**Assessing community noise annoyance: a review of two decades of the international technical specification ISO/TS 15666:2003**

Charlotte Clark,1 Truls Gjestland,2 Lisa Lavia,3 Hilary Notley,4 David Michaud,5and Makoto Morinaga6

1 Population Health Research Institute, St George’s, University of London, London, SW17 0RE, United Kingdom

2 SINTEF Digital, P.O. Box 4760 Torgarden, N-7465 Trondheim, Norway  
3 Noise Abatement Society, Brighton, East Sussex, BN2 9QA, United Kingdom and The Urban Institute, Heriot-Watt University, Edinburgh, EH14 4AS, United Kingdom

4 UK Department for the Environment, Food and Rural Affairs (Defra), Ground Floor, Seacole Building, 2 Marsham Street, London SW1P 4DF, United Kingdom

5 Health Canada, Environmental and Radiation Health Sciences Directorate, Consumer & Clinical Radiation Protection Bureau, Ottawa, ON, Canada

6 Kanagawa University, 3-27-1 Rokkakubashi, Kanagawa-ku, Yokohama, Kanagawa, 221-8686, Japan

Abstract

193/200

The robust assessment of noise annoyance is of key importance given that it is the most prevalent community response in populations exposed to environmental noise. In 1993 the International Commission on Biological Effects of Noise Community Response to Noise team began formalizing a standardized methodology for assessing noise annoyance, which resulted in reporting guidelines and recommendations later published as a Technical Specification (TS) by the International Standards Organization in 2003 (ISO/TS 15666: 2003). This TS, intended to inform the international community on the quantification of the exposure-response relationship between noise exposure and annoyance, has been in circulation for nearly two decades and has been updated in 2021 (ISO/TS 15666: 2021) by an international working group (ISO TC43/SC1/WG62). This paper reviews use of the 2003 TS, identifies common adaptations in use and summarizes the revisions. Methodological issues arising from the use of the 5-point verbal and the 11-point numeric scale questions and the scoring of ‘highly annoyed’ are discussed. The revisions are designed to encourage further standardization in noise annoyance research. The paper highlights research needs that if addressed would strengthen the methodology underlying the assessment of noise annoyance including multidimensional assessments of annoyance.

1. INTRODUCTION

Annoyance is one of the most prevalent reactions among populations exposed to environmental noise, with a long history as the primary metric for quantifying the community response to a wide range of sources including aircraft, road traffic, wind turbines, industry, and railway noise (World Health Organization, 2011; 2018). The assessment of noise annoyance in the home environment has been governed by a Technical Specification (TS) ISO/TS 15666 ‘Acoustics — Assessment of noise annoyance by means of social and socio-acoustic surveys’ since 2003 (ISO/TS15666:2003, 2003), wherein annoyance is defined as one person’s adverse reaction to noise. A TS is not an International Standard and addresses work still under technical development, published for immediate use, and as a means to obtain feedback (ISO, 2021). The aim is that a TS will eventually be transformed and republished as an International Standard.

Prior to the TS, there had been several key developments for the methodological assessment of annoyance over the past few decades starting with the formative Schultz-curves in 1978 (Schultz, 1978), which were the first to synthesize exposure-response relationships between noise exposure and annoyance responses across studies. At a time before meta-analyses were mainstream, Schultz’s paper was considered controversial (Kryter, 1982b; Schultz, 1982; Kryter, 1983; Fidell, 2003) and highlighted many of the difficulties of combining individual study analyses. However, many of the decisions Schultz made about how to synthesize data across annoyance surveys remain with us today and have become conventional wisdom.

Schultz developed exposure-effect functions, also known as dose-response curves, that showed the percentage of people “highly annoyed by noise” as a function of the noise exposure described by the day-night weighted equivalent level, DNL. To be able to compare findings across individual studies using different response scales, Schultz adopted, as a basic rule, that people who responded to the upper 27%-29% of a numeric annoyance scale could be considered “highly annoyed”. The “percent highly annoyed” was chosen as a measure of community response as Schultz considered that the main purpose of surveys on noise annoyance was to seek guidance for regulatory decisions about noise. He argued that when people are highly annoyed by noise, the effects of non-acoustical variables on annoyance, such as attitudes, are reduced, and the correlation between the noise exposure and the expressed subjective reaction is high. Also, the term “highly annoyed” carried a commonsense import that was clear and would be readily understood by politicians and policymakers. It avoided trivializing the concept of annoyance. Others had argued for a measure of central tendency rather than an extreme measure of the distribution of annoyance (Kryter, 1982a).

The US Federal Interagency Committee on Noise (FICON) (1992), declared that annoyance was its preferred summary measure of the general adverse reaction of people to noise, and that the percentage of the area population characterized as "highly annoyed" by long-term exposure to noise was its preferred measure of annoyance. This view, originally proposed by Schultz, thus became more or less a *de facto* standard for reporting adverse reactions to noise.

In his paper Schultz discussed the difficulties of combining results from surveys conducted in different languages according to very different test protocols, using different questionnaires, different noise metrics, and different response scales. In order to simplify future inter-survey comparisons, in 1993 the International Commission on Biological Effects of Noise (ICBEN) Community Response to Noise team commenced seven years of research to formalize a standardized methodology for assessing noise annoyance, which included the development of reporting guidelines for socio-acoustic surveys in 1997 (Fields *et al.*, 1997) and recommendations for standard noise reaction questions in 2001 (Fields *et al.*, 2001). The group focused on requirements that would aid secondary analysis of findings across studies, rather than inform individual studies, *per se*.

The ICBEN group developed a noise reaction measure for social surveys that had several characteristics (Fields *et al.*, 2001). It enabled a valid international comparison of survey results within and between languages; provided a high quality and reliable measure of a general reaction to noise experienced in a residential environment; gave transparent results that could be consistently interpreted by survey respondents and policy makers; and was suitable for all questionnaire administration modes that were prevalent at that time including face-to-face, telephone, and self-administered. The methodology also provided an interval-level measurement scale that could be used in regression analyses and other techniques; with the aim that the approach would be adopted internationally.

The ICBEN team had no intention to develop a standardized procedure for conducting community noise surveys nor did they opt for a standardized questionnaire. They settled for a minimum solution: two standardized questions that could be included in any survey, and that would provide a means of transformation of results from one survey to another and act as a common calibration of the responses.

Two noise reaction questions were recommended. The 5-point verbal scale question was specified 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: with responses give as *Extremely*; *Very*; *Moderately*; *Slightly*; or *Not at all*”. The 11-point numeric scale question included a preamble to the following question: “What number from zero to ten best shows how much you are bothered, disturbed or annoyed by (..source..) noise? Although it is common to evaluate annoyance integrated over a time that would ensure a stable noise situation was being considered (e.g., 12 months or so), there are provisions for this time to be substituted as “the last week or so” or “the last month or so”, as needed.

The resulting method developed by the ICBEN team was later published by the International Organization for Standardization (ISO) as a TS in 2003 - ISO/TS 15666:2003. It has been used to quantify exposure-response relationships between measured or modelled noise exposure and a respondents’ self-reported annoyance at home for a variety of environmental noise exposures covering individual studies of aircraft, road traffic, railway, wind turbines and other industrial noise sources (Brink *et al.*, 2008; Michaud *et al.*, 2016b; Civil Aviation Authority, 2017; Murakami *et al.*, 2018; Brink *et al.*, 2019; Gjestland *et al.*, 2019; Morinaga *et al.*, 2020; Miller *et al.*, 2021a; Miller *et al.*, 2021b), inter alia, as well as in the synthesis of findings across studies (Miedema and Oudshoorn, 2001; European Commission, 2002; Guski *et al.*, 2017) that are frequently used by governments to estimate noise impacts and their potential health effects, and to inform mitigation strategies.

