Short-term annoyance reactions to stationary and time-varying wind turbine and road traffic noise: A laboratory study[a)]

Beat Schäffer[b)]
Empa, Swiss Federal Laboratories for Materials Science and Technology, Laboratory for Acoustics/Noise Control, Überlandstrasse 129, 8600 Dübendorf, Switzerland

Sabine J. Schlittmeier
Work, Environmental and Health Psychology, Catholic University of Eichstätt-Ingolstadt, 85072 Eichstätt, Germany

Reto Pieren and Kurt Heutschi
Empa, Swiss Federal Laboratories for Materials Science and Technology, Laboratory for Acoustics/Noise Control, Überlandstrasse 129, 8600 Dübendorf, Switzerland

Mark Brink
Federal Office for the Environment, Noise and NIR Division, 3003 Bern, Switzerland

Ralf Graf and Jürgen Hellbrück
Work, Environmental and Health Psychology, Catholic University of Eichstätt-Ingolstadt, 85072 Eichstätt, Germany

(Received 25 November 2015; revised 19 April 2016; accepted 26 April 2016; published online 24 May 2016)

Current literature suggests that wind turbine noise is more annoying than transportation noise. To date, however, it is not known which acoustic characteristics of wind turbines alone, i.e., without effect modifiers such as visibility, are associated with annoyance. The objective of this study was therefore to investigate and compare the short-term noise annoyance reactions to wind turbines and road traffic in controlled laboratory listening tests. A set of acoustic scenarios was created which, combined with the factorial design of the listening tests, allowed separating the individual associations of three acoustic characteristics with annoyance, namely, source type (wind turbine, road traffic), A-weighted sound pressure level, and amplitude modulation (without, periodic, random). Sixty participants rated their annoyance to the sounds. At the same A-weighted sound pressure level, wind turbine noise was found to be associated with higher annoyance than road traffic noise, particularly with amplitude modulation. The increased annoyance to amplitude modulation of wind turbines is not related to its periodicity, but seems to depend on the modulation frequency range. The study discloses a direct link of different acoustic characteristics to annoyance, yet the generalizability to long-term exposure in the field still needs to be verified.
 [http://dx.doi.org/10.1121/1.4949566]

[SF] Pages: 2949–2963

I. INTRODUCTION

The production of wind energy is becoming increasingly important worldwide, with wind capacity growing between 1997 and 2014 from 8 to 370 MW by a factor of almost 50 (GWEC, 2015). While the development of wind farms as renewable energy sources is environmentally beneficial, it also results in larger portions of the population being exposed to wind turbine noise (WTN). Wind farms are thus becoming an increasingly important source of industrial noise. WTN has been associated with various health effects, in particular, with annoyance and sleep disturbance (McCunney *et al.*, 2014; Schmidt and Klokker, 2014; Onakpoya *et al.*, 2015). There is evidence from literature that, at comparable sound pressure levels, WTN is associated with higher annoyance reactions than transportation or industrial noise (Janssen *et al.*, 2011). Knowledge of the reasons for these annoyance differences, however, is still relatively scarce. In particular, it is not known which acoustic characteristics of wind turbines alone, i.e., without potential effect modifiers such as the visibility of wind turbines, are associated with annoyance. The objective of this study was therefore to investigate and compare the annoyance reactions to WTN and road traffic noise (RTN) under controlled conditions in the laboratory. The focus was on noise annoyance reactions to short-time exposure (as opposed to annoyance to long-term exposure).

Current literature, as recently reviewed by McCunney *et al.* (2014) and Schmidt and Klokker (2014), suggests that the annoyance reactions to WTN may be explained by a range of factors, namely, by the visibility of wind turbines (Knopper and Ollson, 2011), shadow flicker (Voicescu *et al.*, 2016), the living environment of residents (Pedersen and Larsman, 2008), identifying wind turbines as the noise source leading to window closing (Michaud *et al.*, 2016b), and by individual attributes such as noise sensitivity (Miedema and Vos, 2003), attitude (Pedersen and Persson

[a)]Portions of this work were presented in "Annoyance potential of wind turbine noise compared to road traffic noise," in *Proceedings of Euronoise 2015, 10th European Congress and Exposition on Noise Control Engineering*, Maastricht, The Netherlands, June 2015.
[b)]Electronic mail: beat.schaeffer@empa.ch

Waye, 2004), or economic benefit (Pedersen *et al.*, 2009), in addition to acoustic characteristics. With respect to the latter, periodic amplitude modulation (AM), i.e., periodic temporal level variations sometimes observed for WTN, might be particularly important (van den Berg, 2009; Bockstael *et al.*, 2012; RenewableUK, 2013). However, it is not known which acoustic characteristics alone, i.e., without the consideration of effect modifiers, are associated with (noise) annoyance. This aspect is the focus of the present study.

Recent literature suggests that the acoustic characteristics of WTN are only weakly associated with long-term annoyance assessed in field surveys and that "non-acoustic" (individual, situational) variables play a crucial role (Knopper and Ollson, 2011; McCunney *et al.*, 2014; Michaud *et al.*, 2016b). While the present study exclusively addresses the annoyance reactions to acoustic characteristics of WTN, the important role of non-acoustic variables, albeit not examined here, is acknowledged.

The link of acoustic characteristics to short-term noise annoyance may be investigated in laboratory experiments, as they allow for controlled acoustic situations (e.g., with or without AM) and for exclusion or at least control of potential effect modifiers such as the visual appearance of wind turbines in field surveys. To date various laboratory studies provide evidence of the role of acoustic characteristics of wind turbines for short-term noise annoyance. Sound pressure level is a crucial factor (Lee *et al.*, 2011; Seong *et al.*, 2013). Besides, at a given level annoyance was found to be linked with the type of wind turbine (power, manufacturer) (Persson Waye and Öhrström, 2002; Legarth, 2007), and to increase with the magnitude of periodic AM (Lee *et al.*, 2011). So far, however, laboratory studies either only included WTN as the single sound source (Persson Waye and Öhrström, 2002; Legarth, 2007; Lee *et al.*, 2011; RenewableUK, 2013; Seong *et al.*, 2013) or, when comparing the annoyance to WTN with other noise sources, focused on a single sound pressure level (Van Renterghem *et al.*, 2013). Exposure-response curves for wind turbines in comparison to other sound sources, covering a wide range of sound pressure levels and established under the same controlled laboratory conditions, are currently unavailable.

The objective of the present study therefore was to investigate and compare the short-term annoyance reactions to WTN and RTN over a wide range of sound pressure levels under controlled laboratory conditions. More specifically, annoyance reactions to outdoor WTN and RTN situations during the day (e.g., leisure time) were studied. RTN served as the reference source, as it is the major noise source in the environment (BAFU, 2009). Different WTN and RTN situations covering a wide range of acoustic characteristics (sound pressure level, AM) were studied, which allowed separating the association of source type, sound pressure level, and AM with noise annoyance.

II. METHODS

In this study, the impact of different acoustic characteristics of WTN and RTN on short-term noise annoyance was studied under laboratory conditions. The annoyance ratings correspond to "short-term annoyance" (Bolin *et al.*, 2014) or "psychoacoustic annoyance" (Fastl and Zwicker, 2007), which is different to the long-term annoyance assessed in field surveys (Guski and Bosshardt, 1992). In the following, we refer to the annoyance studied here as "annoyance rating" (for the individual ratings) or "short-term (noise) annoyance."

A. Listening tests—concept

In the listening tests, sound stimuli were systematically varied with respect to the three variables source type, A-weighted sound pressure level, and AM to study their individual associations with the annoyance ratings (Table I).

The A-weighted equivalent continuous sound pressure levels (L_{Aeq}) of 35–60 dB of the stimuli (Table I) cover an environmentally relevant range for WTN and RTN (e.g., Miedema and Oudshoorn, 2001; Janssen *et al.*, 2011; McCunney *et al.*, 2014). WTN was not studied below a L_{Aeq} of 35 dB as annoyance becomes negligible (Schmidt and Klokker, 2014). For the same reason, RTN was not studied below a L_{Aeq} of 40 dB (Miedema and Oudshoorn, 2001; Lercher *et al.*, 2008). WTN was not presented at a L_{Aeq} of 60 dB, as this level occurs only very close to turbines.

"Without AM" corresponds to quasi-stationary (constant over time) RTN or WTN. WTN with time-varying "periodic AM" represents situations with high-frequency "swishing" as well as low-frequency "thumping" sound (Bowdler, 2008). The swishing sound is sometimes referred to as "Normal Amplitude Modulation," and the thumping sound as "Other Amplitude Modulation" (Oerlemans, 2015). "Random AM" is the typical time-varying situation of RTN close to streets with low to intermediate traffic density. To study the association of this source-specific AM to annoyance separately from source type, hypothetical situations of WTN with random AM and of RTN with periodic AM were also included in the study to obtain a complete factorial design. All stimuli contain some natural, random level fluctuations due to atmospheric turbulences.

B. Sound stimuli

For the listening tests, stimuli were generated either by sound synthesis (in the case of WTN) or by mixing of single

TABLE I. Factorial design of the listening tests with sound stimuli covering six different sound pressure levels (L_{Aeq}), two source types, and three AMs. "x" denotes studied stimuli.

	Source type					
	Wind turbine			Road traffic		
	AM					
L_{Aeq} [dB]	without	random	periodic	without	random	periodic
35	x	x	x			
40	x	x	x	x	x	x
45	x	x	x	x	x	x
50	x	x	x	x	x	x
55	x	x	x	x	x	x
60				x	x	x

pass-by recordings (in the case of RTN). No ambient sound was included in the stimuli.

1. WTN

Sound synthesis for WTN was realized using the tools of Pieren *et al.* (2014) and Heutschi *et al.* (2014), which were developed within the research project VisAsim (Manyoky *et al.*, 2014). As a sound source, one single 2 MW Vestas V90 turbine (three blades, hub height = 95 m, rotor diameter = 90 m, Vestas, Aarhus, Denmark) at an operation mode "strong wind" conditions was synthesized. The emission audio files (describing the sound source) with periodic AM and without AM were synthesized as described in Pieren *et al.* (2014). Periodic AM was generated with a standard deviation of the level fluctuation of 3 dB and a fluctuation frequency of 0.75 Hz. Random AM was generated as an amplitude modulated version of an emission file without AM. The AM was adjusted for a standard deviation of 3 dB, the varying fluctuation frequency was set to be comparable to periodic AM (range of 0.3–1.1 Hz). The resulting stimuli with random AM are similar to those with periodic AM except that the temporal pattern of the fluctuations is purely random.

On the emission signals, propagation filtering (Heutschi *et al.*, 2014) was applied for horizontal distances of 600, 350, 200, 100, and 60 m, approximately corresponding to L_{Aeq} values of 35–55 dB, assuming propagation over flat grassy terrain, a receiver height of 2 m, and accounting for geometric spreading, air absorption, ground reflection, and atmospheric turbulences. The stimuli were then fine-tuned in amplitude to exactly match the desired L_{Aeq}. The resulting synthesized single channel audio signals were converted into 2-channel (stereo) files (WAVE PCM format) by channel duplication.

