A measurement model for general negative reaction to noise

Maarten Kroesen[1], Eric J.E. Molin[1], Henk M.E. Miedema[2], Henk Vos[2], Sabine A. Janssen[2], Bert van Wee[1]

1 Delft University of Technology, Faculty of Technology, Policy and Management, P.O. Box 5015, 2600 GA Delft, The Netherlands

2 TNO Department of Environment and Health, P.O. Box 49, 2600 AA Delft, The Netherlands

*corresponding author: Maarten Kroesen, e-mail: M.Kroesen@TUDelft.nl

INTRODUCTION

In order to compare results of community surveys from different countries noise reaction questions have been standardized (Fields et al. 2001). The two selected questions, a 5-point verbal and an 11-point numeric scale, 'seek to obtain general, persistent reactions that allow respondents to integrate their experiences over different times and locations in their home' (Fields et al. 2001, p. 665). The questions are formulated as follows:

"Thinking about the last (12 months or so), when you are here at home, how much does noise from (noise source) bother, disturb, or annoy you?"

However, Job & Sakashita (2007) note that negative reactions to noise encompass more aspects than the mere concept of annoyance. In this respect they refer to a study of Hede et al. (1979) which found that respondents use many different words, other than and unrelated with annoyance, to describe their negative reactions in response to the noise. Hence, the standardized scale captures only a part of subjects' overall (negative) assessment with respect to the impact of a certain noise source on their living conditions. This argument is also substantiated with reference to the work of Job et al. (2001), in which general measures of reaction (dissatisfaction and perceived affectedness) have been shown to have superior psychometric properties in comparison to specific reactions such as annoyance. Job & Sakashita (2007) therefore claim that the inclusion of these general measures is imperative for the valid measurement of community reaction.

With this background the aim of the present study is to assess to which extent noise annoyance captures all negative reactions in response to aircraft noise. For this purpose aircraft noise annoyance is hypothesized to be a manifestation of a more general concept, labeled 'general negative reaction in response to aircraft noise'. Using the analytical framework of Edwards (2001), this concept is measured indirectly via a measurement model by specifying it as a reflective second-order construct. To empirically test the specified structure, data from two available datasets are re-analyzed (Fields' codes NET-371 and GER-531). The use of two datasets provides an effective way to cross-validate the results. Lastly, to assess the validity of the specified model structure, the second-order factor is used to predict two criterion variables, namely residential satisfaction and perceived health.

The remaining part of this paper is structured as follows. The method section describes the methodological approach adopted to answer the question whether noise annoyance indeed captures all relevant negative responses to aircraft noise. Next, the results of two measurement models, which can be used to answer this question, are presented. The last section presents the conclusions and ends with several directions for future research.

DEVELOPMENT OF A MEASUREMENT MODEL

In this section the approach to measure general negative reaction via a measurement model will be described. The idea is that if this concept can be adequately measured, it can also be inferred how well noise annoyance performs as indicator of this construct. The conceptualization of the measurement model is based on two important premises. In the first place it is assumed that such a general negative reaction concept exists in the real world. Evidence for the tenability of this premise is provided by Job et al. (2001) who have shown that general reaction measures have superior psychometric properties in comparison to specific reaction measures. The second premise is that the variance in specific reaction measures is composed of variation specific to the measure plus variation related to the general negative reaction concept. Hence, to a certain extent, specific reaction measures are assumed to reflect (or manifest) the more abstract negative reaction concept. Through specification of an underlying factor the variance common to these specific measures can then be extracted and this abstract concept can be measured.

Via a literature review four distinct and specific measures of negative reaction, which have been found to correlate with physical noise levels in previous research, are identified:

- *Noise annoyance*: Schultz (1978), Job (1988) and Fidell (2003).
- *Perceived disturbance*, alternative labels: sleep disturbance, speech interferences (Taylor 1984), activity interference (Lercher 1996).
- *Non-noise annoyances*, alternative labels: awareness of non-noise problems (Fields 1993), non-noise impacts like odor and vibrations (Lercher 1996).
- *Anxiety and fear*, alternative labels: perceived health effects of noise (McKennell 1963), fear of aircraft accidents (Leonard & Borsky 1973), fear or harm connected with the noise source (Guski 1999).

