A measurement model for general noise reaction in response to aircraft noise

Maarten Kroesen[a)]
Faculty of Technology, Policy and Management, Delft University of Technology, P.O. Box 5015, 2600 GA Delft, The Netherlands

Dirk Schreckenberg
ZEUS GmbH, Sennbrink 46, 58093 Hagen, Germany

(Received 13 January 2010; revised 20 September 2010; accepted 19 October 2010)

In this paper a measurement model for general noise reaction (GNR) in response to aircraft noise is developed to assess the performance of aircraft noise annoyance and a direct measure of general reaction as indicators of this concept. For this purpose GNR is conceptualized as a superordinate latent construct underlying particular manifestations. This conceptualization is empirically tested through estimation of a second-order factor model. Data from a community survey at Frankfurt Airport are used for this purpose ($N = 2206$). The data fit the hypothesized factor structure well and support the conceptualization of GNR as a superordinate construct. It is concluded that noise annoyance and a direct measure of general reaction to noise capture a large part of the negative feelings and emotions in response to aircraft noise but are unable to capture all relevant variance. The paper concludes with recommendations for the valid measurement of community reaction and several directions for further research. © *2011 Acoustical Society of America*.
[DOI: 10.1121/1.3514542]

PACS number(s): 43.50.Rq, 43.50.Qp, 43.50.Lj [BSF] Pages: 200–210

I. INTRODUCTION

As indicated by Job *et al*. (2001) the valid measurement of community reaction in response to environmental noise is important for two reasons: First, it, in itself, can be regarded as an important negative health factor defined broadly as the absence of social well-being, and second, it may contribute to noise-induced health problems (e.g., self-reported symptoms, hypertension, mental health problems). The latter reason becomes even more important since there is evidence that reaction to noise is a better predictor of noise-related health effects than noise exposure itself (Job, 1996; Job *et al*., 2001). Furthermore, Job *et al*. (2001) contend that noise annoyance questions, which are mainly used to assess negative reaction in socio-acoustic surveys, fail to measure many possible and important reactions to noise. According to the authors, noise annoyance measures are too narrow to capture the full breadth of human reaction. Obviously, this is a serious problem if associations between subjective reaction to noise and possible consequences are studied.

Job *et al*. (2001) illustrate the narrowness of noise annoyance as a measure of general negative reaction to noise with reference to Hede *et al*. (1979) who found that respondents use many different words, other than and (semantically) unrelated with annoyance, to describe their feelings in response to noise. Hence, as argued by Job and Sakashita (2007), it is likely that measures, such as the standardized noise annoyance questions of Fields *et al*. (2001), capture only a part of subjects' overall (negative) assessment with respect to the impact of a certain noise source on one's living conditions. Job *et al*. (2001) substantiate this argument by showing that general measures of reaction to noise (being dissatisfied by noise and perceived affectedness) have superior psychometric properties in comparison to specific reactions such as noise annoyance.[1] According to Job and Sakashita (2007) the inclusion of these general measures is therefore imperative for the valid measurement of human reaction to noise.

By measuring general reaction to noise via direct measures, Job *et al*. (2001) implicitly assume the existence of an abstract underlying construct in which all negative feelings and emotions in response to (aircraft) noise are integrated. In addition, by showing that the general noise reaction (GNR) measures are more strongly correlated with the measures of activity disturbance than the noise annoyance measures, they support the conclusion that GNR measures are able to capture more relevant variance in responses to noise. However, although Job *et al*. (2001) show that general measures capture *more* relevant variance in the prediction of other concepts (i.e., activity disturbance), it remains unclear whether these measures indeed capture *all* the relevant variances. In addition, to assess the validity of direct general reaction measures, preferably criteria outside the content domain of general reaction to noise should be used. Similar to noise annoyance, activity disturbance can be regarded as a particular dimension of general reaction and not as a relevant criterion variable of this construct.

With this background the main aim of the present study is to develop a measurement model of GNR, which can be shown to capture all relevant variance in response to aircraft noise. To develop this model we conceptualize GNR as a latent multidimensional construct that underlies particular manifestations such as noise annoyance and activity

[a)]Author to whom correspondence should be addressed. Electronic mail: m.kroesen@tudelft.nl

disturbance. Based on classical test theory we assume that noise annoyance and other measures of subjective reaction to noise contain, in part, "true" variance related to the theoretical concept we wish to measure (i.e., GNR) and, in part, dimension specific variance. By measuring GNR indirectly via multiple dimensions the concept is "stripped" from dimension specific variance. Only the variance common to all dimensions remains. Next, by modeling the specific dimensions as well as the second-order concept (GNR) in a single model together with aircraft noise exposure and relevant criterion variables, it can be assessed whether the variance captured by GNR indeed represents all of the relevant subjective variability resulting from exposure to aircraft noise.

In this study, perceived mental health and physical health are identified as relevant criterion variables. Research has previously established associations between subjective noise reaction and health outcomes, ranging from self-reported effects, such as (self-reported) psychosocial well-being (Job, 1996), to objective medical outcomes, such as increased blood pressure (Babisch *et al.*, 2007). Based on these previously established relationships we assume that GNR influences a person's mental and physical health and that these assumed consequences of GNR are suitable as criteria for validity analysis of GNR.

If sufficient support is found for the validity of the proposed conceptualization, which would support the conclusion that GNR indeed captures all the relevant subjective variability resulting from exposure to aircraft noise, the secondary aim is to examine how well aircraft noise annoyance and a direct measure of general reaction to noise (dissatisfaction with aircraft noise) perform as indicators of this construct.

To attain the above-stated aims a structural equation model is estimated based on data from a field survey conducted near Frankfurt airport. Within this modeling approach the dimensions of GNR, as well as GNR as a second-order factor, can be specified in a single confirmatory factor model, allowing us to assess and compare the predictive ability of GNR versus the dimensions of GNR in relation to the included criterion variables. Additionally, a structural equation model can account for measurement errors (present in observed variables) leading to less biased parameter estimates between the structural variables.

II. DEVELOPMENT OF A MEASUREMENT MODEL FOR GNR

In this section the measurement model for GNR will be developed. First, the theoretical background of multidimensional constructs is described (Sec. II A). Second, the content domain of GNR will be established (Sec. II B). Finally, it concludes with the specification of the measurement model of GNR and its theoretical underpinnings (Sec. II C).

A. Multidimensional constructs

According to Law *et al.* (1998) a multidimensional construct can be conceptualized as an overall abstraction that represents several distinct but interrelated dimensions which can be grouped together into a single theoretical concept. It can therefore be distinguished both from a unidimensional construct, which refers to a single theoretical concept (e.g., noise annoyance), and from multiple dimensions, which may be related but cannot be unified in an overall theoretical concept (Edwards, 2001).