2. RTN

To create the stimuli, 2-channel (stereo) recordings of individual car pass-by events were used and mixed to the desired road traffic scenarios presented in Table I. The recordings were taken at a straight interurban road with a speed limit of 80 km/h in a rural environment with flat terrain, at distances of 30 and 100 m. The car pass-by sound events were dominated by tire/road noise.

The recordings were made during a winter night (no snow) to minimize ambient sound, at a near-ground air temperature of −5 °C and a relative humidity of 86%. At both distances, two omnidirectional microphones (B&K type 4006; Brüel & Kjær, Nærum, Denmark) were installed with windscreens in a Jecklin Disk arrangement at a height of 1.7 m. Prior to the measurements, a calibration tone of 1 kHz emitting 94 dB was recorded on both channels using a B&K type 4231 calibrator (Brüel & Kjaer, Nærum, Denmark). The recording parameters were set to 44.1 kHz sampling frequency and 16 bit sample resolution on both portable digital audio recorders (type SD 702T; Sound Devices, LCC, Reedsburg, WI).

For scenarios with random and periodic AM, the recordings at 30 m were used. Subsequent event mixing was done with software developed for this study, assuming two traffic lanes with a density of 500 vehicles per hour and lane to obtain situations with clearly audible car pass-by events. For periodic AM, with cars of the two lanes passing at the same time, the constant time delay of 7.2 s between events (3600 s/500 vehicles) corresponds to a fluctuation frequency of 0.14 Hz. For random AM, a measured time delay distribution, determined from Swiss traffic meter data, was used. The distribution was strongly positively skewed, with a mode of 1.4 s and a mean value of 7.3 s. For situations without AM, the recordings at 100 m were used, and mixed assuming two traffic lanes with a density of 3000 vehicles per hour and lane (single cars hardly discriminable).

Propagation filtering was applied to the resulting audio signals by performing an overall spectral shaping due to atmospheric absorption and geometric spreading to account for differences in propagation distances between the recordings (30 or 100 m) and the desired situations (distances of 600, 400, 250, 120, and 40 m, corresponding to L_{Aeq} values of approximately 40–65 dB). Other effects were not accounted for in the propagation filtering of the recordings as they were considered not to substantially affect the acoustic impression of the stimuli. After propagation filtering, the stimuli were fine-tuned in amplitude to exactly match the desired L_{Aeq}. The resulting audio signals were 2-channel (stereo) files (WAVE PCM format).

3. Preliminary listening test—length of stimuli

In a preliminary test, an optimal stimuli length was determined, to assure unbiased rating (i.e., adequately long representation of the stimuli, particularly for RTN with random AM), while keeping it as short as possible to avoid unnecessarily long tests and/or impatience and fatigue of the participants.

The test procedure, software, and statistical analysis were very similar to those of the main listening tests described below. The participants were informed about the topic (noise annoyance), but not about the objective to determine stimuli length. Twelve persons (8 males, 4 females) participated in the tests.

The test consisted of two parts. In Part 1, a subset of three WTN and three RTN situations (Table I) was presented, each of them four times, with different lengths of 10, 20, 30, and 40 s (total of 24 stimuli). The participants were exposed to the stimuli in random order and rated them regarding annoyance. In Part 2, the participants were exposed to one of the RTN stimuli with random AM four times, with the above lengths, and classified the perceived length as "*too short*," "*spot-on*," or "*too long*," which was coded as "−1," "0," and "+1," respectively, for the subsequent analysis. The listening test lasted about 15 min.

The data was analyzed by means of linear mixed-effects models. The data of Part 1 did not reveal that stimuli length affects annoyance ($p = 0.52$). Similar results were found by Poulsen (1991) for lengths of 1–30 min. The data of Part 2 showed a quadratic dependence of the perceived length on the real stimulus length (Fig. 1), which was confirmed by the linear mixed effects model ($p = 0.03$). Further, the optimal length (*spot-on*) was found to be 20 s (Fig. 1).

For the main experiments, a stimulus length of 25 s was chosen. It is somewhat longer than the optimal length

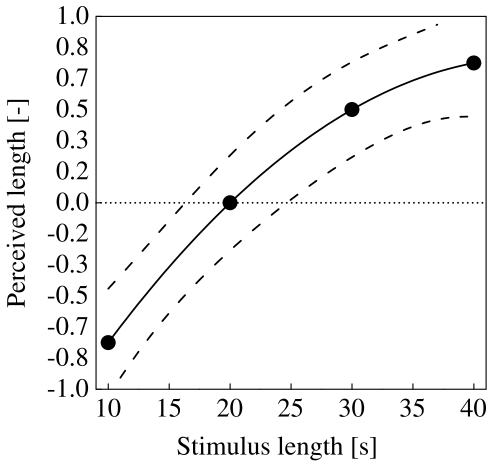

FIG. 1. Averaged perceived length as a function of the physical stimulus length of one RTN stimulus with an A-weighted equivalent continuous sound pressure level of 45 dB and random AM. Symbols represent observed values, and lines the corresponding mixed-effects model (solid line) with 95% CI (dashed lines). The perceived length scale covers values from −1 (*too short*) over 0 (*spot-on*; dotted horizontal line) to +1 (*too long*).

determined in the preliminary test to allow for adequate representation of the RTN situations with slow random AM (Fig. 2). The length of 25 s is comparable to the lengths in other focused listening tests on annoyance, with 5 s (Bolin *et al.*, 2012), 12.5 s (Torija and Flindell, 2015), 15 s (Seong *et al.*, 2013), 30 s (Jeon *et al.*, 2010; Lee *et al.*, 2011), 90 s (Legarth, 2007), or 180 s (Persson Waye and Öhrström, 2002), but substantially shorter than in non-focused listening tests with reading activity, with 450 s (Van Renterghem *et al.*, 2013) or 600 s (Persson Waye and Öhrström, 2002).

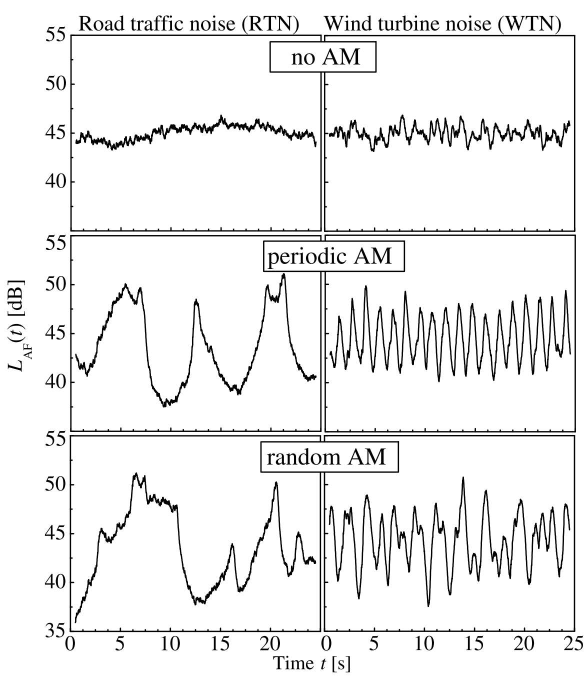

FIG. 2. Level-time histories of the A-weighted and FAST-time-weighted sound pressure level at observation time t [$L_{AF}(t)$] of the stimuli with an A-weighted equivalent continuous sound pressure level of 45 dB, for RTN (left) and WTN (right), without (top), with periodic (middle), or with random (bottom) AM.

4. Final set of stimuli and acoustic characteristics

In total, 30 stimuli representing the sound situations of Table I were established. Figure 2 shows exemplary level-time histories of the A-weighted and FAST-time-weighted sound pressure level at observation time t [$L_{AF}(t)$] and Fig. 3 shows the corresponding spectra.[1] The AM of the WTN and RTN stimuli are inherently different (Fig. 2). RTN has a much lower fluctuation frequency range than WTN (0.14 vs 0.75 Hz), and the AM is more irregular in RTN than in WTN. The standard deviations of the of $L_{AF}(t)$ of the WTN and RTN stimuli, in contrast, are of similar magnitude (WTN: range of 2.2–2.8 dB; RTN: 2.4–4.7 dB).

The synthesized WTN spectra are almost identical irrespective of AM. The recorded RTN spectra, in contrast, vary somewhat between stimuli (Fig. 3). In particular, the RTN stimuli without AM differ from those with periodic and random AM due to different recording distances. While WTN contains more sound energy than RTN at frequencies above 2 kHz, the RTN spectra dominate in the frequency range of 1–2 kHz (peak due to tire/road noise) and show a pronounced dip in the range of 500–600 Hz due to the ground effect (Fig. 3). Overall, WTN spectra contain more energy at low frequencies than RTN. This is also indicated by the differences between C- and A-weighted equivalent continuous sound pressure level, which are 3–6 dB larger for WTN than for RTN in the case of periodic and random AM, and 1–2 dB without AM, the disparate differences for the latter situations being due to the different recording distances of RTN.

Note that while some of the above sound situations do not occur in reality (namely, random AM of WTN and periodic AM of RTN, as well as random AM of RTN at low sound pressure levels where single car pass-by events are hardly discriminable), also these stimuli sounded plausible and realistic. In fact, none of the participants labeled them as being "unrealistic."

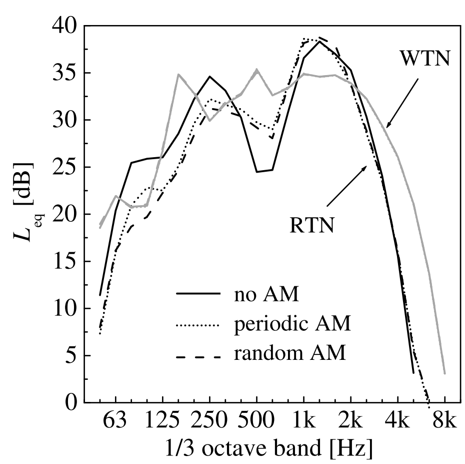

FIG. 3. A-weighted one-third octave band spectra (in L_{eq}) of the stimuli with an A-weighted equivalent continuous sound pressure level of 45 dB, for RTN (black lines) and WTN (gray lines) without (no), with periodic, or with random AM, averaged over the whole stimuli length. Note that the WTN spectra are almost identical.

C. Annoyance ratings and questionnaire

The aim was to study the short-term noise annoyance reactions to outdoor WTN and RTN situations during the day. The participants were therefore asked to rate their annoyance to the stimuli with the ICBEN 11-point scale of ISO/TS 15666 (2003), by answering the following question [in German, modified from ISO/TS 15666 (2003) and Legarth (2007)]: "When you imagine that this is the sound situation in your garden, what number from 0 to 10 represents best how much you would be bothered, disturbed or annoyed by it?"

The listening tests were complemented with a questionnaire. The first part contained questions about hearing (questions of the *Swiss National Accident Insurance Fund*, SUVA) and well-being, and the second part questions on the participants' attributes gender, age, living environment, noise sensitivity, and attitude toward WTN and RTN.