The use of the term ‘estimate’ is deliberate, to acknowledge issues with the precision of quantifying annoyance. Annoyance varies widely for the same noise exposure which could potentially lead to underestimates of annoyance or overly restrictive policies and fuel debate between stakeholders as to the ‘true’ level of annoyance for their context.

Annoyance, and in particular the percentage highly annoyed (%HA), was a key outcome, alongside sleep disturbance informing the World Health Organization’s Environmental Noise Guidelines for the European Region (World Health Organization, 2018), and these outcomes are used to estimate burden of disease (World Health Organization, 2011; European Environment Agency, 2020). Many countries use the %HA in impact assessment for new infrastructure schemes to estimate the health impact of change in noise exposure and/or to monetize the effects of noise on health (Department for Transport, 2013; Kuwano *et al.*, 2014; Health Canada, 2017; van Kamp *et al.*, 2018; van Kamp *et al.*, 2020).

Different strategies may be used to transform the survey results into exposure-response curves as explained in the international standard ISO 1996-1:2016 (ISO 1996-1:2016, 2016). This standard has informative annexes for estimating prevalence of a population highly annoyed as a function of adjusted day-night or day-evening-night sound level using the Community Tolerance Level method (CTL) or a polynomial regression method.

ISO/TS 15666:2003 has now been in operation for nearly 20 years and the aim of this paper is to 1) review its use and recent evidence from the field; 2) discuss the recent technical revision of the TS; and 3) highlight research needs and gaps that, if addressed in the next few years, would help to further strengthen the methodology underlying the assessment of noise annoyance.

The work presented in this paper has been undertaken by an international working group (WG) convened by the United Kingdom, as a work item within ISO TC43/SC1/WG62 entitled “Revision of ISO/TS 15666 – Assessment of noise annoyance by means of social and socio-acoustic surveys”. The team was tasked with revising the TS between 2017-2021 and the updated TS was published in May 2021(ISO/TS15666:2021, 2021 ).

1. THE HISTORICAL APPLICATION OF ISO/TS 15666:2003

The Health Canada librarians performed a literature search in Scopus, Embase, Medline, and Ovid using the following initial search terms “noise” “annoyance” “bother” “disturb” “ISO TS/15666”, “assessment of noise annoyance by means of social and socio-acoustic surveys”, with date restrictions being 2002 to 2021. The initial search resulted in 416 references. Titles and abstracts were used to further screen these references for relevance by one of the authors (DM), resulting in 240 papers that were further reviewed by two of the authors (DM and TG) to evaluate their degree of compliance with the TS. Both DM and TG developed basic metrics to evaluate the studies against, primarily focusing on if both the verbal and numeric questions were used, the level of compliance with the recommended wording in questions and response categories, if a filter question was used based on the audibility of the source being evaluated, and if one’s annoyance at home was evaluated. It should be emphasized that the TS includes many other reporting requirements, however the purpose of the literature screening was primarily to focus on the questions themselves. As this was not intended to serve as a systematic review, DM and TG each reviewed half of the remaining 240 relevant references. Although the TS is specifically intended for social surveys, some of the papers described laboratory studies, where the response scales, either verbal or numerical, were typically adopted, but the rest of the standard was not applied. Overall, a large number of social survey studies did not provide sufficient information in their methods to determine compliance with the TS. A small number of studies amended the TS by including a filter question related to the audibility of a noise source prior to asking about annoyance.

A striking observation was that the ICBEN recommendations by Fields *et al.* (1997) and adopted in the TS for reporting core information from social surveys were followed by only a small number of researchers. The questionnaire, for instance, was only provided by a few of the surveys. Another observation was the mismatch between the reference period for annoyance, typically *twelve months or so,* and the noise exposure period that could be anything from LAeq, 1h, Lday, etc.

Few surveys used the recommended two questions with both a verbal and a numerical scale question. In addition, some surveys used the two questions with a slightly modified wording. There seems to be a preference to include only one of the two questions, with no clear preference: where only one question was used there was equal split between use of the 5-point verbal scale question and the 11-point numeric scale question.

There was a general tendency for the results of the survey to be reported as the %HA, although the manner in which %HA was determined was variable across studies, and quite often not defined at all. This supports the provision of guidance on determining %HA in the updated ISO/TS 15666:2021, even though it was not part of the original publication.

The TS has only been used to its full extent in a very limited number of cases but this might be expected given its status as a TS and not an International Standard. A TS is intended to provide feedback on the methodology to inform the further development of a Standard. Nevertheless, it may be concluded that ISO/TS 15666:2003 has had a clear impact on the way social surveys are being conducted, and the main purpose to facilitate comparison of results across different surveys has been accomplished.

Overall, the review indicated that some common adaptations have occurred in use that may not meet the intent of the original TS. Common adaptations include applying the TS to settings outside the home (e.g., offices, public spaces, laboratories); adding a ‘do not hear’ question prior to the question to exclude respondents from answering the noise reaction questions; administering only one of the questions; changing the wording or presentation such as referring to a certain period of the day or night; and not providing the checklist of reporting requirements that forms part of the TS.

The following sections discuss the key revisions in ISO/TS 15666:2021 covering the format and presentation of the questions; and scoring annoyance. It also discusses further changes and clarifications to aid and encourage consistency in the application and use of the TS and harmonization of results.

1. FORMAT AND PRESENTATION OF THE QUESTIONS
2. Two noise reaction questions

The ICBEN team proposed that two noise reaction questions (the 5-point verbal scale question and the 11-point numeric scale question) should be used in a survey as this was psychometrically robust; offered greater reliability of assessment; and enabled comparison between studies (Fields *et al.*, 1997; Fields *et al.*, 2001). It also facilitates the calculation of %HA, the endpoint that is of most interest internationally.

The formats of the response scales were studied in depth by the ICBEN team. For the verbal scale, a 3-point scale was rejected as not providing a sufficient range of alternatives. On a 3-point scale there would only be two degrees of annoyance for those who were other than “not at all annoyed”. Verbal scales of six points or more were rejected because such long scales were judged to be too cumbersome for telephone interviews. Among the two remaining options some research suggested (Cox, 1980) that five points would provide additional resolution by allowing a more finely graded response than a 4-point scale.

For the numerical scale, a 0-10-point scale was considered to be easily understood by people of all countries and cultures who are familiar with currency in a 10-base monetary system and other familiar counting situations. Later development has shown that the concept “on a scale zero to ten....” has become a preferred way of asking people’s opinion in polls and surveys.

Although specified in ISO/TS 15666:2003, many surveys use only one of the questions, due to space, cost and time constraints of conducting the survey; a historic preference for one of the questions over the other; or the fear that both may be perceived as repetitive. For these reasons, ISO/TS 15666:2021 relaxed the requirement to ask both questions, acknowledging that it may not always be possible or practical to include both questions. The updated specification now includes discussion of the advantages and disadvantages of each question to help users balance the strengths and weaknesses of each question against their research objectives.

There has long been debate about the use and format of Likert response scales in social research and there had been much consideration, research, and review of various formats to inform the development of the 11-point numeric and 5-point verbal scale questions (Schultz, 1978; Fields *et al.*, 2001). Concern about the use of Likert scales continues based on the lack of reporting regarding normality assumptions that should be satisfied for parametric statistical analysis. Likert scales with a greater number of points are more likely to satisfy normality assumptions and have additional precision than Likert scales with only a few number of points (Leung, 2011), but they could also be introducing a higher resolution in response that is not meaningful.