In Figure 1 the measurement model is fully specified. To exclude measurement errors at the level of the dimensions, they are not measured directly, but via multiple observed indicators. The general negative reaction concept is modeled as a second-order reflective factor (Edwards 2001), which extracts the common variance of the specific dimensions.

The validity of this conceptualization is tested in four ways:

- Firstly, the overall fit of the model will be reviewed to assess whether the data supports the second-order factor structure;
- secondly, the strength and significance of the parameter estimates will be examined to assess whether the dimensions converge on the same underlying construct;
- thirdly, the common variance extracted by the second-order construct will be used to predict two outcome variables, namely residential satisfaction and perceived health. It can then be tested whether the specific dimensions can explain variance over and above the general negative reaction construct, or whether the latter concept indeed ‘captures’ all variance relevant for this prediction. If the latter is the case, this would support the construct validity of the general negative reaction factor, for it would indeed measure what it is intended to measure.

- And lastly, it will be assessed whether the results are consistently replicated across two samples.

If the results are supportive for the specified factor structure, the question to which extent noise annoyance captures all relevant negative feelings and emotions can be easily answered through examination of the (standardized) factor loading of this dimension.

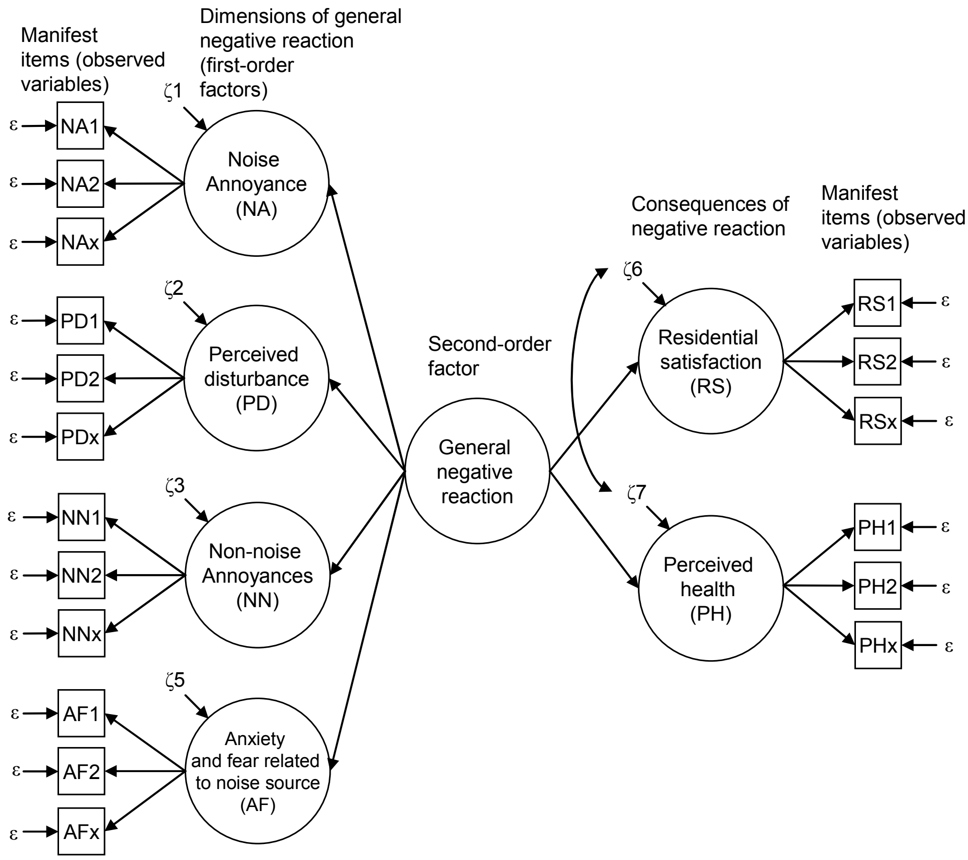

Figure 1: Measurement model and consequences of general negative reaction
Note: because residential satisfaction and perceived health may be causally related or a third variable may influence both, the error terms of these factors are assumed to correlate.