Noise sensitivity was determined with the "Noise-Sensitivity-Questionnaire" NoiSeQ by Schütte *et al.* (2007), which ranges from 0 ("noise-insensitive") to 3 ("highly noise-sensitive"), since noise sensitivity may significantly influence annoyance rating (Schütte *et al.*, 2007).

The participant's attitudes toward WTN and RTN were measured with a questionnaire developed in this study. Some questions were taken from a questionnaire by Pedersen (2007), partly modified, and complemented with further questions to cover the three attitude components affect, behavior, and cognition (Eagly and Chaiken, 1998). The questions are presented in the appendix. They were answered using a five-level rating scale ("strongly agree" = 4, "slightly agree" = 3, "neither/nor" = 2, "slightly disagree" = 1, and "strongly disagree" = 0), with some items having reverse values (see the appendix). To calculate the attitude toward WTN and RTN, the reverse values were first converted (i.e., 0 to 4, 1 to 3, etc.), and the mean values of the 10 items per source type were calculated to obtain a number from 0 to 4 covering a range from very negative to very positive attitude toward the source.

D. Main listening tests

1. Experimental setup

The listening tests were carried out in a semi-anechoic chamber. The stimuli were played back using a 3-channel stereo setup (left, center, right; Fig. 4). The loudspeakers (Focal CMS 50, Focal-JMlab, La Talaudière, France) were installed at a height similar to the seated participants' head at a distance of 150 cm from the participants (Fig. 4). The center speaker reproduced the sum of the left and right channel attenuated by 7 dB. This setup allowed the reproduction of the directional information of pass-by events of RTN, while the monaural WTN signal was more robustly localizable to frontal direction even if the participants' head moved during the listening test.

The background noise L_{Aeq} of the laboratory alone (<20 dB) was distinctly lower than the lowest $L_{AF}(t)$ of 27 dB occurring in the stimuli. Also, the computer used in the listening tests was kept away from the participants to avoid audibility of the ventilation. Background noise was therefore not expected to influence the participants' perception of the stimuli. Prior to the tests, the playback chain was calibrated with a sound level meter located at the position of the seated participants' head.

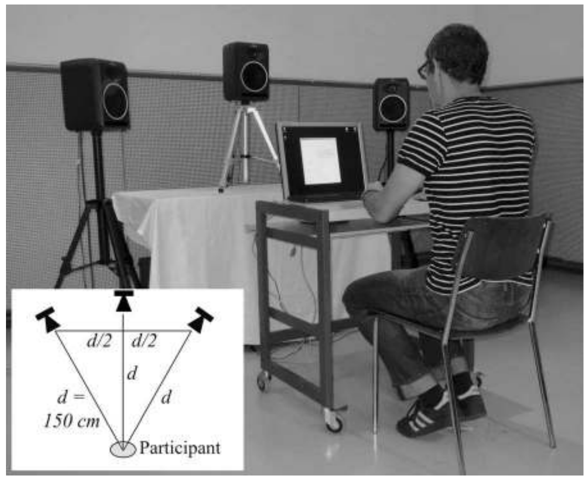

FIG. 4. Photography and layout (inlet figure, with listening distance *d*) of the laboratory setup used for the listening tests.

2. Test procedure

The experiments were done as focused tests, i.e., the participants had to deliberately listen to the stimuli and rate them during or directly after play-back. The stimuli were played once only, one by one, after complete play-back and rating of the previous one, with a break of 1 s between stimuli. Each stimulus was rated once only. While a second rating would have allowed assessing the repeatability of the participants' ratings, it would have substantially prolonged the tests. The participants performed the listening tests individually (one participant at a time). A program developed for this study guided the participants through the whole test, by automatically choosing and playing the stimuli, and by recording the participants' annoyance ratings as well as the time since start of the stimuli used to enter the rating, referred to as "rating time" in the following account. The ratings were entered by the participants via a graphical user interface.

Prior to the experiments, the participants were given a short introduction on the research topic (effects of WTN compared to RTN) and on their task in the experiment, omitting any details potentially biasing their annoyance ratings. The participants signed a consent form to participate in the study. Thereafter, they answered the first part of the questionnaire (hearing, well-being). None of the participants included in the study wore a hearing aid, and all of them declared to have normal hearing and to feel well (without cold).

The participants were then instructed about the program. Thereafter they started the actual listening test. First, as an orientation, they were exposed to five 10-s long stimuli covering the range of situations to be rated. This orientation set the frame of reference ("anchor") for the range of stimuli presented in the subsequent main experiment. Second, to get used to their task and the 11-point scale, they did two exercise ratings. Finally, the main experiment was conducted with the 30 experimental stimuli (Table I). At first, the 24 stimuli with L_{Aeq} of 40–55 dB were reproduced in random order. Thereafter, the remaining 6 stimuli with L_{Aeq} of 35 and 60 dB were reproduced in balanced order. The latter stimuli were reproduced separately to avoid potential bias of the ratings of the stimuli with a L_{Aeq} of 40–55 dB by too large step changes in L_{Aeq} between stimuli. After the experiments, the ratings of the six stimuli were checked for such bias by visual inspection of the data. The individual as well as the averaged ratings as a function of the L_{Aeq} look plausible over the whole studied L_{Aeq} range, including the additional stimuli (cf. Figs. 7 and 8, discussed below). The corresponding potential bias was therefore deemed to be negligible, and also these ratings were included in the analysis.

After the experiment, the participants completed the second part of the questionnaire. The whole listening test including the introduction and the questionnaire lasted about 1 h.

E. Participants

Sixty mostly naive (untrained) participants were recruited for the listening tests. The majority worked at a research institution in Dübendorf, Switzerland, either at the authors' institution, Empa, or at the adjacent institution, Eawag.

Thirty-one males and 29 females, aged from 18 to 60 yrs (median of 35 yrs), with normal hearing (see above), participated in the listening tests. The wide age range allowed checking for a possible dependency of annoyance on age (Van Gerven *et al.*, 2009). The participants covered a wide range of noise sensitivities with values of 0.6–2.6 (median of 1.5), i.e., most participants were moderately noise sensitive. Their attitude toward WTN with values of 1.6–3.8 (median of 3.0), was more positive than toward RTN with values of 0.4–2.9 (median of 1.7). Further, 70% of the participants preferred Swiss politics to focus more on quietness and environmental protection instead of economic growth, and 30% vice versa.

The participants' living environments covered areas from rural (52%) to urban (48%) and from quiet (72%) to loud (28%). Thirty-three percent of the participants lived close to a street with traffic calming, 52% close to a side road, and 15% close to a main road. Only half of the participants had in reality heard WTN prior to the experiments, and none of them lived close to wind turbines.

F. Resulting data set

In the listening tests, a data set of 1800 responses (annoyance ratings and rating times) was recorded (60 participants × 30 stimuli).

In addition, the annoyance ratings were transformed into the binary variable "high annoyance (HA)." HA was defined as 1 ("highly annoyed") for annoyance ratings equal to or larger than 8 (UZH and Empa, 1974; Schultz, 1978), i.e., for the top 27% of the 11-point scale, and else as 0. The same cutoff value has been used in noise effect studies in Switzerland since the 1970s (UZH and Empa, 1974), based on which the limit values of the Swiss legislation (NAO, 1986) were established.

As the cutoff value of 27% is arbitrary, a sensitivity analysis was done, where the results of HA with the cutoff of 27% were compared to those with a cutoff of 36% (ratings $\geq$7) and of 18% (ratings $\geq$9). The analysis revealed that, while the observed relative frequencies of HA strongly depend on the cutoff value, the associations of L_{Aeq}, source, and AM with HA are similar (not shown). Below, only the results for HA with the cutoff value of 27% are presented.

G. Statistical analysis

The statistical analysis was carried out with IBM SPSS Version 22. Tested effects (see below) were considered significant if the probability (p) of the observed results, or more extreme results, under the null hypothesis was $\leq$0.05.

1. Consistency of the individual responses

The consistency of the annoyance ratings and rating time across participants was assessed with the inter-rater reliability (Hallgren, 2012), using a two-way random, consistency, average-measures intraclass correlation [ICC(C,k)] (McGraw and Wong, 1996), where C denotes consistency and k is the number of independent measurements (i.e., the 60 participants) used to determine the average. A large ICC value indicates that the participants generally agree in their annoyance ratings concerning the different stimuli.

2. Annoyance ratings and rating time

The associations of the acoustic characteristics given in Table I with annoyance ratings and rating time were analyzed by means of linear mixed-effects models. These models combine fixed effects (categorical variables with a certain number of levels), covariates (continuous explanatory variables), random effects (randomly chosen from a population with a large set of possible levels, i.e., the participants), and interactions (deviations from the additive model describing how the effect of one variable depends on the levels of another variable) to predict dependent variables (annoyance rating and rating time). Repeated observations per participant (here, 30 ratings and rating times), which have correlated errors, are accounted for by using a hierarchy of levels, the upper level being the participants and the lower level being the repeated ratings/rating times per participant (e.g., Pinheiro and Bates, 2000).

Given the experimental design, the major effects to be included in the model, i.e., L_{Aeq}, source type, and AM (cf. Table I), were *a priori* defined. In addition, interactions between the major effects, the sequence with which the stimuli had been played, and the participants' attributes, i.e., gender, age, noise sensitivity, attitude, preference of political focus (quietness and environmental protection vs economic growth), prior exposure to WTN (yes vs no) and living environment (loud vs quiet, urban vs rural), were studied regarding their link to annoyance. No interaction terms other than those between the major effects were added to the model, as the other variables were not of main interest. Finally, different random effect models (random intercept; random slopes depending on the major effects and sequence; different covariance structures) were tested. Thus, several models of different degrees of complexity were established and compared with respect to completeness (include all relevant variables), performance (data representation, significance of effects), and parsimony (keep the model as simple as possible). The models were compared using the Bayesian Information Criterion (BIC) (Schwarz, 1978), where the model with the lowest BIC is preferred. Non-significant variables and interactions were excluded from the final model. Based on these insights, the final models, presented further below in this article, were chosen.

Compliance with the model assumptions was visually confirmed by means of residual plots. The goodness-of-fit of the final models were assessed with the marginal (R^2_m) and conditional (R^2_c) coefficients of determination (Vonesh *et al.*, 1996). R^2_m represents the variance explained by the fixed factors and R^2_c the variance explained by the fixed plus random factors. R^2_m and R^2_c were quantified according to Nakagawa and Schielzeth (2013) and Johnson (2014).

3. Probability of HA

The association of the binary variable HA with the predictors given in Table I was analyzed by means of logistic regression (Hosmer and Lemeshow, 2000) to obtain the probability of HA to adopt a value of 1 (*p*HA). In this study, we intended to establish exposure-response curves representing an average *p*HA within the population. Therefore, generalized estimating equations (Liang and Zeger, 1986) were used to account for the repeated ratings of the participants, as they predict a population-averaged response (Hu *et al.*, 1998).

Where feasible the same predictor variables were used in the logistic regression model as in the linear mixed-effects model (see above) to allow for model comparison. Different working correlation structures to account for repeated observations were tested.