In principle, there are two ways to specify a multidimensional construct (see Fig. 1). The first assumes the direction of causality to flow from the construct to the dimensions, which is called a reflective (latent factor) model (Nunnally, 1978; Bollen, 1989), and the second assumes the direction of causality to be from the dimensions to the construct, which is called a formative (aggregate composite) model (Fornell and Bookstein, 1982; Bollen and Lennox, 1991). Jarvis *et al.* (2003) give an exhaustive overview of the differences between formative or reflective models, which will be briefly described in the following paragraphs.

A reflective model assumes that the dimensions are manifestations of the underlying construct. Any change in the construct will result in changes in the dimensions and the dimensions are therefore expected to covary. In addition, the

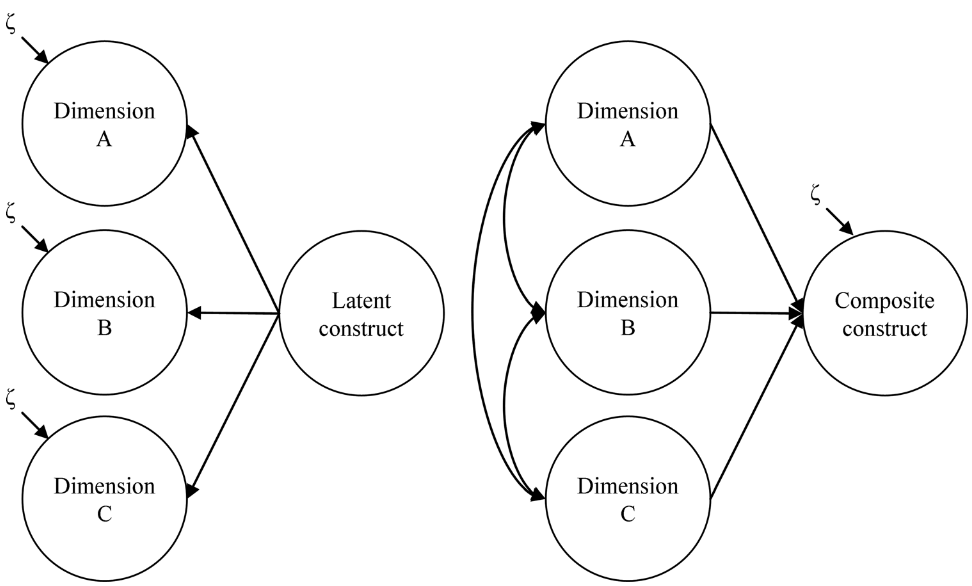

FIG. 1. Reflective (left) and formative (right) measurement models. ζ: Error term of latent construct.

dimensions share a common theme, which means that they are interchangeable and that dropping a dimension does not alter the conceptual domain of the construct. Finally, because the dimensions share the same nomological net they are assumed to have the same consequences and antecedents.

In contrast, the formative model assumes that the dimensions are defining components of the construct. It is therefore important that all components relevant to the conceptual definition of the construct are included as dimensions. Changes in these dimensions are assumed to cause changes in the construct and not vice versa. Hence, these dimensions do not share a common theme, are not necessarily assumed to covary, are not interchangeable, and can have different antecedents and consequences.

Jarvis *et al*. (2003) also mention two similarities between formative and reflective models. First, both models capture surplus meaning beyond the specific components used to measure the constructs. They represent abstract entities that are not wholly reducible to empirical terms. A reflective construct is reduced to the common part (factor) of a set of dimensions, in effect, capturing a more complex and abstract idea than all dimensions individually. A formative construct combines the individual components and, as a result, captures the meaning of the individual dimensions taken together, which also goes beyond the meaning of the dimensions individually. A second similarity, which is related to the previous one, is that neither model can be adequately represented by a scale score, which, if adopted, would lead to inconsistent structural estimates between the construct and other latent variables (Jarvis *et al*., 2003).

To specify a construct it must be identified as either reflective or formative. For this purpose Jarvis *et al*. (2003) developed a comprehensive list of criteria to guide the specification of the relationships between the dimensions and the construct. These are presented in Table I and will be used in Sec. II C to identify the construct of GNR to aircraft noise as either formative or reflective.

B. Dimensions of GNR

In order to measure GNR to aircraft noise its content domain has to be established. Based on a review of the literature three dimensions of GNR are identified: Noise annoyance, activity disturbance, and feelings of anxiety and fear.

1. Noise annoyance

Aircraft noise annoyance is the most often used indicator to assess negative reaction to noise in socio-acoustic surveys (Job *et al*., 2001). According to Guski (1999) this indicator captures the (long-term) evaluative aspect of the reaction to aircraft noise. Since there is ample evidence that this dimension is associated with aircraft noise exposure (see, e.g., Schultz, 1978; Fidell *et al*., 1991), it can be identified as a relevant dimension of GNR.

2. Activity disturbance

This concept has been given different labels by various authors, but they all convey the same general meaning. For instance, Taylor (1984) uses sleep disturbance and speech interferences to capture its nature; Ahrlin (1988) uses the term interference with daily activities which he decomposes into speech interference, interference with resting and sleeping, and the awakening effect; Guski (1999) refers to it as short-term annoyance; and Guski *et al*. (1999), based on the results of an expert review, refer to it as immediate behavioral noise effects. Among others, Kryter (1982) and Taylor (1984) show that activity disturbance increases with higher levels of noise exposure [using metrics as L_{dn} and $L_{eq(24)}$]. It can therefore be concluded that the concept of *activity disturbance* forms a relevant dimension of GNR.

3. Feelings of anxiety and fear

This dimension includes the evaluation of physical risks associated with the presence of the airport, like the fear for an aircraft crash or anxiety related to the negative health effects of noise, as well as the evaluation of non-physical risks, and like the fear that noise levels will increase in the future or the concern about property devaluation. Again, previous research has established significant correlations between indicators of this dimension and aircraft noise exposure (see, e.g., Alexandre, 1976, for the association between noise exposure and "fear for an aircraft crash"). It can therefore be identified as a relevant dimension of GNR.

TABLE I. Criteria for determining whether a construct is formative or reflective.

	Reflective model	Formative model
(1) Direction of causality from construct to dimension implied by the conceptual definition	Direction of causality is from construct to dimensions Dimensions are manifestations of the construct	Direction of causality is from dimensions to construct Dimensions are defining characteristics of the construct
(2) Interchangeability of the dimensions	Dimensions should be interchangeable Dimensions should share a common theme Dropping a dimensions should not alter the conceptual domain of the construct	Dimensions need not be interchangeable Dimensions need not share a common theme Dropping a dimensions may alter the conceptual domain of the construct
(3) Covariation among the indicators	Dimensions are expected to covary	Dimensions need not necessarily covary
(4) Nomological net of the construct dimensions	Nomological net of the construct dimensions should not differ Dimensions are required to have the same antecedents and consequences	Nomological net of the construct dimensions may differ Dimensions are not required to have the same antecedents and consequences

Source: Jarvis *et al*. (2003)

TABLE II. Identified dimensions of GNR to aircraft noise and alternative labels.