METHOD

Data

Data from two community surveys, one at Schiphol Airport in the Netherlands (model I) and the other at Frankfurt Airport in Germany (model II), are used to estimate the hypothesized model in Figure 1.

For the first model a dataset from an aircraft noise study in the Netherlands (Fields' code NET-371), described by TNO/RIVM (1998), Miedema et al. (2000) and Franssen et al. (2004), is used (N=11,812). In this study a stratified random sample was drawn from the population living within a 25 kilometer radius around Schiphol airport, which is the largest airport in the Netherlands. The response rate was 39 %. Cases with more than 10 % missing values are deleted (N=954).

The second model uses a dataset from an aircraft noise study conducted in Germany at Frankfurt Airport (N=2,312; Fields' code GER-531), described in Schreckenberg & Meis (2006). Within this study a random sample was drawn from residents living in 66 residential areas located within a 40 kilometer radius around the Frankfurt Airport.

The response rate was 61 %. Again, cases with more than 10 % of the values missing are deleted (N=106).

The remaining cases in both datasets, N=10,858 and N=2,206, are unequal in size. A random sample of N=2,206 from the NET-371 dataset is therefore selected to ensure that both studies have equal power (i.e. the probability of rejecting a false H0) of the chi-square test to detect discrepancies between the model implied and observed covariance matrix.

Measures

To ensure that the structural estimates of the paths between the constructs are corrected for random measurement errors, each construct in Figure 1 is measured with multiple observed indicators.

With respect to the used observed items of the constructs it needs to be noted that the used questions for noise annoyance in the GER-531 sample exactly match the standardized noise reaction questions developed by Fields et al. (2001). For noise annoyance in the NET-371 sample only the first question is the same as the first standardized question.

The rest of the used indicators are not the same in the two datasets (i.e. different wording, scales, etc.). Due to restrictions in the available space they are not reported.

Table 1: Intercorrelations (all p< .001) and reliability estimates (on the diagonals in italic)

NET-371 sample (N=2,206)								
Dimensions	**Label**	**# items**	**NA**	**PD**	**NN**	**AF**	**RS**	**PH**
Noise annoyance	NA	2	*0.89*					
Perceived disturbance	PD	4	0.92	*0.86*				
Non-noise annoyances	NN	2	0.81	0.82	*0.75*			
Anxiety and fear	AF	3	0.85	0.81	0.80	*0.77*		
Criterion variables								
Residential satisfaction	RS	2	-0.46	-0.48	-0.46	-0.46	*0.60*	
Perceived Health	PH	2	-0.23	-0.31	-0.24	-0.27	0.34	*0.68*
GER-531 sample (N=2,206)								
Dimensions	**Label**	**# items**	**NA**	**PD**	**NN**	**AF**	**RS**	**PH**
Noise annoyance	NA	2	*0.93*					
Perceived disturbance	PD	5	0.91	*0.93*				
Non-noise annoyances	NN	1*	0.77	0.74	*0.91*			
Anxiety and fear	AF	3	0.82	0.84	0.76	*0.88*		
Criterion variables								
Residential satisfaction	RS	2	-0.43	-0.45	-0.34	-0.37	*0.60*	
Perceived Health	PH	2	-0.19	-0.18	-0.20	-0.25	0.12	*0.79*

* For non-noise annoyances in the GER-531 sample only one indicator was present in the dataset. The reliability of this construct was therefore fixed by constraining the error variance of the observed indicator underlying this construct. For this purpose the assumption is made that the reliability of the dimensions is equal to the average reliability of the other dimensions (α=0.91).

Table 1 presents the number of items, the correlation matrices and the reliability estimates (Cronbach alpha's) of the constructs.