The model performance was assessed by determining the rate of correct predictions of the individual HA ratings derived from classification tables (Hosmer and Lemeshow, 2000), as well as by the coefficient of discrimination (Tjur, 2009). In analogy to the coefficient of determination used in ordinary linear regression, it takes values between 0 ("no discriminatory power") and 1 ("perfect discrimination"). Its value increases with increasing difference between the predicted *p*HA of the two (observed) HA categories 1 and 0, i.e., the larger the difference, the better the model can discriminate the two categories. The coefficient of discrimination thus has another interpretation than the ordinary coefficient of determination.

III. RESULTS

A. Analysis of the individual responses

Figure 5 shows boxplots of the individual annoyance ratings and rating times. While the annoyance ratings cover a wide range of the 11-point scale at any L_{Aeq}, there is a clear trend of short-term annoyance increasing with the L_{Aeq}, and of WTN to be associated with higher annoyance reactions than RTN [Fig. 5(a)]. While rating time varied strongly between individual ratings, it tended to be longer at medium L_{Aeq} (∼40–50 dB) than at high or low L_{Aeq}, and longer for RTN than for WTN [Fig. 5(b)]. The ICC of annoyance rating (0.993), resulting binary variable HA (0.983), and rating time (0.904) all lie in the "excellent" range of ICC >0.75 according to Cicchetti (1994), which suggests a high degree of agreement between participants (Hallgren, 2012).

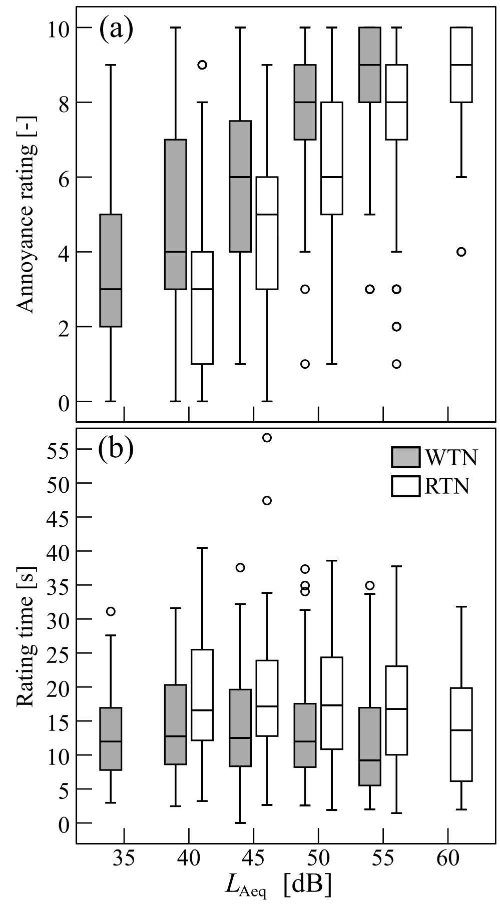

FIG. 5. Boxplots showing the median (50%, horizontal line in boxes), the first and third quantiles (25% and 75%, lower and upper boundaries of boxes), the whiskers comprising the data within 1.5 times the interquartile range, and outliers outside the whiskers, for (a) the individual annoyance ratings and (b) the rating times as a function of the A-weighted equivalent continuous sound pressure level (L_{Aeq}) of the stimuli representing WTN or RTN (pooled data of different situations of AM).

Both annoyance rating and rating time were affected by the sequence, i.e., the playback number, with which the stimuli had been played (Fig. 6). Annoyance rating tended to initially increase before reaching a "plateau," while rating time monotonously decreased. This suggests that the participants initially became increasingly annoyed by the stimuli, while forming their opinion ever quicker as they got accustomed to the sounds. Whether the plateau is (partly) evoked by the 6 (extreme) stimuli with L_{Aeq} of 35 and 60 dB played back at the end of the experiment is not known. The dependence of annoyance rating and

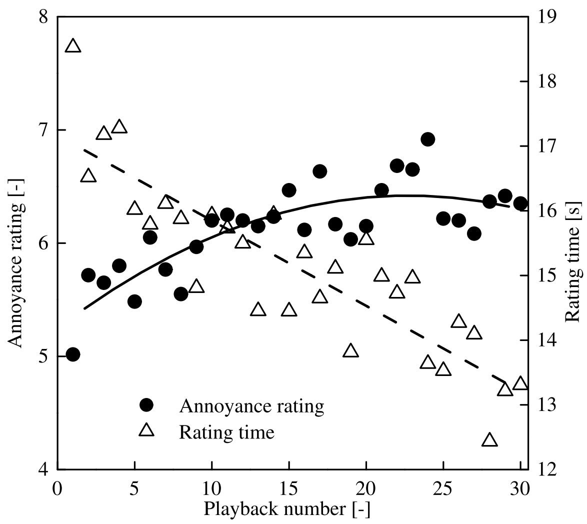

FIG. 6. Scatter diagram of annoyance ratings and rating times vs playback number. Annoyance ratings and rating times are averages of stimuli with the same playback number (pooled data of different situations of WTN, RTN, sound pressure levels, and AM). The lines represent quadratic (annoyance, solid) and linear fits (rating time, dashed).

rating time on the playback number can be described by a quadratic and linear fit, respectively (Fig. 6). The observations corroborate the importance of randomizing stimuli in listening tests. In contrast, none of the collected participants' attributes gender, age, noise sensitivity, or attitude were correlated to annoyance rating or rating time.

Since the annoyance rating is bounded at a value of 10, the participants' ratings tended to have a negatively correlated intercept (rating at low L_{Aeq}) and slope (dependence on L_{Aeq}), i.e., the larger the intercept, the smaller the slope and thus the smaller the dependence of the ratings on the L_{Aeq}, and vice versa (Fig. 7).

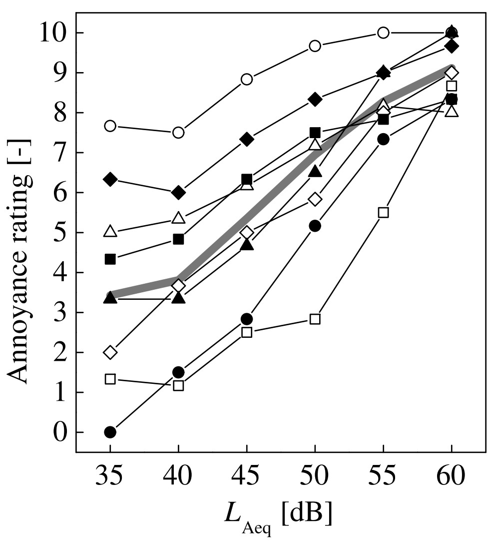

FIG. 7. Individual annoyance ratings (pooled data of different situations of WTN, RTN, and AM, averages per participant and A-weighted equivalent continuous sound pressure level [L_{Aeq}]) of eight participants as a function of the L_{Aeq}. Different symbols connected by lines represent different participants. The gray bold line shows the average of all 60 participants.

B. Evaluation of effects of acoustic characteristics

1. Annoyance

The averaged annoyance ratings are shown in Fig. 8. Annoyance increases linearly with L_{Aeq}, for any combination of source type and AM. Over the whole studied range of L_{Aeq}, WTN is associated with higher annoyance ratings than RTN [Fig. 8(a)], irrespective of whether AM is present or not. The association of AM with annoyance depends on the source type [Fig. 8(b)]. WTN without AM is linked to lower annoyance ratings than WTN with periodic or random AM, while the difference between the latter two is small. For RTN, the association of AM with annoyance is less clear, although periodic AM tends to be linked to lower annoyance ratings than random or no AM. The effects of source type and AM are pronounced at low L_{Aeq} and decrease with increasing levels. This is due to the fact that the ratings adopt values close to the maximum of 10 of the 11-point scale at large L_{Aeq}, irrespective of source and AM.

To describe these observed effects, the following mixed-effects model (SPSS procedure MIXED) was found to be appropriate [Eq. (1)]:

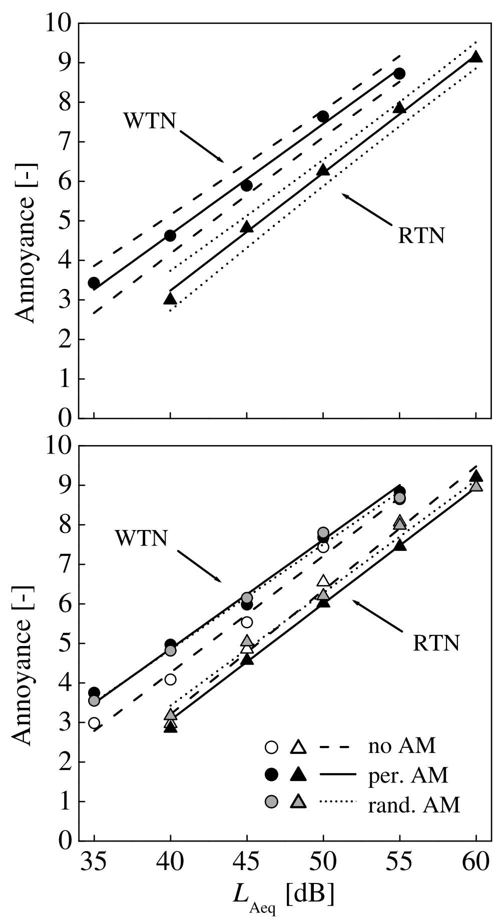

FIG. 8. Short-term annoyance (averaged values) as a function of the A-weighted equivalent continuous sound pressure level (L_{Aeq}) of (a) the pooled data (different situations of AM) of all WTN and RTN stimuli, and (b) WTN and RTN stimuli without (no), with periodic (per.), or random (rand.) AM. Symbols represent observed values (WTN: circles; RTN: triangles), and lines the corresponding mixed-effects model [Eq. (1)], in (a) with 95% CIs (dashed and dotted lines). The curves are shown at the mean playback number of the experiments.

$$
\begin{aligned}
y_{ijk} = {} & \mu + \tau_{\mathrm{Src},i} + \tau_{\mathrm{AM},j} + \beta \cdot L_{\mathrm{Aeq},ijk} \\
& + \tau_{\mathrm{Src}\times\mathrm{AM},ij} + \beta_{\mathrm{Src},i} \cdot L_{\mathrm{Aeq},ijk} + \beta_{\mathrm{AM},j} \cdot L_{\mathrm{Aeq},ijk} \\
& + \gamma \cdot S_{ijk} + \delta \cdot S_{ijk}^2 + u_{0k} + u_{1k} \cdot L_{\mathrm{Aeq},ijk} + \varepsilon_{ijk}. \qquad (1)
\end{aligned}
$$

In Eq. (1), y_{ijk} is the dependent variable (short-term) annoyance, μ is the overall mean, τ_{Src} and τ_{AM} are the fixed effects source type (2 levels: $i = 1, 2$) and AM (3 levels: $j = 1, 2, 3$), $L_{\mathrm{Aeq},ijk}$ and S_{ijk} are the covariates A-weighted sound pressure level and sequence (playback number), and β, γ, and δ are regression coefficients for the covariates. Further, $\tau_{\mathrm{Src}\times\mathrm{AM}}$, β_{Scr}, and β_{AM} represent interactions between the fixed effects (τ_{Src}, τ_{AM}) and the covariate (L_{Aeq}) of Table I. For example, β_{Scr} is the difference in β between WTN and RTN. Finally, the random effect terms u_{0k} and u_{1k} are the participants' random intercept and slope ($k = 1,\ldots,60$), and the error term ε_{ijk} is the random deviation between observed and expected values of y_{ijk}. The index ijk represents the kth replicate observation of the ith source at the jth AM.