Dimension	Alternative labels
Noise annoyance (NA)	Nuisance, unpleasantness, and getting on one's nerves (evaluative aspects of noise) (Guski *et al.*, 1999)
Activity disturbance (AD)	Sleep disturbance, speech interferences (Taylor, 1984)
	Speech interference, interference with resting and sleeping and the awakening effect (Ahrlin, 1988)
	Activity interference (Hall *et al.*, 1985; Lercher, 1996)
	Short-term reactions (Guski, 1999)
	Immediate behavioral noise effects (Guski *et al.*, 1999)
Feelings of anxiety and fear (AF)	Perceived health effects of noise (McKennell, 1963)
	Fear of aircraft accidents (Leonard and Borsky, 1973)
	Fear of danger/health effect (Lercher, 1996)
	Fear of the noise source (Miedema and Vos, 1999)
	Fear or harm connected with the noise source (Guski, 1999)
	Fear of property devaluation due to aircraft noise (Kroesen *et al.*, 2008)

Table II summarizes the identified dimensions and possible alternative labels. It can be argued that several relevant reactions to aircraft noise have not been exhaustively sampled (e.g., related to anger, perceived control, trust in the authorities). These dimensions were not considered because the present questionnaire did not include appropriate indicators to measure them. Yet, since GNR is conceptualized as a *reflective* multidimensional construct, this possible omission is not problematic. For a reflective construct it holds that dropping a dimension does not alter the conceptual domain of the underlying construct (Jarvis *et al.*, 2003). It is therefore not imperative that all manifestations of GNR are included. The choice for a reflective model is justified in Sec. II C.

C. Specification of GNR

Based on the decisional rules presented in Table I GNR to aircraft noise is identified as a *reflective multidimensional construct*. The justification is elaborated below.

Related to the first rule it is assumed that GNR is an abstract negative feeling that becomes manifested in particular responses to aircraft noise. Based on previous research of Bröer (2006), Bröer and Duyvendak (2009), and Kroesen and Bröer (2009), we expect that people develop a general attitude about aircraft noise, which becomes manifested in particular responses to the noise. Therefore, in line with the reflective model, the direction of causality is assumed to flow from the construct to the dimensions. Related to the second rule it can be determined that the identified dimensions of GNR all share a common theme, namely the (negative) response to aircraft noise. Hence, in contrast to the formative model, dropping or omitting a dimension does not alter the conceptual definition of the construct, which is also why it is not imperative that all possible dimensions are included. Related to the third rule it is expected that the dimensions of GNR covary. For example, if one becomes more annoyed by aircraft noise, one is also expected to be more disturbed by it and become more fearful and worried about the risks. Hence, we expect that the dimensions are mutually interrelated. Finally, the dimensions of GNR share the same nomological net. They are expected to have similar antecedents (e.g., the level of aircraft noise exposure) and consequences (e.g., mental and physical health). Hence, each dimension can be placed in a similar nomological network.

To operationalize GNR as a reflective construct it could be specified as a first-order latent factor with its dimensions as observed variables. However, as Edwards (2001) notes, this approach confounds random measurement error (present in observed items) with dimension specificity (i.e., the systematic variance not captured by GNR) and ignores differences in the relationships between each dimension and its measures. Therefore, to exclude measurement errors at the level of the dimensions, the dimensions are not measured directly but indirectly via multiple observed indicators (Bagozzi and Edwards, 1998). In effect, GNR, as a second-order latent factor, will only extract the common variance from these "pure" dimensions. The left side of Fig. 2 presents the second-order factor model. In the terminology of Edwards (2001) this model can be classified as a superordinate construct. In the remainder of this paper we will discuss about the present conceptualization as such.

III. METHOD

A. Validation approach

The validity of GNR as a superordinate construct is tested in three ways. First, the *model fit* will be reviewed to assess whether the data support the second-order factor structure. Given the large sample size, the χ^2 statistics, which indicates the discrepancy between the model-observed and the model-implied covariance matrices, will be significant even with trivial differences between the matrices. The following fit indices, which are not dependent on sample size, are therefore used to evaluate the fit of the estimated model: The root mean square error of approximation (RMSEA) (Browne and 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 0.06 and 0.08 for RSMEA and SRMR, respectively, and a CFI value greater than 0.95 (Hu and Bentler, 1999).

Second, the *convergent validity* will be examined by assessing the size and significance of the parameter estimates (i.e., the factor loadings) and the residual covariances to exclude the presence of local misspecifications (i.e., cross-loadings and/or correlations between the error terms).

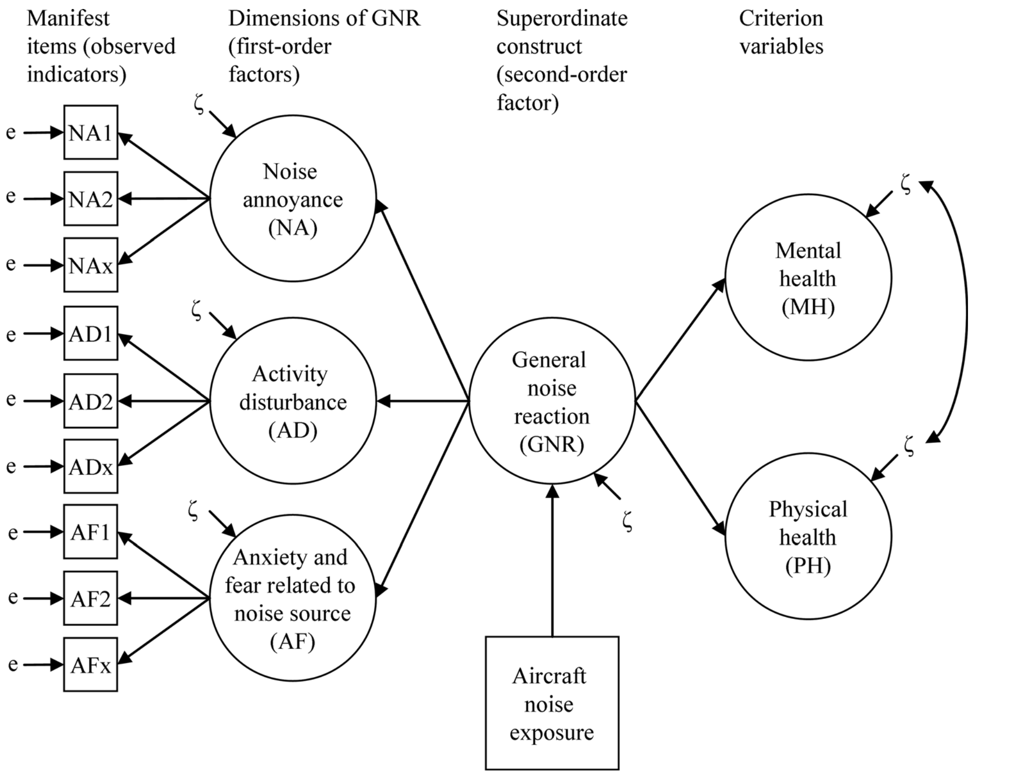

FIG. 2. Superordinate construct model of GNR as a cause of residential satisfaction and perceived health. e: Error term of observed variable and ζ: Error term of latent construct