The 5-point verbal scale question is considered by the WG to be the simpler question. Selecting a one-word verbal response is most likely to be easily performed by all respondents and no further interpretation of the respondent’s choice is necessary. The question was developed to ensure that the commonly understood meaning of the verbal modifiers is consistent with its position on the scale as the verbal modifiers were rated in terms of their intensity of reaction to noise on a scale ranging from 0 to 100 so that different language version of the 5-point verbal scale would have the same positions on an underlying scale of intensity of reaction to noise (Fields *et al.*, 2001). It is generally assumed that the category names “not at all” through “extremely” are not necessarily equidistant on a Thurston scale (Torgerson, 1958), but if undertaken properly, following the procedure developed by the ICBEN team, one will get a scale with verbal categories that can be considered equidistant from each other, making the verbal scale interval. The procedure is based on the choice of modifiers successively bisecting a continuous annoyance scale. The disadvantage is that within cultures the meaning of language can change over time, which could influence the respondents’ perception of the verbal scale (Guski, 2017) and culturally bound terminology could add uncertainty to cross-study comparisons (Fields *et al.*, 2001). For many researchers, including Schultz (1978), scales such as the 5-point verbal question are attractive, as there is greater certainty that those responding ‘very’ or ‘extremely’ annoyed to the question can be assumed to be ‘highly annoyed’.

The simplicity of the 11-point numeric scale question avoids some of the issues related to the verbal question. The spacing along the 11-point scale is equal quantitatively and assumed to be so cognitively as well, which provides a pragmatic rationale for treating the data as “interval” in statistical testing, even though the scale is ordinal by definition. Further advantages are that it may offer a more sensitive assessment of annoyance, as the respondent has a greater number of possible response options but this could contribute to fluctuations in annoyance ratings, without meaningful changes in noise, which can influence definitions of %HA, for example, defining %HA as the top 28% means 7 and 8 on the 11-point scale question are treated very differently (Miedema and Vos, 1998).

Discussions undertaken to inform the revision have highlighted that researchers hold a clear preference for one of the noise reaction questions over the other, yet there is a clear drive within the field to move towards using either the 5-point verbal scale question or the 11-point numeric scale question. However, given the strong preferences of researchers and equal usage of each question in situations where one question has been used, it seems unlikely that agreement will be reached as to which question should be put forward. If having to choose between the verbal and/or the numeric question, ISO/TS 15666:2021 makes the informative recommendation to use the 11-point numeric scale question, as it offers the greatest options for statistical testing and is less likely to be culturally bound. As this is an informative recommendation researchers can continue to use the 5-point verbal scale question if they prefer as the normative text relaxes the requirement to use both questions.

1. Presentation of the verbal modifiers

It was noted during the revision that there were inconsistencies between the original ICBEN recommendations (Fields *et al.*, 1997; Fields *et al.*, 2001) and ISO/TS 15666:2003 regarding the order of presentation of the verbal modifiers. The ICBEN recommendation used a high/low presentation *(extremely, very, moderately, slightly, not at all*) whereas ISO/TS 15666:2003 recommended a low/high presentation (*not at all, slightly, moderately, very, extremely*). This had led to some inconsistencies in the ordering between some languages, with versions developed prior to the TS being more likely to have followed the high/low presentation. It is not clear why the order was changed in the 2003 TS but it has led, in effect, to two “compliant” versions of the 5-point verbal scale question being in use for the past two decades. We have found no plausible explanation for this reversal other than a low/high order of presentation may be a more natural way of presenting these modifiers, particularly for some languages and cultures. There are concerns that starting with ‘*extremely*’ could bias responses towards ‘higher annoyance’ but the same argument can be made for starting with ‘*not at all*’ which could bias towards ‘lower annoyance’. A study which compared annoyance responses for both the 5-point verbal scale and the 11-point numeric scale questions found no significant differences in annoyance responses when using either an ascending or descending presentation of the response options (Brink *et al.*, 2016). We have found no other studies that compared responses using the two approaches, nor could we robustly argue for one version over another. The revised ISO/TS 15666:2021 clarifies that the order should be ‘*Not at all* to *Extremely*’ to maintain consistency with ISO/TS 15666:2003 and to standardize the approach going forward. This approach is also consistent with the presentation of the 11-point numeric scale question which has always been presented from ‘*Not at all* to *Extremely*’ (0 to 10). Those undertaking surveys should check that their questions and their translations follow the low to high specification of the verbal modifiers going forward.

In terms of presentation of the 5-point verbal scale question, the ICBEN team and 2003 and 2021 TS set out that the verbal modifiers should be presented vertically on the page, rather than horizontally, as it is not possible to evenly-space the response options if presented horizontally. Our review suggests that this requirement is not always followed and we remind readers of this requirement to aid further standardization. Incidentally, this approach also facilitates the use of the question in online surveys, where respondents are often completing questionnaires on smartphones that have a portrait screen orientation.

1. Scope of the noise reaction questions
2. Time of day

ISO/TS15666:2003 asks respondents to rate their annoyance for the past 12-months, with an unwritten assumption that this covers the entire 24-hour period within the past 12-months. In recent years, several studies have compared annoyance responses for different times of the day (Wirth *et al.*, 2004; Öhrström *et al.*, 2006; Peris *et al.*, 2012; Michaud *et al.*, 2016a; Civil Aviation Authority, 2017), with noise at night-time consistently found to be more annoying than noise in general (assumed to cover 24-hour period) or in the daytime. This suggests that surveys should specifically include additional questions to assess night-time annoyance in their surveys, where relevant, as an unspecified time period could underestimate annoyance where there is considerable night-time exposure or where daytime and night-time exposure differ markedly. However, not all previous studies examining time of day follow the wording of the TS questions: for example, some studies specifically ask respondents “what time of day is noise most annoying”. A more appropriate approach would be to adopt the TS as specified (i.e. unspecified time period), substituting various time periods to compare to the data for the unspecified time period.

The revised TS maintains flexibility for researchers to assess time-periods that are relevant but proposes that all surveys should ask the TS questions first, which is assumed to be an aggregated response over 12 months covering the 24-hour time period of the day, as this provides a benchmark for comparing annoyance at other times of the day. When subsequently asking questions about annoyance at different times of the day, ISO/TS 15666:2021 recommends that the hours of the day of interest are specified (e.g., morning 07.00-12.00) rather than simply verbally described (e.g., morning) which could be interpreted differently by survey respondents and which may not necessarily be equivalent across different contexts.

1. Time of year

Methodologically there is also a need to build knowledge about how time of the year for data collection and the length of time over which annoyance is being assessed could be influencing annoyance responses. Studies have found that annoyance responses are higher in the summer months than winter months (Michaud *et al.*, 2005; Miedema *et al.*, 2005). A Swiss study found that annoyance responses were higher in the autumn than in the spring (Brink *et al.*, 2016). These findings could be related to window-opening behavior in the summer months, community expectations concerning and/or tolerance towards noise, but could also be explained by atmospheric conditions in the summertime influencing noise propagation and exposure at the dwelling. The Survey of Noise Attitudes (SoNA) 2014 study compared noise metrics calculated for different time-periods and compared it to aircraft noise annoyance assessed using the standard questions which covered the past 12 months, finding that the 92-day summer-average correlated better with mean annoyance scores than the 30-day average or 7-day average (Civil Aviation Authority, 2017). Whether such a difference would be found for other noise sources or contexts is unknown. As in previous publications, ISO/TS 15666:2021 specifies that surveys should report the period of data collection so that the influence of season and different time-periods on annoyance can be further evaluated.