The dependence of annoyance on sequence (cf. Fig. 6) is described by a linear and quadratic term ($\gamma \cdot S_{ijk}$, $\delta \cdot S_{ijk}^2$). The individual annoyance ratings (Fig. 7) are accounted for by correlated u_{0k} and u_{1k} terms, using an unstructured covariance matrix for that purpose. Neither the participants' tested attributes (gender, age, noise sensitivity, attitude, preference of political focus, prior exposure to WTN, living environment; $p = 0.32$–0.89), nor the three-fold interaction between source type, AM and L_{Aeq} ($p = 0.14$) were included as they were not significantly linked to the annoyance ratings. The model parameters are presented in Table II. The parameters can be combined to describe any combination of the variables of Table I.

The mixed-effects model of Eq. (1) explains a large part of the variance, even with the fixed effects alone (R^2_{m} of 0.55, R^2_{c} of 0.84). Accordingly, it predicts the observed annoyance with high accuracy and narrow confidence intervals (CIs) (Fig. 8). The model confirms statistical significance of the above observations. Source type, AM, L_{Aeq}, and sequence (playback number; linear and quadratic term) are all significantly linked to annoyance ($p = 0.00$). There are interactions between source type and AM ($p = 0.00$), AM and L_{Aeq} ($p = 0.01$), and in tendency also between source type and L_{Aeq} ($p = 0.06$), indicated by the slight convergence of the regression lines in Fig. 8.

Over the mutually studied L_{Aeq} range of 40–55 dB, WTN was linked to the same annoyance reactions at ∼4–5 dB lower L_{Aeq} than RTN [Fig. 8(a)]. The significance of this shift on the abscissa (L_{Aeq}) is indicated by the non-overlapping CIs of the model curves in Fig. 8(a), and confirmed by contrast analysis (not shown). Even without AM, WTN was associated with higher annoyance reactions than RTN over the studied L_{Aeq} range, with the same annoyance at ∼3–4 dB lower L_{Aeq} than RTN [Fig. 8(b)]. In the case of WTN, periodic AM was linked to the same annoyance reaction at ∼1–2 dB lower L_{Aeq} as without AM [Fig. 8(b)].

TABLE II. Model coefficients (Coeff.), with 95% CI and probabilities (p) of the linear mixed-effects model for the annoyance ratings and of the population-averaged logistic regression model for the probability of HA, and odds ratio (OR = exp[Coeff.]) with 95% CI for the logistic regression model. The parameters and symbols are explained in Eqs. (1) and (2).

		Linear mixed-effects model [Eq. (1)]			Population-averaged logistic regression model [Eq. (2)]				
Parameter	Symbol	Coeff.	95% CI	p	Coeff.	95% CI	OR	OR 95% CI	p
Intercept	μ	−6.6718	[−8.2022;−5.1414]	0.00	−12.0779	[−14.6398;−9.5159]	0.00	[0.00;0.00]	0.00
Source	$\tau_{\mathrm{Src},i=\mathrm{RTN}}$	−2.1503	[−3.087;−1.2135]	0.00	−2.6744	[−5.3424;−0.0063]	0.07	[0.00;0.99]	0.05
	$\tau_{\mathrm{Src},i=\mathrm{WTN}}$	0[a]			0[a]		1		
AM[b]	$\tau_{\mathrm{AM},j=\mathrm{no}}$	−1.7210	[−2.5691;−0.8729]	0.00	−1.2172	[−3.4954;1.0611]	0.30	[0.03;2.89]	0.30
	$\tau_{\mathrm{AM},j=\mathrm{per.}}$	−0.3509	[−1.1991;0.4972]	0.42	−0.1739	[−1.7383;1.3904]	0.84	[0.18;4.02]	0.83
	$\tau_{\mathrm{AM},j=\mathrm{rand.}}$	0[a]			0[a]		1		
L_{Aeq}	β	0.2666	[0.2386;0.2946]	0.00	0.2359	[0.1862;0.2856]	1.27	[1.20;1.33]	0.00
Source × AM	$\tau_{\mathrm{Src}\times\mathrm{AM},ij=\mathrm{RTN}\times\mathrm{no}}$	0.3616	[0.0855;0.6377]	0.01	0.6130	[0.0625;1.1634]	1.85	[1.06;3.20]	0.03
	$\tau_{\mathrm{Src}\times\mathrm{AM},ij=\mathrm{RTN}\times\mathrm{per.}}$	−0.3875	[−0.6636;−0.1113]	0.01	−0.1678	[−0.5522;0.2165]	0.85	[0.58;1.24]	0.39
	$\tau_{\mathrm{Src}\times\mathrm{AM},ij=\mathrm{RTN}\times\mathrm{rand.}}$	0[a]			0[a]		1		
	$\tau_{\mathrm{Src}\times\mathrm{AM},ij=\mathrm{WTN}\times\mathrm{no}}$	0[a]			0[a]		1		
	$\tau_{\mathrm{Src}\times\mathrm{AM},ij=\mathrm{WTN}\times\mathrm{per.}}$	0[a]			0[a]		1		
	$\tau_{\mathrm{Src}\times\mathrm{AM},ij=\mathrm{WTN}\times\mathrm{rand.}}$	0[a]			0[a]		1		
Source × L_{Aeq}	$\beta_{\mathrm{Src},i=\mathrm{RTN}}$	0.0184	[−0.001;0.0377]	0.06	0.0296	[−0.0224;0.0816]	1.03	[0.98;1.09]	0.26
	$\beta_{\mathrm{Src},i=\mathrm{WTN}}$	0[a]			0[a]		1		
AM × L_{Aeq}	$\beta_{\mathrm{AM},j=\mathrm{no}}$	0.0285	[0.0101;0.0469]	0.00	0.0150	[−0.0328;0.0627]	1.02	[0.97;1.07]	0.54
	$\beta_{\mathrm{AM},j=\mathrm{per.}}$	0.0093	[−0.0091;0.0277]	0.32	0.0037	[−0.0284;0.0358]	1.00	[0.97;1.04]	0.82
	$\beta_{\mathrm{AM},j=\mathrm{rand.}}$	0[a]			0[a]		1		
Seq. no.	γ	0.1101	[0.0831;0.1371]	0.00	0.0526	[0.0348;0.0703]	1.05	[1.04;1.073]	0.00
	δ	−0.0027	[−0.0036;−0.0018]	0.00	—				
Random intercept	u_{0k}	26.6738	[18.1436;39.2145]	0.00	—				
Random slope	u_{1k}	0.0079	[0.0053;0.0117]	0.00	—				
Residual	ε_{ijk}	1.3193	[1.2327;1.4118]	0.00	—				

[a]Redundant coefficients are set to zero.
[b]no = without AM; per. = periodic AM; rand. = random AM.

2. Rating time

The average rating times are shown in Fig. 9. Rating time approximately follows a quadratic function, tending to be longer at medium than at low or high L_{Aeq}. Further, rating time of RTN is 3–5 s longer than of WTN (Fig. 9). Apparently low or high L_{Aeq} are associated with low annoyance or HA, while medium L_{Aeq} seem to be more difficult to rate. Also, the low level fluctuation frequency of RTN (Fig. 2) forced the participants to listen to a large part of the stimuli, while the decision formation was quicker for WTN. Correspondingly, the rating time for RTN increased in the order, without AM < periodic AM < random AM, while the effect of AM was less distinct for WTN (not shown). The mixed-effects model analysis confirms the statistical significance of the above observations (not shown).

3. Probability of HA

Figure 10 shows the averaged observed relative frequencies of HA (HA = 1) for WTN and RTN. The observed data approximately show a sigmoid dependence on L_{Aeq}, for any combination of source type and AM. In line with the annoyance rating, WTN is linked to higher relative frequencies of HA than RTN [Fig. 10(a)]. Further, WTN with random and periodic AM are linked to higher relative frequencies of HA than without AM. For RTN the effect of AM is less pronounced [Fig. 10(b)], although, in contrast to WTN, random and in particular, also periodic AM tend to be associated with lower frequencies of HA than no AM.

To describe these effects, i.e., to predict the averaged probabilities of HA (pHA), the following population-averaged logistic regression model (SPSS procedure GENLIN) was found to be appropriate:

$$\mathrm{logit}\,(p\mathrm{HA}) = \mu + \tau_{\mathrm{Src},i} + \tau_{\mathrm{AM},j} + \beta \cdot L_{\mathrm{Aeq},ijk} + \tau_{\mathrm{Src}\times\mathrm{AM},ij} + \beta_{\mathrm{Src},i} \cdot L_{\mathrm{Aeq},ijk} + \beta_{\mathrm{AM},j} \cdot L_{\mathrm{Aeq},ijk} + \gamma \cdot S_{ijk}. \quad (2)$$

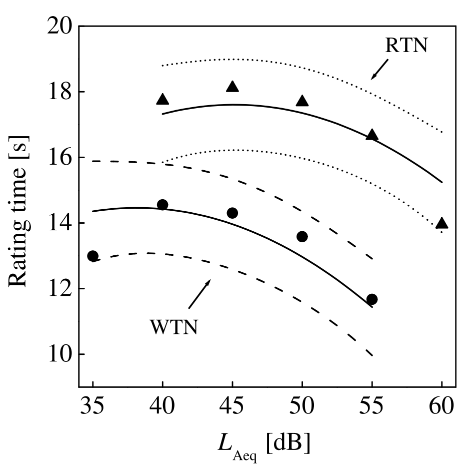

FIG. 9. Rating time (averaged values) as a function of the A-weighted equivalent continuous sound pressure level (L_{Aeq}) of the pooled data (different situations of AM) of all WTN and RTN stimuli. Symbols represent observed values (WTN: circles; RTN: triangles), and lines the corresponding mixed-effects model (solid line) with 95% CIs (dashed and dotted lines). The curves are shown at the mean playback number of the experiments.