Analysis

The chi-square statistic, generally used to test the fit of the model, is due to its sensitivity to large sample sizes (N>2,000), expected to be significant (indicating a lack of fit). The following fit indices, which are not dependent on sample size, are therefore used to evaluate the fit of the estimated models: the Root Mean Square Error of Approximation (RMSEA) (Browne & Cudeck 1993), the Standardized Root Mean Residual (SRMR) and the Comparative Fit Index (CFI) (Bentler 1990). A well-fitting model is defined as having values below .06 and .08 for RSMEA and SRMR respectively and a CFI value greater than .95 (Hu & Bentler 1999).

To avoid biased estimates due to problems with non-normality, the Asymptotic Distribution Free (ADF) estimator of AMOS 7.0, which is developed by Browne (1984), is used to estimate the models.

RESULTS

Test 1: Overall model fit

In Table 2 the fit statistics of both models are presented. Based on these figures it can be concluded that both datasets fit the second-order factor structure well. In addition, a review of the modification indices indicates that adding additional paths or correlations does not lead to substantial decreases in the chi-square statistic. Hence, the model structure, as it is depicted in Figure 1, is supported by the data. This means that, as hypothesized, the specified dimensions and criterion variables are the sole causes for the structural (common) variance in their respective observed indicators and that the general negative reaction construct is the sole cause for the structural (common) variance in the four dimensions.

Table 2: Fit statistics of two models

N=2,206	χ^2	Df	RMSEA	SRMR	CFI
Model 1 (NET-371)	421.0	83	0.043	0.035	0.986
Model 2 (GER-531)	453.1	84	0.045	0.045	0.995

Test 2: Factor loadings

Table 3 presents the standardized factor loadings and regression weights of the dimensions and criterion variables respectively. All estimates fall below the .001 significance level. The factor loadings are all greater than the conventional minimum value of .7. This means that the dimensions converge on the same underlying construct and can be treated as indicators of the same concept.

Table 3: Standardized parameter estimates and proportions of explained variance

	NET-371		GER-531	
Dimensions	Factor Loading	Explained variance (%)	Factor Loading	Explained variance (%)
Noise annoyance	0.96	92.5	0.95	90.1
Perceived disturbance	0.95	89.8	0.96	92.4
Non-noise annoyances	0.87	75.1	0.88	78.0
Anxiety and fear	0.89	78.9	0.90	80.2

Test 3: Predictive accuracy

In the prediction of the two criterion variables the general negative reaction constructs performs well. In residential satisfaction it can explain 25.4 % and 21.7 % of the total variance in the Dutch and German sample respectively. For perceived health these figures are 8.7 % and 5.8 % respectively (see Table 3).

Table 4: Standardized parameter estimates and proportions of explained variance in criterion variables

	NET-371		GER-531	
Criterion variables	**Regression weight**	**Explained variance (%)**	**Regression weight**	**Explained variance (%)**
Residential satisfaction	-0.50	25.4	-0.47	21.7
Perceived health	-0.30	8.7	-0.24	5.8

The real question is whether the variance extracted by the general negative reaction construct is the only relevant variance in the prediction of residential satisfaction and perceived health and thus whether variance specific to the dimensions is irrelevant. This is done through a review of the modification indices related to the paths which can be drawn between the dimensions and the criterion variables (Edwards 2001). These indices indicate the decrease in the chi-square statistic which would be obtained if the extra parameters related to these paths were really estimated. Hence, if such decreases are not statistically significant it can be concluded that the dimensions do not contain additional variance, which could be used to explain variance in the criterion variables over and above the general negative reaction construct. This would mean that this latter construct captures all relevant negative feelings and reactions in response to aircraft noise.

Table 5: Modification indices for the paths from the dimensions to the criterion variables

NET-371	**Criterion variable**	
Dimension	Residential satisfaction	Perceived health
Noise annoyance	0.03	0.45
Perceived disturbance	0.00	0.60
Non-noise annoyances	0.17	0.14
Anxiety and fear	0.03	0.41
GER-531	**Criterion variable**	
Dimension	Residential satisfaction	Perceived health
Noise annoyance	0.01	0.03
Perceived disturbance	0.15	0.13
Non-noise annoyances	0.02	0.00
Anxiety and fear	0.52	2.02

In Table 4 the modification indices for the relationships between the dimensions and the criterion variables are given. Since none of the modification indices in Table 5 exceed the conventional value of 4, it can be concluded that dimension specificities (i.e. variance specific to the dimensions) are irrelevant in the prediction of the two criterion variables. These results provide additional support for the validity of the measurement model and indicate that the general negative reaction construct is effective in capturing all relevant information residing in the dimensions.