1. Studies that compare noise annoyance for indoor and outdoor environments

As one of the changes from the previous edition, ISO/TS 15666:2021 clearly states that the terminology "at home" does not mean strictly indoors, but covers inside the home or outdoors at home, including balconies, gardens, etc. The recommended questions seek to obtain general, consistent reactions that allow respondents to integrate their experiences over different times and locations in and around their home.

According to Öhrström *et al.* (2006), however, there is a difference in response when asked by “at home”, by “indoor with windows closed”, by “indoor with open windows” and by “outdoor”. Annoyance “at home” was lower than annoyance “indoors with open windows” but higher than “indoors with windows closed”. This indicates that when people are asked about annoyance “at home” they probably give an answer that reflects various contextual factors. “Annoyance indoors with open windows” was, somewhat unexpectedly, higher than annoyance “outdoors close to the dwelling”. However, studies of wind turbine noise have found that outdoor annoyance is higher than indoor annoyance (Janssen *et al.*, 2011; Michaud *et al.*, 2016a).

This is an area requiring further investigation as, given the restorative value people place on access to good quality outdoor spaces (Hartig *et al.*, 2014; Marselle *et al.*, 2021), the reasons for the difference in annoyance are unclear without the benefit of contextual data that could further characterize noise exposure and noise reaction for indoor and outdoor environments.

For these reasons, it may be desirable to define the situations precisely if researchers want to know annoyance in a particular situation. Taking this into account, ISO/TS 15666:2021 recommends that if there is a need to evaluate noise annoyance responses in a certain context (e.g., indoors at home or for a particular season or a particular time of the day) the noise reaction questions can be repeated later in the survey, for the combination of conditions. It is a matter for further study whether the wording of the questionnaire for various contexts should form part of future revisions of ISO/TS 15666:2021,

1. Audibility of the noise source

The noise reaction questions were designed so that filter questions based on the audibility of the sound source (i.e., whether the respondent reports hearing the sound source) should not be used to eliminate respondents from answering the noise reaction questions (Fields *et al.*, 2001; ISO/TS15666:2003, 2003). Instead, if audibility is something important for the survey it should be determined later in the survey. This separation of annoyance from audibility makes it clear that respondents who are not annoyed can still report hearing a noise, i.e., avoids respondents conflating low levels of annoyance with not hearing the sound of the evaluate sources (Fields *et al.*, 2001). This would inflate the prevalence of high annoyance because excluding these individuals on the grounds that they claim to not hear a source (when they may mean they are not at all annoyed) reduces the denominator in logistic regression.

Our review of use of the 2003 TS found that some studies have examined audibility alongside annoyance rather than asking about audibility later in the survey. Several different approaches have been taken with some studies adding an additional option at the beginning of the 5-point verbal scale question with the verbal label “do not hear”(Defra, 2014) or amending the original verbal modifiers to account for audibility – e.g. “do not notice” versus “notice but not annoyed” (Pedersen and Waye, 2007). There is an intuitive appeal to this because it provides additional clarification on why a respondent may be selecting “not at all annoyed” however, these types of questions violate the equidistant requirements underpinning the 5-point verbal scale questions. No studies adding a question on audibility to the 11-point numerical scale question were identified nor had studies compared responses on the 5-point scale for the original and amended versions. Often, these surveys are applied in a specific location where the sample has been selected because they are exposed to a certain noise source. Sometimes national (or wide-scale) surveys are conducted, where exposure to the noise source being asked about is not guaranteed (Defra, 2014). In these cases, being able to distinguish the %HA as a proportion of different samples is an important tool for policy makers, i.e. being able to consider the proportion who “hear and do not feel they are annoyed” as well as those “who do not hear, so cannot be annoyed”. The WG decided given the lack of empirical evidence comparing these additions to the noise reaction questions that they could not include further changes regarding audibility at this time.

1. Cultural comparisons and translations
2. Construction of the verbal scale for translations

Special care must be taken when selecting the modifiers for the verbal scale. These questions were constructed by several rounds of translation and back-translation to maintain the exact meaning. The construction of the scale is described in detail in Fields *et al.* (2001) and has been already discussed. In brief, various verbal modifiers were rated in terms of their intensity of reaction to noise on a scale ranging from 0 to 100 so that different language version of the 5-point verbal scale would have the same positions on an underlying scale of intensity of reaction to noise. If done properly this will give a scale with verbal categories that can be considered to be equidistant from each other. Fields *et al.* (2001) list words for the 5-point verbal scale question in nine different languages together with a translation of the two standardized questions.

Since the publication of the TS in 2003, work has been done to construct verbal scales and standardized questions in eight new languages (Preis *et al.*, 2003; Yano and Ma, 2004; Kvist and Holm Pedersen, 2006; Günther *et al.*, 2007; Gjestland *et al.*, 2017; Lertsawat, 2020). ISO 15666:2021 now comprises this information for a total of 17 languages (Chinese (Pinin), Danish, Dutch (including Flemish), English, French, German, Hungarian, Japanese, Korean, Norwegian, Romanian, Polish, Portuguese-Brazilian, Spanish, Thai, Turkish, and Vietnamese).

1. Cultural comparison

It is challenging to untangle cultural differences attributable to nuances of the question scales from differences in responses across contexts, countries, and cultures because the social and cultural differences include many factors, including different social system and economic activities; different social and modal infrastructure; differences in the sound environment of local communities; and different attitudinal factors to environmental problems and noise sources.

Although it is difficult to determine what is meant by "cultural differences," there have been several studies comparing annoyance responses in different countries and regions. Morinaga *et al.* (2017) reported using *L*ct defined in ISO 1996-1 that the response was more severe in Japan than in Europe and the United States. In addition, it has been reported that in Vietnam, the response to road traffic noise was lower than in Europe and the U.S., while annoyance to aircraft noise was about the same or slightly higher (Nguyen *et al.*, 2016a; Nguyen *et al.*, 2016b). The authors proposed that the reasons for this might be that the benefits from road traffic were perceived to be greater than the benefits from aircraft. However, there is wide variation in annoyance responses generally across studies, making it unclear if differences between countries and regions represent real cultural differences or usual variation in response.

It is beyond the remit of the TS to further address the differences in response due to regional or cultural differences beyond issues relating to the most appropriate terminology for the scale questions.

1. SCORING ANNOYANCE
2. “Highly annoyed”

Starting with the Schultz curves (Schultz, 1978) it has become the norm to plot the percentage of the population ‘highly annoyed’ by noise exposure to obtain an annoyance exposure-response function. Schultz observed that high magnitudes of annoyance were more strongly associated with noise exposure than lower magnitudes of annoyance, which were more susceptible to non-acoustic factors. This meant that HA would have a better chance of being addressed by noise policy. However, use, and quantification, of ‘highly annoyed’ is not standardized. ISO/TS 15666:2003 did not prescribe how the 5-point verbal scale question and the 11-point numeric scale question should be scored, allowing for flexibility in how countries may wish to score annoyance. Nevertheless, over-time it has become apparent that the lack of standardized scoring can impede cross-study and cross-country comparisons.