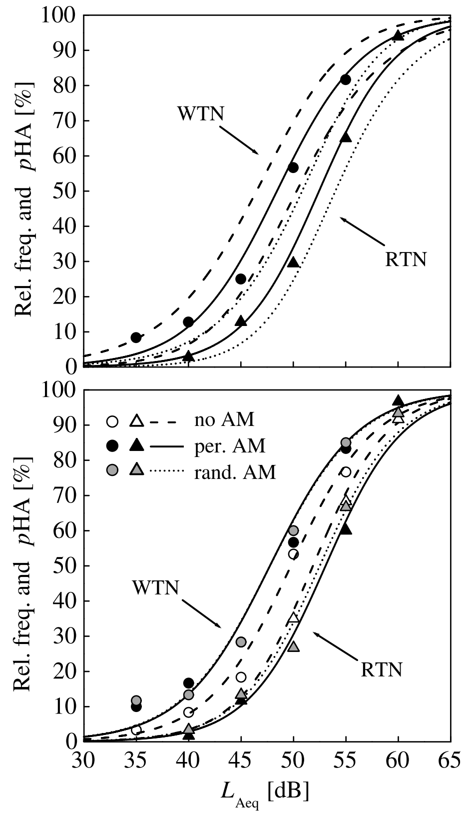

FIG. 10. Relative frequencies (rel. freq.; symbols) and predicted probability of HA (pHA; lines) as a function of the A-weighted equivalent continuous sound pressure level (L_{Aeq}) of (a) the pooled data (different situations of AM) of all WTN and RTN stimuli, and (b) WTN and RTN stimuli without (no), with periodic (per.), or random (rand.) AM. Symbols represent observed values (WTN: circles; RTN: triangles), and lines the corresponding logistic regression model [Eq. (2)], in (a) with 95% CIs (dashed and dotted lines). The curves are shown at the mean playback number of the experiments. Note that in (b) the WTN curves with periodic and random AM are almost identical.

In Eq. (2), logit(pHA) = ln(pHA/[1 − pHA]) is the logit for pHA (for details see, e.g., Hosmer and Lemeshow, 2000), and the other variables have the same notation as in Eq. (1) for short-term annoyance ratings. Repeated observations are accounted by an exchangeable working correlation structure (Hu *et al.*, 1998), which is a practical choice for small samples (Jang, 2011). In contrast to Eq. (1) no quadratic term for sequence was included in Eq. (2) as it was highly non-significant ($p = 0.85$). Apart from that, the same variables (also non-significant ones) were included to allow for direct comparison with the annoyance model of Eq. (1). The model parameters are presented in Table II. Again, the parameters can be combined to describe any combination of the variables of Table I.

The model of Eq. (2) predicts the individual ratings satisfyingly, with a coefficient of discrimination (Tjur, 2009) of 0.42 and rate of correct HA predictions of 82%. Further, it closely represents the observed averaged relative frequencies of HA which are of interest here (Fig. 10). The model confirms the statistical significance of the effects observed above. L_{Aeq} and sequence ($p = 0.00$) but also source type ($p = 0.06$) are linked to pHA, while AM is associated with pHA by a significant interaction with source type ($p = 0.02$), i.e., its effect differs between WTN and RTN [Fig. 10(b)]. In contrast to short-term annoyance [Fig. 8 and Eq. (1)] there were no significant interactions between L_{Aeq} and source type ($p = 0.26$) or L_{Aeq} and AM ($p = 0.82$). The associations of the investigated variables with *p*HA are thus comparable to, but somewhat less pronounced than with annoyance. Over the studied L_{Aeq} range of 40–55 dB, the resulting shifts of the model curves of *p*HA on the abscissa (Fig. 10) are very similar to those of short-term annoyance (Fig. 8). The WTN and RTN curves (pooled over different AM situations) are shifted by ~3–5 dB L_{Aeq} [Fig. 10(a)]. Further, the curves of WTN without AM and RTN are shifted by ~2–3 dB, and those of WTN with periodic and without AM by ~2 dB [Fig. 10(b)].

IV. DISCUSSION

In this study, focused laboratory listening tests were conducted to investigate and compare the short-term annoyance reactions to different WTN and RTN situations and to establish exposure-response curves for the probability of HA (*p*HA). The factorial design and the fully controlled sound stimuli not only allowed for exclusion of effect modifiers inherent to field surveys, but also for separation of the individual associations of sound pressure level, source type, and AM with annoyance reactions. The observed differences between WTN and RTN are therefore exclusively attributable to acoustic characteristics.

A. Acoustic characteristics associated with noise annoyance

Within the studied L_{Aeq} range of 35–55 dB, strong short-term annoyance reactions to WTN were observed. The annoyance ratings of 3–9 on the ICBEN 11-point scale (Fig. 8) are similar to those of other focused tests with values of 1–8 for comparable sound pressure levels (Legarth, 2007; Lee *et al.*, 2011; RenewableUK, 2013; Seong *et al.*, 2013). Unfocused tests (including a reading task) by Persson Waye and Öhrström (2002), in contrast, yielded somewhat lower ratings of 2–3 at a L_{Aeq} of 40 dB. Further, the annoyance ratings of 3–9 for RTN in the L_{Aeq} range of 40–60 dB (Fig. 8) are higher than in a focused test by Jeon *et al.* (2010) with ratings of 1–4 for the same sound pressure level range. In this study the resulting *p*HA of WTN within an L_{Aeq} range of 35–45 dB was found to be 2%–34% [Fig. 10(b)] which, interestingly enough, is very similar to the *p*HA of ~4%–30% (outdoor annoyance) found in field studies by Janssen *et al.* (2011), while larger than the ~7%–16% found by Michaud *et al.* (2016a), both for similar sound pressure levels. For RTN, a *p*HA of 3%–91% was found within a L_{Aeq} range of 40–60 dB [Fig. 10(b)]. This *p*HA range is substantially larger than the 1%–12% determined by Miedema and Oudshoorn (2001) in a meta-analysis of earlier field studies and the 5%–25% found in field studies by Yokoshima *et al.* (2012), but of similar magnitude as the 6%–60% determined in a field study by Lercher *et al.* (2008), for similar sound pressure levels.

The observed annoyance is strongly linked to the L_{Aeq}. This confirms recent findings of other laboratory experiments that an A-weighted metric is an appropriate predictor at least for (source-specific) short-term annoyance to WTN (Bolin *et al.*, 2014) as well as RTN (Jeon *et al.*, 2010; Torija and Flindell, 2015), and thus possibly also for annoyance reactions to long-term exposure. In interpreting these results, one has to consider the strong relation between short-term annoyance and perceived loudness, and also the weak association of acoustic characteristics with long-term annoyance assessed in field surveys (see Sec. IV B). Further, as propagation filtering was applied in generating the stimuli, the L_{Aeq} was varied along with the spectrum. These variables (L_{Aeq} and spectrum) are thus confounded, i.e., their effects cannot be distinguished. However, for the considered propagation distances of ≤600 m the L_{Aeq} is expected to be the dominant effect. Despite the strong dependence of short-term annoyance on the L_{Aeq}, the differences between WTN and RTN prove that other acoustic characteristics need to be considered as well.

In particular, source type is important. WTN was found to be more annoying than RTN (Figs. 8 and 10). This result is in line with findings from field surveys (Janssen *et al.*, 2011), while only small differences between WTN and RTN were observed in a study by Pedersen *et al.* (2010). Over the L_{Aeq} range of 40–55 dB, WTN was linked to the same *p*HA at ~3–5 dB lower L_{Aeq} than RTN. While this "purely acoustic" shift is pronounced, it is much smaller than the shift of ~15–20 dB determined by Janssen *et al.* (2011) for outdoor WTN with a L_{Aeq} of ~35–40 dB, or of 6–9 dB according to Kuwano *et al.* (2014) for WTN with a L_{Aeq} of ~30–50 dB, or of 16 dB revealed by Michaud *et al.* (2016b). The larger shift determined in field surveys may reflect that other, non-acoustic variables play an important role, which were excluded in the present study. Contrasting our findings, in a laboratory study by Van Renterghem *et al.* (2013), WTN was found to be similarly or even less annoying than RTN, depending on the road situation. In the latter study, however, an unfocused listening test including a reading task was performed for indoor noise, without disclosing to the participants which sound sources they were going to be exposed to.

In addition, also AM (partly) determines annoyance. The increased annoyance reactions to WTN with periodic AM are in agreement with previous studies (Lee *et al.*, 2011; RenewableUK, 2013; Ioannidou *et al.*, 2016). The limited influence of AM in the case of RTN (Figs. 8 and 10) contrasts with findings of Lercher *et al.* (2008) and Van Renterghem *et al.* (2013) that RTN with random AM ("local roads," "main roads") was linked to significantly higher annoyance than without AM (highway). However, as it is not known to what degree the acoustic characteristics (vehicle mix, traffic density; AM, spectra) of the above studies coincide with those of the present study, also the comparability of the results is limited. Regarding AM, two findings are particularly interesting. First, the effect of AM on annoyance was different for WTN and RTN. While the standard deviation of the level fluctuation of WTN and RTN was of similar magnitude, level fluctuation frequency range strongly differed (Fig. 2). This indicates that possibly the latter influences annoyance. The (subjective) hearing sensation of AM at level fluctuation frequencies below 20 Hz is described with the psychoacoustic parameter fluctuation strength (Fastl and Zwicker, 2007). Fluctuation strength reaches its maximum at a fluctuation frequency of 4 Hz (Fastl, 1982). The level fluctuation frequency of WTN (0.75 Hz) is relatively close to 4 Hz. The level fluctuation frequency of RTN (∼0.14 Hz), in contrast, was apparently too low to evoke this sensation. Second, the participants did not discriminate between periodic and random AM in their annoyance rating of WTN, i.e., periodicity was not a particularly annoying acoustic characteristic. However, this might have been different if the participants had lived close to wind turbines, thus being accustomed to WTN and potentially recognizing random AM as unrealistic. The results suggest that annoyance reactions to WTN may be at least partially reduced if the occurrence of periodic AM can be ruled out or at least strongly reduced, e.g., by blade pitch control (Makarewicz and Gołębiewski, 2015) or an operational approach (Bockstael *et al.*, 2012).

B. Comparability of the results with field surveys

In the above discussion it is worth noting that results from field and laboratory studies are of limited comparability due to inherent differences. In field surveys, people are not exposed to specific sound situations while being interviewed, but rather rate their annoyance based on their memory of the last "12 months or so" (ISO/TS 15666, 2003) which comprises different (outdoor and indoor) sound exposures. In particular, also recollection of nighttime sound exposure (and thus of sleep disturbance) is included. Further, in field surveys individual attributes of the participants such as noise sensitivity or attitude were found to significantly affect annoyance (see Sec. I), which was neither observed here nor in a laboratory study by Legarth (2007), and only partly in a laboratory study by Crichton *et al.* (2015). This is most probably due to the fact that in laboratory experiments, participants' ratings are closely related to the sensory perception of the sounds present at the time of rating. Consistent with this, laboratory annoyance ratings are usually highly correlated with perceived loudness (Guski and Bosshardt, 1992), which in turn strongly depends on the (physical) sound pressure level. However, loudness and annoyance seem discriminable also in the laboratory (e.g., Kuwano *et al.*, 1988). In the field, in contrast, various other factors, besides sound pressure level, may play a (more pronounced) role. Context (field vs laboratory studies) therefore is an important influencing factor for annoyance and needs to be accounted for when comparing studies.