Test 4: Cross-validation

A remarkable (and desirable) result is that both models, which are estimated based on data from different populations and used different observed indicators (i.e. alternative question wording, number of items and overall questionnaire design), are very much alike in terms of overall model fit as well as the parameter estimates. The two patterns of factor loadings and regression estimates match each other very well (the estimates from both models are almost the same). The fact that the results can be replicated so well and under such different circumstances provides strong evidence for the validity of the model.

Noise annoyance: a good measure of general negative reaction?

Lastly, the question, whether noise annoyance is a good indicator for general negative reaction, can be answered. From Table 3 it can be inferred that noise annoyance and perceived disturbance perform equally well as indicators of general negative reaction. Non-noise annoyances and anxiety and fear related to the noise source also perform well in an absolute sense but are relatively worse indicators in comparison to noise annoyance and perceived disturbance.

Overall, it can be concluded that noise annoyance is a strong reflection of general negative reaction, but does not capture all relevant variance. Table 5 provides several figures to further illustrate this point. The given values indicate the loss in predictive accuracy if noise annoyance, instead of general negative reaction, would be used in the prediction of the two criterion variables. It can be concluded that, using noise annoyance, a smaller proportion of the total variance in the criterion variables can be explained.

Table 5: Explained variance in criterion variables

	NET-371		**GER-531**	
Criterion variables	**Predictor: general negative reaction**	**Predictor: noise annoyance**	**Predictor: general negative reaction**	**Predictor: noise annoyance**
Residential satisfaction	25.4 %	20.6 %	21.7 %	17.3 %
Perceived health	8.7 %	5.0 %	5.8 %	3.9 %

CONCLUSION

In this study a measurement model is developed to measure general negative reaction in response to aircraft noise. Estimation of the model using two different datasets yielded a good fit to the data and supported the second-order factor structure. Additional support for the specified structure is found in the superior predictive accuracy of the general negative construct and the fact that the results are consistently replicated in two different samples. Using the general negative construct it is inferred that noise annoyance is a strong reflection of this construct, but does not capture all relevant information.

Based on a reflection on the present study two directions for future research are identified. The first is related to identifying old as well as developing new theories which can explain the particular strong factor structure found in the data. Such theories should be able to answer the question why these very different responses (e.g. annoyance, fear and disturbance) are so strongly interrelated. It might be that several theoretical notions, which relate to the individual associations between variables, underlie the present data structure, or that there is one overarching theory which can provide a holistic explanation. A theory of the latter kind might be the social-psychological consistency theory of Festinger (1957), which is previously applied to the appraisal of aircraft noise by Bröer (2006). According to Bröer subjects appraise noise within a holistic frame which consists of a set of consonant feelings and beliefs. This would provide an explanation for the consistent responses found in this study.

The second research direction is related to the assumption that the factor structure as hypothesized in Figure 1 holds for all people, which is, of course, an assumption underlying the models presented in this study. Additionally, if this structure is indeed the same for different people, which would indicate so-called configural invariance, it might be that the factor loadings connecting the factors with their indicators are not of equal size for different groups, which would indicate the presence of measurement variance. An interesting question, for example, would be whether the pattern of factor loadings is different for people living close to the airport in comparison to people living distant from it. More specifically, it can hypothesized that within the former group reactions like fear or non-noise annoyances (i.e. vibrations) play a greater role within the general reaction construct and hence would receive a greater factor loading. This would mean that the meaning of the concept of general reaction would differ for this group and also that it might be stronger related to criterion variables like residential satisfaction and perceived health.