Schultz originally defined ‘highly annoyed’ as the top 27-29% of an 11-point numeric scale question (Schultz, 1978) – the scores of 8, 9 or 10 on the scale. As specified in the extensive ICBEN analysis (Fields *et al.*, 2001) and consistent with Schultz’s analysis, ‘highly annoyed’ can be defined on the 5-point verbal scale question as the top two verbal response categories (i.e., *Very* and *Extremely*). At face value, it would appear that this gives the two questions different cut-offs to define ‘highly annoyed’, as the cut off is based on the top 28% (mid-point between 27-29%) for the numeric scale question and on the top 40% (the top two responses on the 5-point Likert scale – so two out of the five responses) for the verbal question. Given the different scoring methods for the verbal scale question and the numeric scale question, differences in annoyance scores between the two questions are therefore expected (Nguyen *et al.*, 2013; Brink *et al.*, 2016; Civil Aviation Authority, 2017; Nguyen *et al.*, 2017; Morinaga *et al.*, 2021). Miedema and Vos (1998) subsequently proposed converting both scale questions on a common scale to run from 0 to 100, which used a cut-off of the top 28%, which they stated was consistent with Schultz (1978). These different methodologies can lead to confusion as to how %HA has been scored across studies. Further, our review of use of ISO/TS 15666:2003 has identified that several studies provided numeric values to the verbal categories, e.g., 1-5 or 1-50, which is not consistent with the TS.

A common adaptation in the past decade based on comparative analyses in large-scale studies (Gjestland *et al.*, 2015; Brink *et al.*, 2016; Wothge *et al.*, 2017) have suggested that comparisons between the numeric and verbal scale questions can be improved by a weighting scheme, whereby: “highly annoyed” is defined as the top two verbal response categories on the 5-point verbal scale question (i.e., *Very* and *Extremely*) with ‘*Extremely’* counted in full, and ‘*Very’* weighted by a factor 0.4, which creates a greater mathematical similarity between the 5-point verbal and 11-point numeric questions. A comparison of these scoring methods, suggested that the 5-point verbal scale question resulted in lower %HA at all noise exposure levels in comparison to the 11-point numeric scale question (Brink *et al.*, 2016). The authors concluded that even weighting the 5-point verbal scale question to align cut off points on the two questions “does not automatically produce congruent exposure-annoyance relationships”. In comparison, the same study suggested that there was less of a difference between the average annoyance scores using the 11-point numeric scale and the 5-point verbal scale questions (Brink *et al.*, 2016). There are some concerns that making the 5-point verbal scale question cut-off similar to the 11-point numeric scale question omits a large proportion of those who are annoyed from the analysis. Michaud *et al.* (2005) suggested that a cut-off based on the scores of 7, 8, 9 and 10 (the top four responses representing a cut-off of the top 36% for %HA), also used in other studies (Lercher *et al.*, 2008), were more strongly correlated with the top two verbal response categories of ‘*Very*’ and ‘*Extremely*’ and that the conventional cut-off for %HA for the 11-point numeric scale question could be underestimating %HA.

If presenting both questions, evidence suggests that annoyance responses are not modified by the order of presentation of the question but that placement of the questions near the beginning of the survey produces higher annoyance responses than placement later in the survey (Brink *et al.*, 2016).

ISO/TS 15666: 2021 highlights that when reporting percentage ‘highly annoyed’ it is important to clearly state which scoring has been used, to support comparison and harmonisation of data. To aid standardization of reporting, a naming system for ‘highly annoyed’ has been set out in the ISO/TS 15666:2021 revision as described in Table I.

TABLE I. Summary of scoring and naming conventions for ‘highly annoyed’ in ISO/TS 15666:2021.

|  |  |  |
| --- | --- | --- |
| Question format | Scoring ‘highly annoyed’ | Naming of the outcome |
| 11-point numeric question | Values 8, 9, and 10 | HAN |
| 5-point verbal question | Values 4 and 5/*’very’* and ‘*extremely’* | HAV |
| 5-point verbal question | Values 4 and 5/*’very’* and ‘*extremely’* with ‘*extremely’* counted in full, and ‘*very’* weighted by a factor of 0.4 | HAVW |

It is however recognised that users and countries may conventionally adopt other forms of scoring due to preference or historical need, which may not be based on a need to report the prevalence of high annoyance. This is allowed for within the specification, but in these instances, it is encouraged that the data is also presented in terms of HAN, HAV and HAVW so that the data could later be included in meta-analyses and/or compared to other study findings. Where the numerical scale is used, presenting a cumulative distribution of the data would also aid in cross-study comparisons and meta-analysis.

A recent paper published after the revised TS was published, has provided additional discussion about issues to consider when scoring HA including rules for converting annoyance response data that originate from different response scales on a unified scale from 0 to 100 in order to be able to pool annoyance scores and compare resulting linear exposure-annoyance relationships (Brink *et al.*, In press 2021). The paper highlights how the annoyance exposure-response relationship is shifted on the x-axis by the use of different decision making in relation to scoring and that this also influences the steepness of the regression lines (Brink *et al.*, In press 2021).

1. Type of Data and Statistical Analysis

Whilst developed as an ordinal scale (Fields *et al.*, 1997; Fields *et al.*, 2001), the 5-point verbal scale is often being treated as a numerical scale. ISO/TS15666:2021 has clarified the original requirement that if the verbal question is converted to a numeric scale for analysis purposes, there needs to be clear evidence that the step-difference between each category is perceived by the respondent to be equal (Fields *et al.*, 1997; Fields *et al.*, 2001).

1. STANDARDIZATION

It is clear that many researchers, with the best of intentions, are using the noise reaction questions to assess annoyance but are not following some of the specific details of the specification. We encourage researchers planning to undertake an annoyance survey to take time to understand the philosophy and detail that lies behind the specification (Fields *et al.*, 1997; Fields *et al.*, 2001), which set out how the survey was designed and should be conducted, as well as following the 2021 revised TS including the Annexes to ensure that any further changes are taken into account.

The checklist about information that should be reported for an annoyance survey, has been updated in ISO/TS 15666:2021. The checklist should be consulted at the design stage of the survey and completed when the survey is published. Editors and peer-reviewers should make sure the checklist is reported. This detail not only ensures a robust survey has been carried out; but will be useful for meta-analyses pooling findings across studies; and will improve the overall standard of reporting.

Predicting annoyance at any given sound level has large uncertainty, with wide-ranging estimates of annoyance being found for the same sound level across studies (Guski *et al.*, 2017; Gjestland, 2020). The TS focuses on the y-axis of the exposure-response function - how annoyance is assessed. However, scrutiny should also be applied to the x-axis of the relationship – the estimation of noise exposure which also introduces uncertainty into exposure-response functions. Imprecision in estimates within and across studies can be introduced by poor quality input data for environmental models on factors that influence source emissions and dispersion in the environment such as land use, building types and locations, terrain, vehicle types, speed and traffic patterns, weather, sources, *inter alia*; methodological and computational differences between models; the lack of validated models; as well as issues relating to the measurement of noise *in situ*. Uncertainty relating to noise exposure and its influence on exposure-response functions is often overlooked (Horonjeff, 2021) and is likely to also have a significant impact.

1. RESEARCH ROADMAP

Further studies examining several methodological aspects of the standardization are needed to feed into the next revision of the specification. Studies should seek to explore methodological variations in the noise reaction questions, after implementation of the TS questions.

Further studies are needed to inform discussions about the assumptions that respondents are rating their annoyance over the past 12 months; over a 24-hour period; and for the home environment as a whole. It is not known how respondents integrate or recall their experience over the past 12 months. It seems plausible that people may be recalling the past month or so and basing their 12-month evaluation on their more recent experience. This could result in weaker associations between annoyance and long-term noise metrics (e.g., Lden). It is not known how participants integrate assessments for different areas of the home or how they might combine to reflect the assessment in the overall noise reaction questions. Are respondents reporting annoyance for the ‘worst’ area in their home or are they integrating the reaction across different areas? Further studies comparing data on annoyance for different times of the day to the TS noise reaction questions would be helpful. There is also support within the field for extending use of the noise reaction questions beyond residential settings where these issues would also be relevant and a need to explore the issue of equidistance in responses for both questions in more detail. Digital technology could be used to collect noise reaction responses more frequently for participants to shed light on these issues. We also need to understand how electronic methods of presenting the TS questions compare to the traditional face-to-face or telephone data collection methods. In particular, there are concerns that the 11-point numeric scale question may not translate well to presentation on mobile devices.

