, I. & Delauzun, F. (1977). Individual differences in sensitivity to traffic noise: an empirical study. *Journal of Sound and Vibration* **55,** 93–107.

Gulliksen, H. (1964). *Theory of Mental Tests.* New York: Wiley.

Guski, R. (1987). *Lärm. Wirkungen unerwünschter geräusche* [*Noise. Effects of unwanted sounds*]. Bern: Huber.

Hautzinger, M. & Bailer, M. (1993). *ADS Allgemeine Depressionsskala.* Weinheim: Beltz.

Job, R. F. S. (1988). Community response to noise: a review of factors influencing the relationship between noise exposure and reaction. *Journal of the Acoustical Society of America* **83,** 991–1001.

Job, R. F. S. (in press). Noise sensitivity as a factor influencing human reaction to noise. *Noise and Health.*

Jones, D.M. & Davies, D.R. (1984). Individual and group differences in the response to noise. In D. M. Jones & A. J. Chapman (Eds) *Noise and Society.* New York: Wiley, pp. 125–153.

Jones, D. M. & Morris, N. (1992). Irrelevant speech and cognition. In D .M. Jones & A. P. Smith (Eds) *Handbook* *of Human Performance, Vol. 1.* London: Academic Press, pp. 29–53.

Kjellberg, A., Landström, U., Tesarz, M., Söderberg, L. & Akerlund, E. (1996). The effects of non-physical noise characteristics, ongoing task and noise sensitivity on annoyance and distraction due to noise at work. *Journal of Environmental Psychology* **16,** 123–136

Langdon, F. (1976*a*). Noise nuisance caused by road traffic in residential areas: part I. *Journal of Sound and Vibration* **47,** 243–263.

Langdon, F. (1976*b*). Noise nuisance caused by road traffic in residential areas: part II. *Journal of Sound and Vibration* **47,** 265–282.

Lercher, P. (1996). Environmental noise and health: an integrated research perspective. *Environment International* **22,** 117–129.

Laux, L., Glanzmann, P., Schaffner, P. & Spielberger, C. (1981). *Das State-Trait-Angstinventar STAI.* Weinheim: Beltz.

Lienert, G. A. & Raatz, U. (1994). *Testaufbau und Testanalyse* [Test construction and test evaluation] (5th Edn). Weinheim, Germany: Beltz.

Lopez Barrio, I. & Carles, J. (1993). Subjective response to traffic noise. The importance given to noise environment in choosing a place of residence. In M. Vallet (Ed.) *Noise as a Public Health Problem. Proceedings of the 6th International Congress (Volume II).* Nice: Institut National de Rechere sur les Transports et leur Sécurité, pp. 205–208.

Mabe, P. A. & West, S. W. (1982). Validity of self-evaluation of ability: review and meta-analysis. *Journal of Applied Psychology* **67,** 280–296.

McKennell, A. (1969). Methodological problems in a survey of aircraft noise annoyance. *The Statistician* **19,** 1–29.

Moreira, N. & Bryan, M. (1972). Noise annoyance susceptibility. *Journal of Sound and Vibration* **21,** 449–462.

Müller-Andritzky, M, Chassein, J. & Schick, A. (1990). *Cross cultural investigations into the assessment of self-produced noise and noise from neighbours in the residential neighbourhood.* Reports from the Institute for Research into Man–Environmental Relations, University of Oldenburg, No. 12.

Öhrström, E., Björkman, M. & Rylander, R. (1988). Noise annoyance with regard to neurophysiological sensitivity, subjective noise sensitivity and personality variables. *Psychological Medicine* **18,** 605–613.

Raw, G. & Griffiths, I. (1988). Individual differences in response to road traffic noise. *Journal of Sound and Vibration* **121,** 463–471.

Salamé, P. & Baddeley, A.D. (1982). Disruption of short-term memory by unattended speech: implications for the structure of working memory. *Journal of Verbal Learning and Verbal Behavior* **21,** 150–164.

Schwenkmezger, P., Hodapp, V. & Spielberger, C. (1991). *Das State-Trait-Ärgerausdrucks-Inventar STAXI.* Bern: Huber.

Stansfeld, S. (1992). Noise, noise sensitivity and psychiatric disorder: epidemiological and psychophysiological studies. *Psychological Medicine. Monograph supplement 22.*

Stansfeld, S., Clark, C., Jenkins, L. & Tarnopolsky, A. (1985). Sensitivity to noise in a community sample: I. Measurement of psychiatric disorder and personality. *Psychological Medicine* **15,** 243–254.

Staples, S. L. (1996). Human response to environmental noise: psychological research and public policy. *American Psychologist* **51,** 143–150.

Taylor, S. M. (1984). A path model of aircraft noise annoyance. *Journal of Sound and Vibration* **96,** 243–260.

Weinstein, N. D. (1978). Individual differences in reactions to noise: a longitudinal study in a college dormitory. *Journal of Applied Psychology* **63,** 458–466.

Zimmer, K. & Ellermeier, W. (1997). Eine deutsche Version der Lärmempfindlichkeitsskala von Weinstein [A German version of Weinstein's noise sensitivity scale]. *Zeitschrift für Lärmbekämpfung* **44,** 107–110.

Zimmer, K. & Ellermeier, W. (1998*a*). Konstruktion und Evaluation eines Fragebogens zur Erfassung der individuellen Lärmempfindlichkeit [Construction and evaluation of a noise sensitivity questionnaire]. *Diagnostica* **44,** 11–20.

Zimmer, K. & Ellermeier, W. (1998*b*). Ein Kurzfragebogen zur Erfassung der individuellen Lärmempfindlichkeit [Short form of a questionnaire measuring individual noise sensitivity]. *Umweltpsychologie* **2,** 54–63.