REFERENCES

Bentler PM (1990). Comparative fit indices in structural models. Psychol Bull 107: 238-246.

Bröer C (2006). Beleid vormt overlast. Amsterdam: Aksant.

Browne MW (1984). Asymptotic distribution free methods in the analysis of covariance structures. Br J Math Stat Psychol 37: 62-83.

Browne MW, Cudeck R (1993). Alternative ways of assessing model fit. In: Bollen KA, Long JS (eds.): Testing structural equation models (pp 136-162). Thousand Oaks, CA: Sage.

Edwards JR (2001). Multidimensional constructs in organizational behaviour research: an integrative analytical framework. Organ Res Methods 4: 144-192.

Festinger L (1957). A theory of cognitive dissonance. Evanston, IL: Row, Perterson & Company.

Fidell S (2003). The Schultz curve 25 years later: a research perspective. J Acoust Soc Am 114: 3007-3015.

Fields JM (1993). Effect of personal and situational variables on noise annoyance in residential areas. J Acoust Soc Am 46: 2753-2763.

Fields JM, de Jong RG, Gjestland T, Flindell IH, Job RFS, Kurra S, Lercher P, Vallet M, Yano T, Guski R, Schumer R (2001). Standardised general-purpose noise reaction questions for community noise surveys: research and a recommendation community response to noise team of ICBEN (the International Commission on the Biological Effects of Noise). J Sound Vibr 242: 641-679.

Franssen EAM, van Wiechen CMAG, Nagelkerke NJD, Lebret E (2004). Aircraft noise around a large international airport and its impact on general health and medication use. Occup Environ Med 61: 405-413.

Hede AJ, Bullen RB, Rose JA (1979). A social study of the nature of subjective reaction to aircraft noise. National Acoustic Laboratories Report no. 79. Canberra: Australian Government Publishing Services.

Hu L, Bentler PM (1999). Cutoff criteria for fit indices in covariance structure analysis: conventional criteria versus new alternatives. Struct Equat Model 6: 1-55.

Job RFS (1988). Community response to noise: a review of factors influencing the relationship between noise exposure and reaction. J Acoust Soc Am 83: 991-1001.

Job RFS, Sakashita C (2007). Assessment of community reaction to noise: methodology in question wording and delivery. Proceedings of the Internoise 2007 congress. Istanbul, Turkey.

Job RFS, Hatfield J, Carter NL, Peploe P, Taylor R, Morrell S (2001). General scales of community reaction to noise (dissatisfaction and perceived affectedness) are more reliable than scales of annoyance. J Acoust Soc Am 110: 939-946.

Leonard S, Borsky PN (1973). A causal model for relating noise exposure, psychological variables and aircraft noise annoyance. Proceedings of the International Congress on Noise as a Public Health Problem (pp 627-644). Dubrovnik.

Lercher P (1996). Environmental noise and health: an integrated research perspective. Environ Int 22: 117-129.

McKennell AC (1963). Aircraft noise annoyance around London (Heathrow) Airport. U.K. Government Social Survey Report SS337. London: HMSO.

Miedema HME, Vos H, de Jong RG (2000). Community reaction to aircraft noise: time-of-day penalty and tradeoff between levels of overflights. J Acoust Soc Am 107: 3245–3253.

Schreckenberg D, Meis M (2006). Belästigung durch Fluglärm im Umfeld des Frankfurter Flughafens: Gutachten im Auftrag des Regionalen Dialogforums Flughafen Frankfurt; Endbericht. Bochum, Oldenburg: AG Fluglärmwirkung.

Schultz TJ (1978). Synthesis of social surveys on noise annoyance. J Acoust Soc Am 64: 377-405.

Taylor SM (1984). A path model of aircraft noise annoyance? J Sound Vibr 96: 243-260.

TNO/RIVM (1998). Hinder, slaapverstoring, gezondheids en belevingsaspecten in de regio Schiphol, resultaten van een vragenlijstonderzoek. TNO-PG, Leiden, The Netherlands, Report 98.039/RIVM, Bilthoven, The Netherlands, Report 441520010.