For the present study, the comparability of the results with annoyance associated with long-term exposure in the field is limited due to the following reasons. First, the participants of the study represent a wide and balanced range of age, gender, noise sensitivity, and attitude, but only a limited geographic region and working environment. In particular, the study includes no residents living close to wind farms, who might react differently. Bolin *et al.* (2014) found that residents close to wind farms were more annoyed by WTN than non-affected participants, which might be linked to (increased) recognition of WTN (Van Renterghem *et al.*, 2013). Second, the loudspeakers used in this experiment reproduce frequencies down to ∼50 Hz, while WTN has considerable sound energy also below (Møller and Pedersen, 2011). Thus, low-frequency noise (∼20–200 Hz), which may additionally contribute to annoyance (Pawlaczyk-Łuszczyńska *et al.*, 2003), was only partly covered. Third, the annoyance question of this study ("*When you imagine…*") is different to the original question of ISO/TS 15666 (2003) ["*Thinking about the last (12 months or so)…*"], the former involving imagination of a hypothetic location and the latter an integration of annoyance over a longer time period. Fourth, WTN does not cease or at least decrease at night, in contrast to many other sources, which might additionally contribute to annoyance (Pedersen *et al.*, 2009). Finally, non-acoustic effect modifiers (e.g., individual characteristics such as attitude), which are always present in the field, were excluded or at least controlled to study acoustic characteristics alone and thus to establish a closer relationship to (short-term) annoyance. However, such non-acoustic variables may be crucial for long-term annoyance assessed in field surveys (Janssen *et al.*, 2011), and the association of annoyance with WTN characteristics alone may be weak. In a recent field survey, the Health Canada study (see overview by Schomer and Fidell, 2016), WTN characteristics yielded an R^2 of only 9%, while 10 additional variables increased R^2 to 58% (Michaud *et al.*, 2016b). However, the survey covered sound pressure levels of up to 46 dB only, while the present study included L_{Aeq} of up to 55 dB. Also, only two variables were found to be equally or more important than WTN characteristics, namely, "*annoyance with blinking lights*," increasing R^2 by +9%, and "*closure of bedroom window due to wind turbines [as noise source]*" (+30%), and the latter by necessity is related to WTN characteristics. Acoustic characteristics, while one of various variables only, are therefore not negligible.

Thus, this laboratory study reliably discloses acoustic characteristics of WTN and RTN linked to short-term annoyance. Yet, the generalizability of the results to long-term exposure in the field still needs to be verified. The high control of effect modifiers, which is the strength of laboratory studies, is at the expense of ecological validity. For field surveys, the opposite is true (less control, but higher ecological validity). Laboratory studies and field surveys are therefore complementary.

V. CONCLUSIONS

In the present laboratory study, WTN was found to be associated with higher annoyance reactions than RTN at the same L_{Aeq}, particularly when AM was present, but also for quasi-stationary (constant over time) signals. The increased annoyance reactions to AM of wind turbines are not related to the periodicity, whereas they seem to depend on the modulation frequency range. The AM of RTN, in contrast, was less clearly linked to annoyance. As visual factors were excluded from the experiments, the observed differences in annoyance reactions to wind turbines and road traffic are associated exclusively with their acoustic characteristics. The study discloses a direct link of acoustic characteristics of wind turbines and road traffic to annoyance reactions, yet the generalizability to long-term exposure in the field still needs to be verified, even more so as in field surveys non-acoustic variables were found to be at least as crucial for annoyance reactions as acoustic characteristics of WTN.

ACKNOWLEDGMENTS

The authors are very grateful to the participants of the preliminary and main experiments of this study. Further, they would like to thank Sarah Grimm and Madeleine Manyoky for their valuable inputs regarding the statistical analysis. This study was funded by the Swiss Federal Office for the Environment (Assignment No. 5211.00723.100.02). According to Swiss cantonal and federal law this study was not subject to approval by any ethics committee. However, the Cantonal Ethics Committee KEK Zurich, after having checked the project, stated that from an ethical point of view there was no objection to carrying out the study [Waiver No. 40-2015 (KEK-ZH-No. 2014-0294) from April 22, 2015]. The authors declare no conflicts of interest regarding the work presented in this article.

APPENDIX: ATTITUDE QUESTIONNAIRE

Item No.	Attitude component	Reverse value[a]	Item[b]
1	Affect	—	Ich finde Windkraftanlagen gut. (*I think that wind turbines are good.*)
2	Cognition	—	Windkraftanlagen sind nützlich für die Gesellschaft. (*Wind turbines are beneficial for society.*)
3	Behavior	Yes	Ich würde nicht in die Nähe von Windkraftanlagen ziehen. (*I would not move to the vicinity of wind turbines.*)
4	Cognition	Yes	Windkraftanlagen sind ungesund für Anwohner. (*Wind turbines are unhealthy for residents.*)
5	Behavior	—	Ich würde für den Ausbau von Windkraftanlagen stimmen. (*I would vote for the development of wind turbines.*)
6	Cognition	Yes	Windkraftanlagen tragen zur Umweltverschmutzung bei. (*Wind turbines contribute to environmental pollution.*)
7	Affect	Yes	Windkraftanlagen wirken auf mich bedrohlich. (*Wind turbines are threatening to me.*)
8	Cognition	Yes	Windkraftanlagen stören die Landschaft. (*Wind turbines disturb the landscape.*)
9	Behavior	—	Ich wäre bereit, für die Förderung von Windkraftanlagen mehr zu bezahlen. (*I would be willing to pay more for the funding of wind turbines.*)
10	Affect	Yes	Windkraftanlagen nerven mich. (*Wind turbines annoy me.*)
11	Affect	—	Ich finde Strassen gut. (*I think that roads are good.*)
12	Cognition	—	Strassen sind nützlich für die Gesellschaft. (*Roads are beneficial for society.*)
13	Behavior	Yes	Ich würde nicht in die Nähe verkehrsreicher Strassen ziehen. (*I would not move to the vicinity of busy roads.*)
14	Cognition	Yes	Strassenverkehr ist ungesund für Anwohner. (*Road traffic is unhealthy for residents.*)
15	Behavior	—	Ich würde für den Ausbau des Strassenverkehrsnetzes stimmen. (*I would vote for the development of the road network.*)
16	Cognition	Yes	Strassenverkehr trägt zur Umweltverschmutzung bei. (*Road traffic contributes to environmental pollution.*)
17	Affect	Yes	Strassenverkehr wirkt auf mich bedrohlich. (*Road traffic is threatening to me.*)
18	Cognition	Yes	Strassen stören die Landschaft. (*Roads disturb the landscape.*)
19	Behavior	—	Ich wäre bereit, für die Förderung des Strassenverkehrsnetzes mehr zu bezahlen. (*I would be willing to pay more for the funding of the road network.*)
20	Affect	Yes	Strassenverkehr nervt mich. (*Road traffic annoys me.*)

[a]Values of 0–4, 0 indicating a very negative and 4 a very positive attitude for non-reverse values, and vice versa for reverse values.
[b]The German questions were used in the listening tests. The English translations in parentheses are added for readers' convenience.

[1]See supplementary material at http://dx.doi.org/10.1121/1.4949566 for the compressed audio files (MP3 format) of these stimuli to get an audio impression.

BAFU (**2009**). Noise Pollution in Switzerland. Results of the SonBase National Noise Monitoring Programme. State of the Environment No. 0907. [Federal Office for the Environment (BAFU/FOEN), Berne, Switzerland], available online at http://www.bafu.admin.ch/publikationen/publikation/01036/index.html?lang=en (Last viewed March 17, 2016).

Bockstael, A., Dekoninck, L., Can, A., Oldoni, D., De Coensel, B., and Botteldooren, D. (**2012**). "Reduction of wind turbine noise annoyance: An operational approach," Acta Acust. Acust. **98**, 392–401.

Bolin, K., Bluhm, G., and Nilsson, M. E. (**2014**). "Listening test comparing A-weighted and C-weighted sound pressure level as indicator of wind turbine noise annoyance," Acta Acust. Acust. **100**, 842–847.

Bolin, K., Kedhammar, A., and Nilsson, M. E. (**2012**). "The influence of background sounds on loudness and annoyance of wind turbine noise," Acta Acust. Acust. **98**, 741–748.

Bowdler, D. (**2008**). "Amplitude modulation of wind turbine noise. A review of the evidence," Acoust. Bull. Inst. Acoust. **33**, 31–35.

Cicchetti, D. V. (**1994**). "Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology," Psychol. Assess. **6**, 284–290.

Crichton, F., Dodd, G., Schmid, G., and Petrie, K. J. (**2015**). "Framing sound: Using expectations to reduce environmental noise annoyance," Environ. Res. **142**, 609–614.

Eagly, A. H., and Chaiken, S. (**1998**). "Attitude structure and function," in *The Handbook of Social Psychology*, 4th ed., edited by D. T. Gilbert, S. T. Fiske, and G. Lindzey (McGraw-Hill, New York, NY), Vol. 1, Chap. 7, pp. 269–322.

Fastl, H. (**1982**). "Fluctuation strength and temporal masking patterns of amplitude-modulated broadband noise," Hear. Res. **8**, 59–69.

Fastl, H., and Zwicker, E. (**2007**). *Psychoacoustics: Facts and Models*, 3rd ed. (Springer-Verlag, Berlin, Heidelberg, Germany), 462 pp.

Guski, R., and Bosshardt, H.-G. (**1992**). "Gibt es eine 'unbeeinflußte' Lästigkeit?" ("Does 'unbiased annoyance' exist?"), Z. Lärmbekämpf. **39**, 67–74.

GWEC (**2015**). Global Wind Report, Annual Market Update 2014 [Global Wind Energy Council (GWEC), Brussels, Belgium], available online at http://www.gwec.net/wp-content/uploads/2015/03/GWEC_Global_Wind_2014_Report_LR.pdf (Last viewed March 17, 2016).

Hallgren, K. A. (**2012**). "Computing inter-rater reliability for observational data: An overview and tutorial," Tutor. Quant. Methods Psychol. **8**, 23–34.

Heutschi, K., Pieren, R., Müller, M., Manyoky, M., Wissen Hayek, U., and Eggenschwiler, K. (**2014**). "Auralization of wind turbine noise: Propagation filtering and vegetation noise synthesis," Acta Acust. Acust. **100**, 13–24.

Hosmer, W. D., and Lemeshow, S. (**2000**). *Applied Logistic Regression*, 2nd ed. (John Wiley & Sons, Inc., Toronto, Canada), 375 pp.

Hu, F. B., Goldberg, J., Hedeker, D., Flay, B. R., and Pentz, M. A. (**1998**). "Comparison of population-averaged and subject-specific approaches for analyzing repeated binary outcomes," Am. J. Epidemiol. **147**, 694–703.

Ioannidou, C., Santurette, S., and Jeong, C.-H. (**2016**). "Effect of modulation depth, frequency, and intermittence on wind turbine noise annoyance," J. Acoust. Soc. Am. **139**, 1241–1251.

ISO/TS 15666 (**2003**). "Technical specification: Acoustics–Assessment of noise annoyance by means of social and socio-acoustic surveys" (International Organization for Standardization, Geneva, Switzerland).