Recent surveys have highlighted how the exposure-response curve for annoyance is shifted considerably in relation to non-acoustic factors such as attitudes, demographic factors, and environmental and contextual factors (Guski, 1999; Civil Aviation Authority, 2017; Guski *et al.*, 2017) and can have a larger impact on annoyance than the noise exposure, per se (Job, 1988; Job, 1996). Non-acoustic factors have been hypothesized to account for between-study variation in exposure-response curves. There is an appetite to develop a standardized assessment survey addressing non-acoustic factors to aid further harmonization across surveys and this is currently being undertaken as part of a separate work item, led by the UK, within ISO TC43/SC1/WG62 by a different ISO Working Group (Fenech *et al.*, 2021). The rationale for the proposed new standard is that there are currently four international standards related directly to the human perception/evaluation of sound: ISO TS 15666:2021 (assessment of noise annoyance) and ISO 12913 parts 1, 2 and 3 (measurement and assessment of soundscape quality) (ISO 12913-1:2014, 2014; ISO/TS 12913-2:2018, 2018; ISO/TS 12913-3:2019, 2019). These standards aim to harmonize the characterization of perceptions/reactions to a specific sound/sound environment. Personal, social and situational variables (often referred to as non-acoustic factors) are as important as acoustic features in determining human evaluation of sound. Currently socio-acoustic surveys on annoyance and soundscapes attempt to quantify the influence of these non-acoustic factors using study-specific questions. This limits opportunities to merge different survey datasets in order to a) evaluate the effectiveness of specific questions as measuring instruments b) improve the interpretation of survey data and c) identify effective non-acoustic interventions.

Research is increasingly questioning whether the two noise reaction questions are adequate for describing the multi-faceted complex concept that is annoyance (Schreckenberg *et al.*, 2018). The assessment of annoyance, by two single items that are not combined in any way, differs to that for most other health and psychological constructs, which are assessed using multi-item scales that demonstrate robust psychometric properties when the items are considered together. Using data from the German NORAH study (Noise-Related Annoyance, Cognition and Health) the authors developed a Multiple Item Annoyance Scale (MIAS). Drawing on definitions of annoyance (Guski, 1999; Guski *et al.*, 2017), three components of annoyance were examined including aircraft noise-related disturbances; affective evaluations and attitudes; and perceptions of control and coping capacity. The study developed two additional scales for use alongside the ISO/TS 15666 5-point verbal scale question which examined ‘noise source-specific disturbances’ and ‘perceived coping capacity’ (Schreckenberg *et al.*, 2018). The authors proposed that the MIAS could have advantages over the single noise reaction question, helping to further explain ‘non-acoustic factors’ and to assess the impact of mitigation or operational change. It is also likely that a multidimensional approach may shed further light on the role of annoyance as a moderator or mediator on environmental noise effects on health (Peris and Fenech, 2020).

Another approach has been put forward by Michaud *et al.* (2018b) who argued that for some noise sources, like wind turbines, noise is one of several features capable of causing annoyance. By asking about annoyance toward several features it is possible to numerically combine them and develop an overall score which is more meaningful than reporting on a single feature in isolation because that gives the impression that mitigation for one feature alone will reduce overall annoyance. Michaud *et al.* (2018a) found that a mean aggregate annoyance score could reliably distinguish participants who self-reported health effects (or noise complaints) from those who did not, and that with further research, this approach could be one of several factors considered for determining noise limits, rather than basing limits on %HA.

Within the field, in the next few years we are likely to see an expansion of the assessment of annoyance to acknowledge the complexities of community annoyance responses. However, to achieve this, we need further evidence from psychometrically robust studies from different contexts (countries and noise sources) (Boateng *et al.*, 2018) which explore a multidimensional approach of annoyance for different sources.

1. CONCLUSION

The methodologically robust work of the ICBEN team, several decades ago, standardizing the methodology has placed annoyance at the forefront of discussions about the public health effects of environmental noise. However, we have to appreciate the considerable expectations that rest upon the noise reaction questions: they are expected to cover all people at all times in all locations within their residential environments, for all environmental noise sources, in many languages, providing a 5-point verbal scale question and 11-point numeric scale question to cover all disciplinary and analysis preferences, that should be comparable for all cultures and contexts. These questions are expected to work extremely hard and the time has come to reflect and act to further develop the field in the next few years. To quote Schultz (1978), over forty years later, “There is so much work yet to be done in understanding how people respond to noise that one might say the task has barely begun”.

ACKNOWLEGEMENTS

The authors wish to acknowledge and thank all the participants of ISO TC43/SC1/WG62 committee for their considerable input and work over several years. The authors also wish to acknowledge current and past members of the ICBEN Community Response to Noise team for responding to requests for evidence and discussions about key aspects of the TS; and Philip Dunbavin convenor of the ISO working group WG62, which is responsible for the revision of ISO/TS 15666:2021

REFERENCES

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., and Young, S. L. (**2018**). "Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer," Front Public Health **6**, 149.

Brink, M., Giorgis-Allemand, L., Schreckenberg, D., and Evrard, A. S. (**In press 2021**). "Pooling and comparing noise annoyance scores and "high annoyance" (HA) responses on the 5-point and 11-point scales: Principles and practical advice," Int. J. Environ. Res. Public Health.

Brink, M., Schaffer, B., Vienneau, D., Foraster, M., Pieren, R., Eze, I. C., Cajochen, C., Probst-Hensch, N., Roosli, M., and Wunderli, J. M. (**2019**). "A survey on exposure-response relationships for road, rail, and aircraft noise annoyance: Differences between continuous and intermittent noise," Environ. Int. **125**, 277-290.

Brink, M., Schreckenberg, D., Vienneau, D., Cajochen, C., Wunderli, J.-M., Probst-Hensch, N., and Roosli, M. (**2016**). "Effects of scale, question location, order of response alternatives, and season on self-reported noise annoyance using ICBEN Scales: A field experiment," Int. J. Environ. Res. Public Health **13**, 1163.

Brink, M., Wirth, K. E., Schierz, C., Thomann, G., and Bauer, G. (**2008**). "Annoyance responses to stable and changing aircraft noise exposure," J. Acoust. Soc. Am. **124**, 2930-2941.

Civil Aviation Authority (**2017**). "CAP1506 Survey of Noise Attitudes 2014," (Civil Aviation Authority, London; United Kingdom).

Cox, E. P. (**1980**). "The optimal number of response alternatives for a scale: a review," Journal of Marketing Research **17**, 407-422.

Defra (**2014**). "NANR322 Survey of Noise Atttitudes (SoNA 2013) - NO0242," (Department of Environment, Food and Rural Affairs United Kingdom ).

Department for Transport (**2013**). "Transport analysis guidance: WebTAG," (Department of Transport).

European Commission (**2002**). "Position paper on dose response relationships between transportation noise and annoyance," (Office for Official Publicaitons of the European Commission, Luxembourg.).

European Environment Agency (**2020**). "Environmental Noise in Europe 2020," (Publications of the European Union, Luxembourg).