Third, in line with the strategy outlined by Edwards (2001), the *construct validity* of GNR will be assessed. This will be done in three ways. First, it will be examined whether the specific dimensions (i.e., noise annoyance, activity disturbance, and feelings of anxiety and fear) can explain variance in the two outcome variables, i.e., (self-reported) physical health and mental health, over and above the GNR construct, or whether GNR indeed captures all relevant variance for this prediction. If the latter is the case, this would support the construct validity of the GNR construct; for it would show that variance specific to the dimensions is irrelevant in the prediction of the criterion variables. In other words, only the variance that is common to all dimensions (captured by GNR) is relevant.

Second, it will be assessed whether GNR can effectively mediate the effects of aircraft noise exposure on the dimensions of GNR. If, on the one hand, the direct paths between aircraft noise exposure and the dimensions of GNR are zero (after controlling for the indirect paths via GNR), it can be inferred that variance specific to the dimensions of GNR cannot be attributed to exposure to aircraft noise. It would then be legitimate to ignore specific dimension variance and regard the common variance captured by the second-order construct (GNR) as the only true variance. If, on the other hand, the specific variance in the dimensions *can* be explained by aircraft noise exposure (over and above GNR), dimension specificity can and should not be neglected, since it constitutes subjective variability resulting from exposure to aircraft noise.

Third, it will be assessed whether GNR can effectively mediate the effects of aircraft noise exposure on the criterion variables of GNR, i.e., mental and physical health. This would reflect the idea that only those consciously affected by aircraft noise (as a form of stress) will suffer from (mental and/or physical) health effects due to the noise. If the direct relationships between aircraft noise exposure and the criterion variables are insignificant, it can be inferred that there are no other cognitive mediators present in the noise–health relationship (other than GNR). In other words, GNR is the only relevant cognitive mediator. It should be noted that, however, the presence of direct effects between noise exposure and the criterion variables does not directly invalidate GNR. Previous research has shown that direct effects between noise and health are present for which it is less likely that they are mediated by cognitive variables [see, e.g., Haralabidis *et al.* (2008) who found a significant direct effect between night-time noise exposure and blood pressure]. Hence, the presence of significant direct effects does not necessarily indicate that there are other intervening psychological variables at work in the noise–health relationship.

To summarize, evidence of construct validity of GNR is present when (1) the direct paths from the dimensions to the criterion variables are zero, (2) the direct paths from aircraft noise exposure to the dimensions are zero, and (3) the direct paths from aircraft noise exposure to the criterion variables are zero.

Mental health and physical health are chosen as criterion variables because (1) the concepts lie outside the content domain of noise reaction, (2) the concepts are sufficiently broad to function as relevant criterion variables (Edwards, 2001), and (3) health-related variables have previously been shown to be associated with noise reaction. With respect to the last point it has previously been established that noise reaction is associated with mental health (Van Kamp *et al.*, 2007), blood pressure (Babisch *et al.*, 2007), (self-reported) high blood pressure (Black *et al.*, 2007), antihypertensive treatment use (Neus *et al.*, 1983), and (self-reported) physical health problems (Hatfield *et al.*, 2001). These studies were based on cross-sectional data and so do not provide definitive evidence of causality. It might be that, for example, the direction of causation is opposite, i.e., bad health causes increased noise reaction (Job, 1996; Tarnopolsky *et al.*, 1978). Alternatively, a "third" underlying variable might explain the association between noise reaction and health, e.g., noise sensitivity (Stansfeld, 1992), positive affect (Cohen and Pressman, 2006), or negative affect (Watson, 1988). Theoretical and empirical considerations, however, suggest that a causal relationship exists between psychological reaction and health (Job, 1996). In all, we believe that (self-reported) mental health and physical health are suited as criterion variables to assess the validity of GNR.

The criterion variables are included in the model (see the right side of Fig. 2). Since these variables may covary for reasons other than sharing the superordinate construct as a cause (e.g., they may influence each other or both be influenced by a third variable), the error terms of these constructs are allowed to correlate.

If the results are supportive for the specified factor structure, the question to what extent noise annoyance captures the relevant variance resulting from exposure to aircraft noise can be easily answered through examination of the (standardized) factor loadings of this dimension and its indicators. In addition, the correlation coefficient between GNR and a direct measure of general reaction to noise (dissatisfaction with aircraft noise) will be computed to provide information as to how well this measure performs as an indicator of GNR.

B. Data

The data to test the hypothesized model in Fig. 2 is obtained from an aircraft noise study conducted in Germany at Frankfurt Airport ($N = 2312$), described in Schreckenberg and Meis (2006). This survey was conducted in the period April through December 2005. Within this study a random sample was drawn from residents living in 66 residential areas located within a 40 km radius around the Frankfurt Airport. The selection of these areas resulted from the combination of different strata for the level of noise exposure and four directions from the airport (northwest, east, south, and west). In practice, this resulted in an oversampling of residents living close to the airport. The response rate was 61%. Cases with more than 10% of the values missing are deleted ($N = 106$), and the remaining missing values (1.1%) are imputed via the expectation-maximization algorithm of SPSS.

C. Measures

To exclude the presence of measurement errors at the level of the dimensions they are treated as latent constructs and measured with multiple observed indicators. Table III presents the used indicators of the dimensions of GNR. The items used for noise annoyance (NA1 and NA2) exactly match the standardized noise reaction questions developed by Fields *et al.* (2001). These questions are formulated as follows: (1) "Thinking about the last 12 months or so, when you are here at home, how much does noise from aircraft

TABLE III. The constructs and their indicators (label, range, mean and SD[a]).