Jang, M. J. (**2011**). "Working correlation selection in generalized estimating equations," Ph.D. dissertation, University of Iowa, Iowa City, IA, available online at http://ir.uiowa.edu/etd/2719 (Last viewed March 17, 2016).

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.

Jeon, J. Y., Lee, P. J., You, J., and Kang, J. (**2010**). "Perceptual assessment of quality of urban soundscapes with combined noise sources and water sounds," J. Acoust. Soc. Am. **127**, 1357–1366.

Johnson, P. C. D. (**2014**). "Extension of Nakagawa & Schielzeth's R^2_{GLMM} to random slopes models," Methods Ecol. Evol. **5**, 944–946.

Knopper, L. D., and Ollson, C. A. (**2011**). "Health effects and wind turbines: A review of the literature," Env. Heal. **10**, 78.

Kuwano, S., Namba, S., and Fastl, H. (**1988**). "On the judgment of loudness, noisiness and annoyance with actual and artificial noises," J. Sound Vib. **127**, 457–465.

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

Lee, S., Kim, K., Choi, W., and Lee, S. (**2011**). "Annoyance caused by amplitude modulation of wind turbine noise," Noise Control Engr. J. **59**, 38–46.

Legarth, S. V. (**2007**). "Auralization and assessments of annoyance from wind turbines," in *Second International Meeting on Wind Turbine Noise*, September 20–21, 2007, Lyon, France [available from Institute of Noise Control Engineering (INCE) Europe, Merseyside, United Kingdom].

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 *Proceedings of the 9th Congress of the International Commission on the Biological Effects of Noise (ICBEN)*, July 21–25, 2008, Foxwoods, CT, edited by B. Griefahn (available from Institut für Arbeisphyisologie an der Universität Dortmund, Dortmund, Germany), pp. 562–570.

Liang, K. Y., and Zeger, S. L. (**1986**). "Longitudinal data analysis using generalized linear models," Biometrika **73**, 13–22.

Makarewicz, R., and Gołębiewski, R. (**2015**). "Influence of blade pitch on amplitude modulation of wind turbine noise," Noise Control Engr. J. **63**, 195–201.

Manyoky, M., Wissen Hayek, U., Heutschi, K., Pieren, R., and Grêt-Regamey, A. (**2014**). "Developing a GIS-based visual-acoustic 3D simulation for wind farm assessment," ISPRS Int. J. Geo-Inf. **3**, 29–48.

McCunney, R. J., Mundt, K. A., Colby, W. D., Dobie, R., Kaliski, K., and Blais, M. (**2014**). "Wind turbines and health: A critical review of the scientific literature," J. Occup. Environ. Med. **56**, e108–e130.

McGraw, K. O., and Wong, S. P. (**1996**). "Forming inferences about some intraclass correlation coefficients," Psychol. Meth. **1**, 30–46.

Michaud, D. S., Feder, K., Keith, S. E., Voicescu, S. A., Marro, L., Than, J., Guay, M., Denning, A., McGuire, D., Bower, T., Lavigne, E., Murray, B. J., Weiss, S. K., and van den Berg, F. (**2016a**). "Exposure to wind turbine noise: Perceptual responses and reported health effects," J. Acoust. Soc. Am. **139**, 1443–1454.

Michaud, D. S., Keith, S. E., Feder, K., Voicescu, S. A., Marro, L., Than, J., Guay, M., Bower, T., Denning, A., Lavigne, E., Whelan, C., Janssen, S. A., Leroux, T., and van den Berg, F. (**2016b**). "Personal and situational variables associated with wind turbine noise annoyance," J. Acoust. Soc. Am. **139**, 1455–1466.

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

Miedema, H. M. E., and Vos, H. (**2003**). "Noise sensitivity and reactions to noise and other environmental conditions," J. Acoust. Soc. Am. **113**, 1492–1504.

Møller, H., and Pedersen, C. S. (**2011**). "Low-frequency noise from large wind turbines," J. Acoust. Soc. Am. **129**, 3727–3744.

Nakagawa, S., and Schielzeth, H. (**2013**). "A general and simple method for obtaining R^2 from generalized linear mixed-effects models," Methods Ecol. Evol. **4**, 133–142.

NAO (**1986**). Lärmschutz-Verordnung (LSV) vom 15. Dezember 1986 (Stand am Stand am 1. Januar 2016) [Noise Abatement Ordinance (NAO) of December 15 1986 (Status as of January 1, 2016)]. SR 814.41, available online at https://www.admin.ch/opc/de/classified-compilation/19860372/201601010000/814.41.pdf (Last viewed March 17, 2016).

Oerlemans, S. (**2015**). "Effect of wind shear on amplitude modulation of wind turbine noise," Int. J. Aeroacoust. **14**, 715–728.

Onakpoya, I. J., O'Sullivan, J., Thompson, M. J., and Heneghan, C. J. (**2015**). "The effect of wind turbine noise on sleep and quality of life: A systematic review and meta-analysis of observational studies," Environ. Int. **82**, 1–9.

Pawlaczyk-Łuszczyńska, M., Dudarewicz, A., Waszkowska, M., and Śliwińska-Kowalska, M. (**2003**). "Assessment of annoyance from low frequency and broadband noises," Int. J. Occup. Med. Environ. Health **16**, 337–343.

Pedersen, E., and Larsman, P. (**2008**). "The impact of visual factors on noise annoyance among people living in the vicinity of wind turbines," J. Environ. Psychol. **28**, 379–389.

Pedersen, E., and Persson Waye, K. (**2004**). "Perception and annoyance due to wind turbine noise—a dose-response relationship," J. Acoust. Soc. Am. **116**, 3460–3470.

Pedersen, E., van den Berg, F., Bakker, R., and Bouma, J. (**2009**). "Response to noise from modern wind farms in The Netherlands," J. Acoust. Soc. Am. **126**, 634–643.

Pedersen, E., van den Berg, F., Bakker, R., and Bouma, J. (**2010**). "Can road traffic mask sound from wind turbines? Response to wind turbine sound at different levels of road traffic sound," Energ. Policy **38**, 2520–2527.

Pedersen, T. H. (**2007**). The "Genlyd" Noise Annoyance Model, Dose-Response Relationships Modelled by Logistic Functions. Journal No. AV

1102/07 (DELTA Danish Electronics, Light & Acoustics, Hørsholm, Denmark), available online at http://share.madebydelta.com/wp-content/publications/akustik/paper_og_rapport/The_Genlyd_Noise_Annoyance_Model.pdf (Last viewed March 17, 2016).
Persson Waye, K., and Öhrström, E. (**2002**). "Psycho-acoustic characters of relevance for annoyance of wind turbine noise," J. Sound Vib. **250**, 65–73.
Pieren, R., Heutschi, K., Müller, M., Manyoky, M., and Eggenschwiler, K. (**2014**). "Auralization of wind turbine noise: Emission synthesis," Acta Acust. Acust. **100**, 25–33.
Pinheiro, J. C., and Bates, D. M. (**2000**). *Mixed-Effects Models in S and S-Plus*, 1st ed. (Springer-Verlag, New York, NY), 528 pp.
Poulsen, T. (**1991**). "Influence of session length on judged annoyance," J. Sound Vib. **145**, 217–224.
RenewableUK (**2013**). Wind Turbine Amplitude Modulation: Research to Improve Understanding as to its Cause and Effect (RenewableUK, London, United Kingdom), available online at http://www.renewableuk.com/en/publications/index.cfm/wind-turbine-amplitude-modulation (Last viewed March 17, 2016).
Schmidt, J. H., and Klokker, M. (**2014**). "Health effects related to wind turbine noise exposure: A systematic review," PLoS One **9**, e114183.
Schomer, P., and Fidell, S. (**2016**). "Introductory remarks for special issue on wind turbine noise," J. Acoust. Soc. Am. **139**, 1430.
Schultz, T. J. (**1978**). "Synthesis of social surveys on noise annoyance," J. Acoust. Soc. Am. **64**, 377–405.
Schütte, M., Marks, A., Wenning, E., and Griefahn, B. (**2007**). "The development of the noise sensitivity questionnaire," Noise Health **9**, 15–24.
Schwarz, G. (**1978**). "Estimating the dimension of a model," Ann. Stat. **6**, 461–464.
Seong, Y., Lee, S., Gwak, D. Y., Cho, Y., Hong, J., and Lee, S. (**2013**). "An experimental study on annoyance scale for assessment of wind turbine noise," J. Renew. Sust. Ener. **5**, 052008.
Tjur, T. (**2009**). "Coefficients of determination in logistic regression models—a new proposal: The coefficient of discrimination," Am. Stat. **63**, 366–372.
Torija, A. J., and Flindell, I. H. (**2015**). "The subjective effect of low frequency content in road traffic noise," J. Acoust. Soc. Am. **137**, 189–198.
UZH, and Empa (**1974**). Sozio-psychologische Fluglärmuntersuchung im Gebiet der drei Schweizer Flughäfen Zürich, Genf, Basel (Sociopsychological Investigation of Aircraft Noise in the Vicinity of the Three Swiss Airports Zurich, Geneva, Basel). (Arbeitsgemeinschaft für sozio-psychologische Fluglärmuntersuchungen [University of Zurich (UZH) and Swiss Federal Laboratories for Materials Science and Technology (Empa)], Berne, Switzerland).
van den Berg, F. (**2009**). "Why is wind turbine noise noisier than other noise?" in *Proceedings of Euronoise 2009, 8th European Conference on Noise Control*, Edinburgh, Scotland (October 26–28, 2009), edited by Institute of Acoustics [available from European Acoustical Society (EAA), Institute of Acoustics (UK), ISBN: 9781615676804], Paper No. 0087.
Van Gerven, P. W. M., Vos, H., Van Boxtel, M. P. J., Janssen, S. A., and Miedema, H. M. E. (**2009**). "Annoyance from environmental noise across the lifespan," J. Acoust. Soc. Am. **126**, 187–194.
Van Renterghem, T., Bockstael, A., De Weirt, V., and Botteldooren, D. (**2013**). "Annoyance, detection and recognition of wind turbine noise," Sci. Total Environ. **456–457**, 333–345.
Voicescu, S. A., Michaud, D. S., Feder, K., Marro, L., Than, J., Guay, M., Denning, A., Bower, T., van den Berg, F., Broner, N., and Lavigne, E. (**2016**). "Estimating annoyance to calculated wind turbine shadow flicker is improved when variables associated with wind turbine noise exposure are considered," J. Acoust. Soc. Am. **139**, 1480–1492.
Vonesh, E. F., Chinchilli, V. M., and Pu, K. (**1996**). "Goodness-of-fit in generalized nonlinear mixed-effects models," Biometrics **52**, 572–587.
Yokoshima, S., Yano, T., Kawai, K., Morinaga, M., and Ota, A. (**2012**). "Representative dose-response curves for individual transportation noises in Japan," in *Proceedings of the Inter-Noise 2012, 41st International Congress and Exposition on Noise Control Engineering*, New York, NY (August 19–22, 2012), edited by C. Burroughs and S. Conlon (available from Institute of Noise Control Engineering, Indianapolis, IN), Paper No. 796.