Fenech, B., Lavia, L., Rodgers, G., and Notley, H. (**2021**). "Development of a new ISO Technical Specification on non-acoustic factors to improve the interpretation of annoyance and soundscape datasets," in *13th Congress on Noise as a Public Health Problem (ICBEN)* (Stockholm, Sweden).

Fidell, S. (**2003**). "The Schultz curve 25 years later: A research perspective," J. Acoustic. Soc. Am **114**, 3007-3015.

Fields, J. M., Jong, R. G., Brown, A. L., Flindell, I. H., Gjestland, T., Job, R. F. S., Kurra, S., Lercher, P., Schuemer-Kohrs, A., Vallet, M., and Yano, T. (**1997**). "Guidelines for reporting core information from community noise reaction surveys," J. Sound Vib **206**, 685-695.

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

Gjestland, T. (**2020**). "On the Temporal Stability of People’s Annoyance with Road Traffic Noise," Int. J. Environ. Res. Public Health **17**, 1374.

Gjestland, T., Gelderblom, F. B., and Fidell, S. (**2019**). "Sample size implications for calculations of community tolerance level values from social surveys of noise-induced annoyance," J. Acoust. Soc. Am. **146**, 1212.

Gjestland, T., Nguyen, T. L., and Yano, T. (**2015**). "Community response to noise in Vietnam: exposure-response relationships based on the community tolerance level," J. Acoust. Soc. Am. **137**, 2596-2601.

Gjestland, T., Toma, A., Dragasanu, L., Deaconu, M., and Oprea, B. (**2017**). "Romanian Standardized Noise Reaction Questions for Community Noise Surveys," Acta Acustica united with Acustica **103**, 232-235.

Günther, H., Iglesias, F., and de Sousa, J. M. (**2007**). "Note on the development of a Brazilian version of a noise annoyance scale," J. Sound Vib **308**, 343-347.

Guski, R. (**1999**). "Personal and social variables as co-determinants of noise annoyance," Noise Health **1**, 45-56.

Guski, R. (**2017**). "The increase of aircraft noise annoyance in communities. Causes and consequences " in *12th International Commission on the Biological Effects of Noise (ICBEN)* (Zurich, Switzerland).

Guski, R., Schreckenberg, D., and Schuemer, R. (**2017**). "WHO Environmental Noise Guidelines for the European Region: A systematic review on environmental noise and annoyance," Int. J. Environ. Res. Public Health **14**, 1539.

Hartig, T., Mitchell, R., de Vries, S., and Frumkin, H. (**2014**). "Nature and health," Annu. Rev. Public Health **35**, 207-228.

Health Canada (**2017**). "Guidance for Evaluating Human Health Impacts in Environmental Assessment: Noise. ."

Horonjeff, R. D. (**2021**). "An examination of dose uncertainty and dose distribution effects on community noise attitudinal survey outcomes," J. Acoustic. Soc. Am **150**, 1691-1701.

ISO 1996-1:2016 (**2016**). "Acoustics - Description, measurement and assessment of environmental noise - Part 1: Basic quantities and assessment procedures," edited by I. O. f. Standardization (International Organization for Standardization, Geneva; Switzerland).

ISO 12913-1:2014 (**2014**). "Acoustics — Soundscape — Part 1: Definition and conceptual framework " (International Organization for Standardization, Geneva; Switzerland).

ISO (**2021**). "The different types of ISO publications," edited by International Standards Organisation.

ISO/TS15666:2003 (**2003**). "Acoustics - Assessment of noise annoyance by means of social and socio-acoustic surveys," (International Organization for Standardization, Geneva; Switzerland).

ISO/TS15666:2021 (**2021** ). "Acoustics - Assessment of noise annoyance by means of social and socio-acoustic surveys," (International Organization for Standardization, Geneva; Switzerland).

ISO/TS 12913-2:2018 (**2018**). "Acoustics — Soundscape — Part 2: Data collection and reporting requirements," (International Organization for Standardization, Geneva; Switzerland).

ISO/TS 12913-3:2019 (**2019**). "Acoustics — Soundscape — Part 3: Data Analysis," (International Organization for Standardization, Geneva; Switzerland).

Janssen, S. A., Vos, H., Eisses, A. R., and Pedersen, E. (**2011**). "A comparison between exposure-response relationships for wind turbine annoyance and annoyance due to other noise sources," J. Acoust. Soc. Am. **130**, 3746-3753.

Job, R. F. S. (**1988**). "Over-reaction to changes in noise exposure: the possible effect of attitude " J. Sound Vib **126**, 550-552.

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

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

Kryter, K. D. (**1982b**). "Rebuttal by Karl D. Kryter to comments by T. J. Schultz," J. Acoustic. Soc. Am **72**, 1253-1257.

Kryter, K. D. (**1983**). "Response of K. D. Kryter to modified comments by T. J. Schultz on K. D. Kryter’s paper, ‘‘Community annoyance from aircraft and ground vehicle noise’’[J. Acoust. Soc. Am. 72, 1243–1252 (1982)]," J. Acoustic. Soc. Am **73**, 1066-1068.

Kuwano, S., Yano, T., Kageyama, T., Sueoka, S., and Tachibana, H. (**2014**). "Social survey on wind turbine noise in Japan," Noise Control Eng. J. **62**, 503-520.

Kvist, P., and Holm Pedersen, T. (**2006**). "Translation into Danish of the questions and modifiers for socio-acoustic surveys," in *Euronoise* (Tampere, Finland).

Lercher, P., de Greve, B., Botteldooren, D., and Rüdisser, J. (**2008**). "A comparison of regional noise-annoyance-curves in alpine areas with the European standard curves. ," in *9th International Conference on Noise as a Public Health Problem (ICBEN 2008)* (Foxwoods, CT, USA).

Lertsawat, K. (**2020**). "Verbal scale and standardized questions in Thai ".

Leung, S. O. (**2011**). "A Comparison of Psychometric Properties and Normality in 4-, 5-, 6-, and 11-Point Likert Scales," Journal of Social Service Research **37**, 412-421.

Marselle, M. R., Hartig, T., Cox, D. T. C., de Bell, S., Knapp, S., Lindley, S., Triguero-Mas, M., Böhning-Gaese, K., Braubach, M., Cook, P. A., de Vries, S., Heintz-Buschart, A., Hofmann, M., Irvine, K. N., Kabisch, N., Kolek, F., Kraemer, R., Markevych, I., Martens, D., Müller, R., Nieuwenhuijsen, M., Potts, J. M., Stadler, J., Walton, S., Warber, S. L., and Bonn, A. (**2021**). "Pathways linking biodiversity to human health: A conceptual framework," Environ. Int. **150**, 106420.

Michaud, D. S., Feder, K., Keith, S. E., Voicescu, S. A., and et al (**2016a**). "Exposure to wind turbine noise: perceptual responses and reported health effects," J. Acoust. Soc. Am. **139**.

Michaud, D. S., Keith, S. E., Feder, K., Voicescu, S. A., Marro, L., Than, J., Guay, M., Bower, T., and al, e. (**2016b**). "Personal and situational variables associated with wind turbine noise annoyance," J. Acoust. Soc. Am. **139**, 1455-1466.

Michaud, D. S., Keith, S. E., and McMurchy, D. (**2005**). "Noise annoyance in Canada," Noise Health **7**, 39-47.

Michaud, D. S., Marro, L., and McNamee, J. (**2018a**). "The association between self-reported and objective measures of health and aggregate annoyance scores toward wind turbine installations," Can. J. Public Health. **109**, 252-260.