Dimension	Observed indicator	Label	Range	Mean	SD
Noise annoyance	Aircraft noise annoyance (past 12 months)	NA1	(0 = not at all annoyed to 10 = extremely annoyed)	4.85	3.17
	Aircraft noise annoyance (past 12 months)	NA2	(1 = not at all, 2 = slightly, 3 = moderately, 4 = very, 5 = extremely)	3.07	1.34
Activity disturbance	Disturbance during relaxation	AD1	(1 = not at all, 2 = slightly, 3 = moderately, 4 = very, 5 = extremely)	2.36	1.32
	Disturbance during conversations/telephone calls	AD2	(1 = not at all, 2 = slightly, 3 = moderately, 4 = very, 5 = extremely)	2.23	1.28
	Disturbance during reading, concentrating	AD3	(1 = not at all, 2 = slightly, 3 = moderately, 4 = very, 5 = extremely)	2.19	1.29
	Disturbance during domestic coziness or visitation	AD4	(1 = not at all, 2 = slightly, 3 = moderately, 4 = very, 5 = extremely)	2.13	1.23
	Disturbance during sleeping	AD5	(1 = not at all, 2 = slightly, 3 = moderately, 4 = very, 5 = extremely)	2.00	1.28
Anxiety and fear related to the noise source	Health threatened by stress caused by aircrafts	AF1	(1 = not, 2 = a little, 3 = reasonably, 4 = fairly, 5 = very)	2.16	1.34
	Hearing threatened by aircraft noise	AF2	(1 = not, 2 = a little, 3 = reasonably, 4 = fairly, 5 = very)	1.92	1.16
	Threatened by the low altitude of over flying aircrafts	AF3	(1 = not, 2 = a little, 3 = reasonably, 4 = fairly, 5 = very)	2.27	1.35
	Threatened by property devaluation due to aircraft noise	AF4	(1 = not, 2 = a little, 3 = reasonably, 4 = fairly, 5 = very)	2.37	1.53
Direct measure					
General reaction to noise	Dissatisfaction with aircraft noise	—	(1 = very satisfied, 2 = fairly satisfied, 3 = rather satisfied, 4 = rather dissatisfied, 5 = dissatisfied) (recoded)	3.63	1.03
Cause					
Aircraft noise exposure	L_{den} dB(A)	L_{den}	(43.8–70.3)	56.9	6.87

[a]Standard deviation (SD).

noise bother, disturb, or annoy you?" and (2) "Next is a zero to ten opinion scale for how much aircraft noise bothers, disturbs, or annoys you when you are here at home. If you are not at all annoyed choose zero, if you are extremely annoyed choose ten, if you are somewhere in between choose a number between zero and ten. Thinking about the last 12 months or so, what number from zero to ten best shows how much you are bothered, disturbed, or annoyed by aircraft noise?" (Fields *et al.*, 2001).

The direct measure of general reaction to noise (presented in the second last row of Table III) related to the question "How satisfied are you with respect to the environmental condition of aircraft noise?" (Responses are recoded.)

For each address individual aircraft noise levels were calculated on the basis of flight movements of the six busiest months of the year 2005 according to the German regulation for aircraft noise calculation. For the present study L_{den} [i.e., level day–evening–night in dB(A)] is selected as a measure of the level of aircraft noise exposure.[2] Univariate statistics for this measure are presented in the last row of Table III.

Physical health and mental health are measured using the previously validated scales of the 12-item Short Form (SF-12) Health Survey (Ware *et al.*, 1996). The SF-12 Health Survey constitutes a subset of 12 items from the SF-36 Health Survey (with 36 items) and covers eight health concepts: Physical functioning, role limitations due to physical health, bodily pain, general health, vitality, social functioning, role limitations due to emotional problems, and mental health (Ware and Sherbourne, 1992). Studies of the factor structure of the SF-36 Health Survey consistently revealed two underlying dimensions of these concepts, namely physical health and mental health (Ware *et al.*, 1998). This factor structure (eight first-order factors and two second-order factors) has been validated in a confirmatory factor model with data from multiple countries (Keller *et al.*, 1998). It has been shown that the SF-12 Health Survey can adequately reproduce the physical and mental components of the SF-36 (Gandek *et al.*, 1998). For a German sample, correlations between the SF-12 and the SF-36 summary measures were 0.96 and 0.94 for the physical and mental health summary measures, respectively (Gandek *et al.*, 1998). For the present study physical and mental component measures are computed based on population normative data derived from a German sample (Bullinger and Kirchberger, 1998). These measures are included as observed indicators of two respective latent variables (i.e., physical and mental health) in the structural equation model. The reliability of the latent variables is fixed based on previously observed test–retest correlations of 0.89 and 0.76 for the physical and mental health summary measures, respectively (Ware *et al.*, 1996). In effect, the parameter estimates of the paths between GNR and the criterion variables are corrected for random measurement errors.

D. Estimation procedure

With the exception of aircraft noise exposure the indicators in Table III are measured on ordinal scales. For these measures polychoric correlations are computed. Compared to three other types of correlations (e.g., Pearson, Spearman, and Kendall) the polychoric correlation has been shown to be the least biased in the case of ordinal variables (Jöreskog and Sörbom, 1996). In addition, this bias becomes negligible for moderate to large sample sizes (Jöreskog and Sörbom, 1996). Finally, the polychoric correlation estimates have been shown to be robust to moderate violations of normality of the assumed underlying continuous variables (Flora and Curran, 2004).

Substituting the polychoric correlation matrix with the product-moment correlation matrix and applying the usual maximum likelihood estimation function will yield consistent parameter estimates, but incorrect test statistics and standard errors. In response, the weighted least squares (WLS) approach has been developed to yield unbiased estimates and standard errors (Browne, 1984). In this study, robust WLS approach is used to estimate the model. Based on the results of a simulation study Flora and Curran (2004) concluded that this estimation method performs well under various conditions (i.e., at varying sample sizes, underlying distributions of the continuous variables, numbers of indicators, and numbers of categories of the indicators). The authors recommended its use especially for medium-to-large models with ordinal variables.

The polychoric correlation matrix and the asymptotic covariance matrix, which are necessary for the use of the robust WLS method, are calculated in PRELIS 2, and LISREL 8.8 is used to estimate the structural equation model.

IV. RESULTS

A. Model fit

Based on the fit statistics it can be concluded that the data set fits the second-order factor structure described in Fig. 2 well. All statistics are above or below their respective lower and upper limits ($\chi^2_{df=73} = 430.8$, RMSEA $= 0.042 < 0.06$, SRMR $= 0.036 < 0.08$, and CFI $= 0.99 > 0.95$).

B. Convergent validity

The estimated factor loadings are all significant ($p < 0.001$) and exceed the preferable minimum criterion of 0.70. Hence, the measures and the dimensions converge on their hypothesized underlying constructs. Additional evidence for the convergent validity of the hypothesized model is provided by the variance-extracted estimates and the construct reliabilities, which are presented within each latent construct in Fig. 2. The average variance-extracted estimates are above the conventional minimum criterion of 50% and the construct reliabilities all exceed the minimum criterion of 0.70. Taken together, the evidence supports the convergent validity of the measurement model.

A review of the residual covariances shows that the fit of the model cannot be substantially improved by adding additional parameters. Therefore the presence of local misspecifications—i.e., covariances between the error terms of the measures/dimensions and cross-loadings (i.e., factor loadings of indicators on constructs other than the one they were intended to measure)—can be excluded. This means

TABLE IV. Performance of noise annoyance and a direct measure of general negative reaction in comparison with GNR as a superordinate construct.

		Explained variance (%)	
		Mental health	Physical health
GNR (superordinate construct; Fig. 3)		7.6	3.4
Dimension	Factor loading		
Noise annoyance	0.98	5.3	1.9
Observed indicator	Factor loading (indirect)[a]		
NA1	0.93	5.3	1.8
NA2	0.92	4.0	1.5
Observed indicator	Correlation with GNR		
General reaction to noise	0.85	6.1	2.8

[a]Computed by taking the product of the factor loading of the observed indicator on noise annoyance and the factor loading of noise annoyance on GNR.

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 GNR as superordinate construct is the sole cause for the structural (common) variance in the three dimensions.

C. Construct validity

The third step in the adopted validation approach is to assess the significance of (1) the direct effects of the dimensions of GNR on the included criterion variables ($4 \times 2 = 8$ paths); (2) the direct effects of aircraft noise exposure on the dimensions of GNR ($1 \times 4 = 4$ paths); and (3) the direct effects of aircraft noise exposure on the criterion variables ($1 \times 2 = 2$ paths). For this purpose the modification indices related to these paths are reviewed. These indices indicate the expected drop in the χ^2 test statistics if an additional path is drawn.

To avoid capitalization on chance in finding a significant effect while there is in fact none (i.e., a type I error), the usually adopted significance level of 0.05 is divided by the number of modification indices examined (i.e., 14), yielding a critical p-value of 0.00357 and a corresponding χ^2 value of 8.49. Based on this criterion it is concluded that none of the reviewed modification indices are significant.

Substantively, this means that (1) there is no specific variance in the dimensions which can be used to explain additional variance in the criterion variables (over and above the variance explained by GNR); (2) dimension specific variance is unrelated to aircraft noise exposure and thus originates from another source; and (3) GNR effectively mediates the relationship between the aircraft noise exposure and the criterion variables. The findings are supportive of the conclusion that GNR captures all the relevant subjective variability resulting from exposure to aircraft noise.

D. Summary

Based on the overall model fit and the convergent and construct validity it is concluded that GNR as a superordinate construct is a valid conceptualization and indeed measures what it is intended to measure. Substantively, this supports the notion that people develop a general attitude of aircraft noise which is reflected in particular responses to the noise. The variance which is shared by the dimensions of GNR is effectively captured by the superordinate construct. Only this shared variance, and not variance specific to the dimensions, is relevant in the prediction of two criterion variables.

E. Performance of noise annoyance and a general measure of negative reaction

Now that the validity of GNR as a latent superordinate construct is sufficiently supported it can be examined how well noise annoyance, its indicators, and a direct measure of general reaction to noise (i.e., dissatisfaction with the noise) perform as manifestations of this construct. This information is presented in Table IV (Fig. 3).

It can be concluded that noise annoyance is a strong reflection of GNR (factor loading of 0.98). Hence, a large part of all negative feelings and emotions in response to the aircraft noise is captured by noise annoyance. However, as expected, noise annoyance does not reflect all relevant variances. This also becomes apparent when noise annoyance is used as a sole determinant in predicting the two criterion variables. The percentages of explained variance in these variables are substantially lower in comparison to GNR as determinant.

Examination of the factor loadings and proportions of explained variance for the indicators of noise annoyance (NA1 and NA2) clearly shows the additional effect of measurement error (on top of dimension specificity) which is present in the observed indicators and which biases the structural estimates of the relationships between the latent variables. In comparison to the noise annoyance dimension the indicators are (by definition) weaker reflections of GNR and, when used as direct determinants of the criterion variables, can also explain less variance in these variables. Hence, the measurement error present in these observed indicators suppresses the real associations between the factors. Preferably, the observed indicators should therefore not be used in isolation.

Finally, the performance of the direct measure of general reaction to noise (dissatisfaction with aircraft noise) is assessed. This measure is also strongly correlated with GNR ($r = 0.85$). In addition, bearing in mind that this measure is also "contaminated" with measurement error, its performance

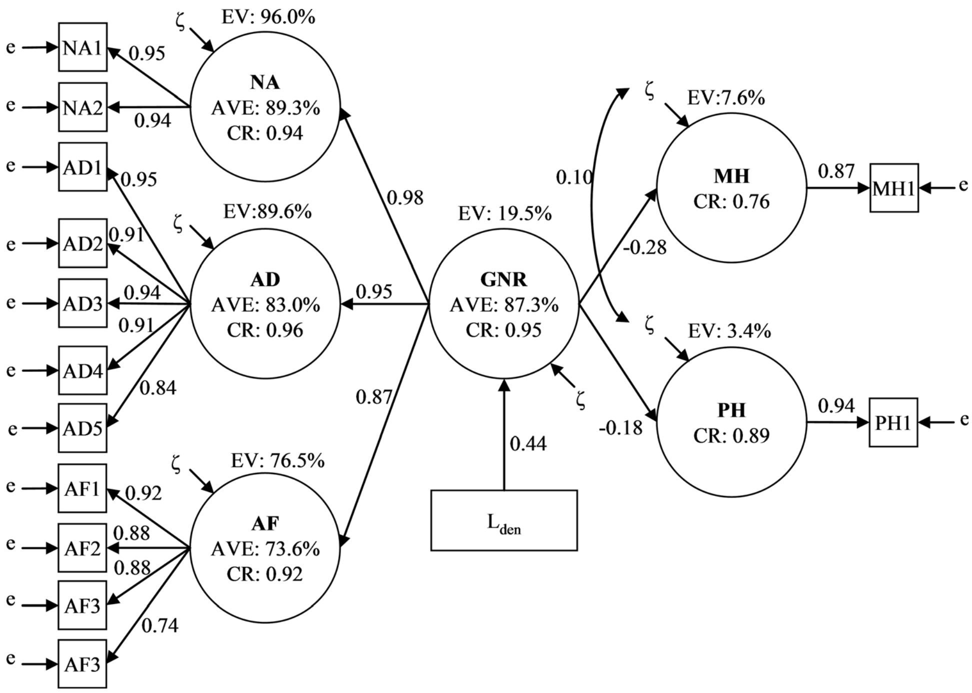

FIG. 3. Standardized solution of the superordinate construct model ($N = 2206$). Fit statistics: $\chi^2 = 430.8$, $df = 73$, RMSEA = 0.042, SRMR = 0.036, CFI = 0.99. EV: Explained variance; AVE: Average variance extracted; CR: Construct reliability; e: Error term of observed variable; and ζ: Error term of latent construct.

in terms of the explained variance in the criterion variables is remarkable. Fitting the previous results of Job *et al.* (2001) the conclusion can be drawn that measures of general reaction perform better than specific measures of noise reaction, such as noise annoyance.