Michaud, D. S., Marro, L., and McNamee, J. (**2018b**). "Derivation and application of a composite annoyance reaction construct based on multiple wind turbine features," Can. J. Public Health. **109**, 242-251.

Miedema, H. M., Fields, J. M., and Vos, H. (**2005**). "Effect of season and meteorological conditions on community noise annoyance," J. Acoust. Soc. Am. **117**, 2853-2865.

Miedema, H. M., and Oudshoorn, C. G. (**2001**). "Annoyance from transportation noise: relationships with exposure metrics DNL and DENL and their confidence intervals," Environ. Health Perspect. **109**, 409-416.

Miedema, H. M., and Vos, H. (**1998**). "Exposure-response relationships for transportation noise," J. Acoust. Soc. Am. **104**, 3432-3445.

Miller, N. P., Czech, J. J., Hellauer, K. M., Nicholas, B. L., Lohr, S., Jodts, E., Broene, P., Morganstein, D., Kali, J., Zhu, X., Cantor, D., Hudnall, J., and Melia, K. (**2021a**). "Analysis of the Neighborhood Environmental Survey," (U.S. Department of Transportation, Federal Aviation Administration).

Miller, N. P., Czech, J. J., Hellauer, K. M., Nicholas, B. L., Lohr, S., Jodts, E., Broene, P., Morganstein, D., Kali, J., Zhu, X., Cantor, D., Hudnall, J., and Melia, K. (**2021b**). "Analysis of the Neighborhood Environmental Survey: TC-21-4 Analysis of NES," edited by U. S. D. o. T. Federal Aviation Administration (Federal Aviation Administration, U.S. Department of Transportation. Atlantic City; New Jersey).

Morinaga, M., Nguyen, T. L., Shimoyama, K., Yokoshima, S., and Yano, T. (**2020**). "Effects of step change in aircraft noise exposure on activity disturbances: Socio-acoustic surveys around Hanoi Noi Bai International Airport," Acoust. Sci. Technol. **41**, 590-597.

Morinaga, M., Nguyen, T. L., Yokoshima, S., Shimoyama, K., Morihara, T., and Yano, T. (**2021**). "The Effect of an Alternative Definition of "Percent Highly Annoyed" on the Exposure-Response Relationship: Comparison of Noise Annoyance Responses Measured by ICBEN 5-Point Verbal and 11-Point Numerical Scales," Int. J. Environ. Res. Public Health **18**.

Morinaga, M., Yokoshima, S., Morihara, T., Kawai, K., and Yano, T. (**2017**). "Community tolerance level for transportation noises derived from the Socio-Acoustic Survey Data Archive, SASDA," in *The 12th Congress on Noise as a Public Health Problem (ICBEN 2017)* ( Zurich, Switzerland.).

Murakami, Y., Yano, T., Morinaga, M., and Yokoshima, S. (**2018**). "Effects of Railway Elevation, Operation of a New Station, and Earthquakes on Railway Noise Annoyance in Kumamoto, Japan," Int. J. Environ. Res. Public Health **15**, 1417.

Nguyen, T., Yano, T., and Morihara, T. (**2013**). "A method to compare the prevalence of annoyance measured with different scales," in *Internoise 2013* (Innsbruck, Austria).

Nguyen, T., Yano, T., Morihara, T., Yokoshima, S., and Morinaga, M. (**2017**). "Comparison of annoyance response measuerd with ICBEN 5-point verbal and 11-point numerical scales in Japanese and Vietnamese," in *12th International Conference on Noise as a Public Health Problem (ICBEN 2008)* (Zurich, Switzerland).

Nguyen, T., Yano, T., Nishimura, T., and Sato, T. (**2016a**). "Exposure-response relationships for road traffic and aircraft noise in Vietnam," Noise Control Eng. J. **64**, 243-258.

Nguyen, T. L., Nguyen, T. L., Yano, T., Nishimura, T., Sato, T., Morinaga, M., and Yamada, I. (**2016b**). "Social surveys on community response to a change in aircraft noise exposure before and after the operation of the new terminal building in Hanoi Noi Bai International Airport " in *The 45th International Congress on Noise Control Engineering Inter-noise 2016) Hamburg, Germany.* (Hamburg, Germany.).

Öhrström, E., Skånberg, A., Svensson, H., and Gidlöf-Gunnarsson, A. (**2006**). "Effects of road traffic noise and the benefit of access to quietness," J. Sound Vib **295**, 40-59.

Pedersen, E., and Waye, K. P. (**2007**). "Wind turbine noise, annoyance and self-reported health and well-being in different living environments," Occup. Environ. Med. **64**, 480-486.

Peris, E., and Fenech, B. (**2020**). "Associations and effect modification between transportation noise, self-reported response to noise and the wider determinants of health: a narrative synthesis of the literature," Sci. Total Environ. **748**.

Peris, E., Woodcock, J., Sica, G., Moorhouse, A. T., and Waddington, D. C. (**2012**). "Annoyance due to railway vibration at different times of the day," J. Acoust. Soc. Am. **131**, El191-196.

Preis, A., Kaczmarek, T., Wojciechowska, H., Zera, J., and Fields, J. M. (**2003**). "Polish version of standardized noise reaction questions for community noise surveys," Int. J. Occup. Med. Environ. Health **16**, 155-159.

Schreckenberg, D., Belke, C., and Spilski, J. (**2018**). "The Development of a Multiple-Item Annoyance Scale (MIAS) for Transportation Noise Annoyance," Int J Environ Res Public Health **15**.

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

Schultz, T. J. (**1982**). "Comments on K. D. Kryter’s paper, ’’Community annoyance from aircraft and ground vehicle noise’’," J. Acoustic. Soc. Am **72**, 1243-1252.

Torgerson, W. S. (**1958**). *Theory and Methods of Scaling. J Wiley & Sons, Inc*.

US Federal Interagency Committee on Noise (FICON) (**1992**). "Federal Agency review of selected airport noise analysis issues " (FICON, Washington DC, US.).

van Kamp, I., Schreckenberg, D., van Kempen, E., Basner, M., Brown, A. L., Clark, C., Houthuijs, D., Breugelmans, O., van Beek, A. J., and Janssen-Stelder, B. M. (**2018**). "Study on methodology to perform environmental noise and health assessment," edited by National Institute for Public Health and the Environment (RIVM) (National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands).

van Kamp, I., Simon, S., Notley, H., Baliatas, C., and van Kempen, E. (**2020**). "Evidence Relating to Environmental Noise Exposure and Annoyance, Sleep Disturbance, Cardio-Vascular and Metabolic Health Outcomes in the Context of IGCB (N): A Scoping Review of New Evidence," Int. J. Environ. Res. Public Health **17**, 3016.

Wirth, K., Brink, M., and Schierz, C. (**2004**). "Aircraft Noise Annoyance at Different Times of Day."

World Health Organization (**2011**). "Burden of Disease from Environmental Noise," (World Health Organization, Europe, Copenhagen; Denmark).

World Health Organization (**2018**). "The World Health Organization Guidelines for Environmental Noise Exposure for the European Region," (World Health Organization Europe, Copenhagen: Denmark).

Wothge, J., Belke, C., Möhler, U., Guski, R., and Schreckenberg, D. (**2017**). "The Combined Effects of Aircraft and Road Traffic Noise and Aircraft and Railway Noise on Noise Annoyance-An Analysis in the Context of the Joint Research Initiative NORAH," Int. J. Environ. Res. Public Health **14**, 871.

Yano, T., and Ma, H. (**2004**). "Standardized noise annoyance scales in Chinese, Korean and Vietnamese," J. Sound Vib **277**, 583-588.