V. CONCLUSION

In this study a measurement model is developed to measure general negative reaction to aircraft noise. Estimation of the model yielded a good fit to the data and supported the second-order factor structure. Additional support for the specified structure is found in the convergence and construct validity of GNR as a superordinate construct. Based on the factor loadings on GNR it is inferred that noise annoyance and its observed indicators are strong reflections of this construct but do not capture all relevant variances. To a lesser extent the same conclusion holds for a direct measure of general negative reaction to noise, i.e., dissatisfaction with aircraft noise.

The results of the present study are in line with the previous findings of Job *et al.* (2001) and indicate that general measures are more valid indicators of negative reaction to (aircraft) noise than specific dimensions such as annoyance or disturbance. Therefore, we also endorse their recommendation to include such general measures in future community surveys. In addition, we advise the use of multiple observed indicators to control for random measurement errors. As the present study has confirmed, these random errors suppress the real associations between the constructs of interest.

As a by-product of our approach the developed model provides us insight into the overall experience of aircraft noise. Based on the results it is apparent that this experience is multifacetted and includes at least three, but possibly many other, dimensions. In addition, from the factor loadings on GNR it can be inferred that dimensions such as noise annoyance and activity disturbance lie at the core of GNR, while the anxiety and fear dimension operates at a more distant level.

Based on the results of the present study several interesting directions for future research can be identified. One would be to explore additional dimensions of general negative reaction such as perceived control or the attitude toward the noise source authorities. Using the approach followed in this study it can be assessed whether such factors also form an integral part of people's general reaction toward aircraft noise or whether these should be viewed as independent variables.

A second research direction is to explore the factor structure for different subsets of the population. 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 be hypothesized that within the former group reactions like fear plays 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 differs for this group and also that it might be differently related to the criterion variables.

To conclude, we emphasize that reaction to aircraft noise must not be understood as a narrow experience like noise annoyance but as an aspect of a broader multidimensional construct that compromises many other feelings, emotions, and beliefs related to aircraft noise in the residential environment.

ACKNOWLEDGMENTS

The authors would like to thank two anonymous reviewers for their valuable comments on an earlier draft of this paper.

[1]The study of Job *et al*. (2001) was based on a relatively small number of subjects ($N = 97$) in a change situation. Their results with respect to the reliability and validity of general reaction measures should therefore be interpreted with care.

[2]L_{den} is an equivalent sound level of 24 h expressed in decibels (dB) on the "A" weighted scale dB(A). Sound levels during the evening (7 p.m. to 11 p.m.) and during the night (11 p.m. to 7 a.m.) are increased by a penalty of 5 and 10 dB(A), respectively.

Ahrlin, U. (**1988**). "Activity disturbances caused by different environmental noises," J. Sound Vib. **127**, 599–603.

Alexandre, A. (**1976**). "An assessment of certain causal models used in surveys on aircraft noise annoyance," J. Sound Vib. **44**, 119–125.

Babisch, W., Houthuijs, D., Pershagen, G., Katsouyanni, K., Velonakis, M., Cadum, E., Dudley, M. L., Bluhm, G., Breugelmans, O., Charalampidis, A., Dimakopoulou, K., Savigny, P., Seiffert, I., Selander, J., Sourtzi, P., Swart, W., Vigna-Taglianti, F., and Järup, L. (**2007**). "Association between noise annoyance and high blood pressure. Preliminary results from the HYENA study," *Proceedings of Internoise Conference*, Istanbul, Turkey.

Bagozzi, R. P., and Edwards, J. R. (**1998**). "A general approach for representing constructs in organizational research," Organ. Res. Methods **1**, 45–87.

Bentler, P. M. (**1990**). "Comparative fit indexes in structural models," Psychol. Bull. **107**, 238–246.

Black, D. A., Black, J. A., Issarayangyun, T., and Samuels, S. E. (**2007**). "Aircraft noise exposure and resident's stress and hypertension: A public health perspective for airport environmental management," J. Air Transp. Manag. **13**, 264–276.

Bollen, K., and Lennox, R. (**1991**). "Conventional wisdom on measurement—A structural equation perspective," Psychol. Bull. **110**, 305–314.

Bollen, K. A. (**1989**). *Structural Equations with Latent Variables* (Wiley, New York), pp. 1–514.

Bröer, C. (**2006**). *Beleid vormt overlast, hoe beleidsdiscoursen de beleving van geluid bepalen (Policy annoyance, how policy discourses shape the experience of aircraft sound)* (Aksant, Amsterdam), pp. 1–271.

Bröer, C., and Duyvendak, J. W. (**2009**). "Discursive opportunities, feeling rules, and the rise of protests against aircraft noise," Mobilization **14**, 337–356.

Browne, M. W. (**1984**). "Asymptotically distribution-free methods for the analysis of covariance structures," Brit. J. Math. Stat. Psychol. **37**, 62–83.

Browne, M. W., and Cudeck, R. (**1993**). "Alternative ways of assessing model fit," in *Testing Structural Equation Models*, edited by K. A. Bollen, and J. S. Long (Sage Publications, Newbury Park, CA), pp. 136–162.

Bullinger, M., and Kirchberger, I. (**1998**). *SF-36 Fragebogen zum Gesundheitszustand—Manual* (The SF-36 health survey–manual) (Hogrefe, Göttingen).

Cohen, S., and Pressman, S. D. (**2006**). "Positive affect and health," Curr. Dir. Psychol. Sci. **15**, 122–125.

Edwards, J. R. (**2001**). "Multidimensional constructs in organizational behavior research: An integrative analytical framework," Organ. Res. Methods **4**, 144–192.

Fidell, S., Barber, D. S., and Schultz, T. J. (**1991**). "Updating a dosage-effect relationship for the prevalence of annoyance due to general transportation noise," J. Acoust. Soc. Am. **89**, 221–233.

Fields, J. M., De Jong, R. G., Gjestland, T., Flindell, I. H., Job, R. F. S., Kurra, S., Lercher, P., Vallet, M., Yano, T., Guski, R., Felscher-Suhr, U., and Schumer, R. (**2001**). "Standardized general-purpose noise reaction questions for community noise surveys: Research and a recommendation," J. Sound Vib. **242**, 641–679.

Flora, D. B., and Curran, P. J. (**2004**). "An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data," Psychol. Methods **9**, 466–491.

Fornell, C., and Bookstein, F. L. (**1982**). "Two structural equation models—Lisrel and PLS applied to consumer exit-voice theory," J. Marketing Res. **19**, 440–452.

Gandek, B., Ware, J. E., Aaronson, N. K., Apolone, G., Bjorner, J. B., Brazier, J. E., Bullinger, M., Kaasa, S., Leplege, A., Prieto, L., and Sullivan, M. (**1998**). "Cross-validation of item selection and scoring for the SF-12 Health Survey in nine countries: Results from the IQOLA Project," J. Clin. Epidemiol. **51**, 1171–1178.

Guski, R. (**1999**). "Personal and social variables as co-determinants of noise-annoyance," Noise Health **3**, 45–56.

Guski, R., Felscher-Suhr, U., and Schuemer, R. (**1999**). "The concept of noise annoyance: How international experts see it," J. Sound Vib. **223**, 513–527.

Hall, F. L., Taylor, S. M., and Birnie, S. E. (**1985**). "Activity interference and noise annoyance," J. Sound Vib. **103**, 237–252.

Haralabidis, A. S., Dimakopoulou, K., Vigna-Taglianti, F., Giampaolo, M., Borgini, A., Dudley, M. L., Pershagen, G., Bluhm, G., Houthuijs, D., Babisch, W., Velonakis, M., Katsouyanni, K., Jarup, L., and Consortium, H. (**2008**). "Acute effects of night-time noise exposure on blood pressure in populations living near airports," Eur. Heart J. **29**, 658–664.

Hatfield, J., Job, R. F. S., Carter, N. L., Peploe, P., Taylor, R., and Morrell, S. (**2001**). "The influence of psychological factors on self-reported physiological effects of noise," Noise Health **3**, 1–13.

Hede, A. J., Bullen, R. B., and Rose, J. A. (**1979**). *A Social Study of the Nature of Subjective Reaction to Aircraft Noise*, edited by A. G. P. Services (National Acoustic Laboratories, Canberra).

Hu, L., and Bentler, P. M. (**1999**). "Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives," Struct. Equation Model. **6**, 1–55.

Jarvis, C. B., MacKenzie, S. B., and Podsakoff, P. M. (**2003**). "A critical review of construct indicators and measurement model misspecification in marketing and consumer research," J. Consum. Res. **30**, 199–218.

Job, R. F. S. (**1996**). "The influence of subjective reactions to noise on health effects of the noise," Environ. Int. **22**, 93–104.

Job, R. F. S., Hatfield, J., Carter, N. L., Peploe, P., Taylor, R., and 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.

Job, R. F. S., and Sakashita, C. (**2007**). "Assessment of community reaction to noise: Methodology in question wording and delivery," *Proceedings of the Internoise Conference*, Istanbul, Turkey.

Jöreskog, K. G., and Sörbom, D. (**1996**). *Prelis 2: User's Reference Guide* (Scientific Software International, Lincolnwood, IL), pp. 1–224.

Keller, S. D., Ware, J. E., Bentler, P. M., Aaronson, N. K., Alonso, J., Apolone, G., Bjorner, J. B., Brazier, J., Bullinger, M., Kaasa, S., Leplege, A., Sullivan, M., and Gandek, B. (**1998**). "Use of structural equation modeling to test the construct validity of the SF-36 Health Survey in ten countries: Results from the IQOLA Project," J. Clin. Epidemiol. **51**, 1179–1188.

Kroesen, M., and Bröer, C. (**2009**). "Policy discourse, people's internal frames and declared aircraft noise annoyance: An application of Q methodology," J. Acoust. Soc. Am. **126**, 195–207.

Kroesen, M., Molin, E. J. E., and Van Wee, B. (**2008**). "Testing a theory of aircraft noise annoyance: A structural equation analysis," J. Acoust. Soc. Am. **123**, 4250–4260.

Kryter, K. D. (**1982**). "Community annoyance from aircraft and ground vehicle noise," J. Acoust. Soc. Am. **72**, 1222–1242.

Law, K. S., Wong, C. S., and Mobley, W. H. (**1998**). "Toward a taxonomy of multidimensional constructs," Acad. Manage. Rev. **23**, 741–755.

Leonard, S., and Borsky, P. N. (**1973**). "A causal model for relating noise exposure, psychosocial variables and aircraft noise annoyance," in *Proceedings of the International Congress on Noise as a Public Health Problem*, pp. 691–705.

Lercher, P. (**1996**). "Environmental noise and health: An integrated research perspective," Environ. Int. **22**, 117–129.

McKennell, A. C. (**1963**). *Aircraft Noise Annoyance Around Heathrow Airport* (Her Majesty's Stationery Office, London).

Miedema, H. M. E., and Vos, H. (**1999**). "Demographic and attitudinal factors that modify annoyance from transportation noise," J. Acoust. Soc. Am. **105**, 3336–3344.

Neus, H., Ruddel, H., and Schulte, W. (**1983**). "Traffic noise and hypertension—An epidemiological-study on the role of subjective reactions," Int. Arch. Occup. Environ. Health **51**, 223–229.

Nunnally, J. C. (**1978**). *Psychometric Theory* (McGraw Hill, New York), pp. 1–640.

Schreckenberg, D., and Meis, M. (**2006**). *Belästigung durch Fluglärm im Umfeld des Frankfurter Flughafens. Gutachten im Auftrag des Regionalen Dialogforums Flughafen Frankfurt (Endbericht)* [Aircraft noise annoyance in the vicinity of Frankfurt airport. Advice on behalf of the Regional Dialogue Forum Frankfurt (final report)] (AG Fluglärmwirkung, Bochum, Oldenburg), pp. 1–340.

Schultz, T. J. (**1978**). "Synthesis on social surveys on noise annoyance," J. Acoust. Soc. Am. **64**, 377–405.

Stansfeld, S. A. (**1992**). "Noise, noise sensitivity and psychiatric disorder: Epidemiological and psychophysiological studies," Psychol. Med. **22**, 1–44.

Tarnopolsky, A., Barker, S. M., Wiggins, R. D., and McLean, E. K. (**1978**). "The effect of aircraft noise on the mental health of a community sample: A pilot study," Psychol. Med. **8**, 219–233.

Taylor, S. M. (**1984**). "A path model of aircraft noise annoyance," J. Sound Vib. **96**, 243–260.

Van Kamp, I., Houthuijs, D., Van Wiechen, C., and Breugelmans, O. (**2007**). "Environmental noise and mental health: Evidence from the Schiphol monitoring program," *Proceedings of Internoise Conference*, Istanbul, Turkey.

Ware, J. E., Kosinski, M., Gandek, B., Aaronson, N. K., Apolone, G., Bech, P., Brazier, J., Bullinger, M., Kaasa, S., Leplege, A., Prieto, L., and Sullivan, M. (**1998**). "The factor structure of the SF-36 Health Survey in 10 countries: Results from the IQOLA Project," J. Clin. Epidemiol. **51**, 1159–1165.

Ware, J. E., Kosinski, M., and Keller, S. D. (**1996**). "A 12-item short-form health survey—Construction of scales and preliminary tests of reliability and validity," Med. Care **34**, 220–233.

Ware, J. E., and Sherbourne, C. D. (**1992**). "The Mos 36-Item Short-Form Health Survey (SF-36).1. Conceptual-framework and item selection," Med. Care **30**, 473–483.

Watson, D. (**1988**). "Intraindividual and interindividual analyses of positive and negative affect—Their relation to health complaints, perceived stress, and daily activities," J Pers. Soc. Psychol. **54**, 1020–1030.