The role of task interference and exposure duration in judging noise annoyance

Zimmer, Karin; Ghani, Jody Ann; Ellermcier, Wolfgang

Published in:
Journal of Sound and Vibration

DOI (link to publication from Publisher):
10.1016/j.jsv.2007.10.002

Publication date:
2008

Document Version
Publisher's PDF, also known as Version of record

Available online at www.sciencedirect.com
ScienceDirect
Journal of Sound and Vibration 311 (2008) 1039–1051
JOURNAL OF SOUND AND VIBRATION
www.elsevier.com/locate/jsvi

The role of task interference and exposure duration in judging noise annoyance ☆

Karin Zimmer[a,1], Jody Ghani[a,*], Wolfgang Ellermeier[b]

[a] *Sound Quality Research Unit (SQRU), Department of Acoustics, Aalborg University, Fredrik Bajers Vej 7 B5, DK-9220 Aalborg Ø, Denmark*
[b] *Institut für Psychologie, Technische Universität Darmstadt, Germany*

Received 27 April 2006; received in revised form 26 September 2007; accepted 2 October 2007

Abstract

To determine whether the amount of performance disruption by a noise has an effect on the annoyance that noise evokes, a laboratory situation was created in which the participants rated a number of sounds before, after, and while performing a cognitively demanding memory task. The task consisted of memorizing, and later reproducing, a visually presented sequence of digits while being exposed to irrelevant sound chosen to produce different degrees of disruption. In two experiments, participants assessed these background sounds (frequency-modulated tones, broadband noise and speech) on a rating scale consisting of thirteen categories ranging from 'not annoying at all' to 'extremely annoying.' The judgments were collected immediately before, after, and concomitant to, the memory task. The results of the first experiment ($N = 24$) showed that the annoyance assessments were indeed altered by the experience of disruption, most strongly during, and to a lesser extent after task completion, whereas ratings of the non-disruptive sounds remained largely unaffected. In the second experiment ($N = 25$), participants were exposed to the same sounds, but for longer intervals at a time: 10 min as opposed to 14 s in the first experiment. The longer exposure resulted in increased annoyance in all noise conditions, but did not alter the differential effect of disruption on annoyance, which was replicated. The results of these laboratory experiments support the notion that annoyance cannot be conceived of as a purely perceptual sound property; rather, it is influenced by the degree of interference with the task at hand.

1. Introduction

It is common knowledge that noise may adversely affect one's well-being. A prevalent non-auditory effect of noise is annoyance: According to a recent report by the World Health Organization, more than 50% of the citizens in the European Union live in an environment that is prone to give acoustical discomfort [1].

☆ Portions of this work were presented at the 33rd International Congress and Exposition on Noise Control Engineering, Internoise 2004, Prague, Czech Republic, August 2004.

*Corresponding author. Tel.: +45 61672042.
E-mail address: jody.ghani@gmail.com (J. Ghani).
URL: http://www.soundquality.dk (J. Ghani).
[1] Currently at OECD, Directorate for Education, Paris, France.

0022-460X/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jsv.2007.10.002

Noise annoyance is defined as an evaluative response towards the sound and its source, including both emotional ('nuisance', 'unpleasantness'), and cognitive ('disturbance', 'interference') aspects (for an overview on current conceptualizations, see Ref. [2]), and is commonly assessed using rating scales. The degree to which a person is annoyed by a sound is to some extent related to acoustical properties of the sound: Level and frequency content, as well as their variation over time, have been identified as major contributors to annoyance [3,4].

Conceptualizing annoyance as an evaluative response largely determined by the physical characteristics of the auditory input naturally leads to attempts to predict annoyance judgments from a combination of acoustical, or psycho-acoustical metrics, the best-known being Zwicker's *unbiased annoyance* [5]. This approach has been criticized from the point of view of psychological acoustics, arguing that the situation in which a sound is perceived should also be considered. While overhearing a conversation between neighbours may not raise much feeling when preparing dinner, for example, the same sounds may severely annoy the listener when on the verge of sleep. Therefore, some researchers have asserted that annoyance can only be assessed in relation to a (primary) activity which the sound, potentially, or factually, interferes with (e.g. Refs. [6–9]). While the *unpleasantness* of sound may thus be a purely perceptual matter (to be judged entirely on the basis of the auditory input), these authors postulate that in order to speak of (auditory) *annoyance*, the interference aspect is essential.

1.1. Task interference and annoyance

It appears that the literature on noise annoyance on the one hand, and that on noise effects on the other, are so disparate, that only a handful of laboratory studies [6,7,9–12] have looked at objective impairment and subjective annoyance simultaneously. Two of these, that are most relevant to the present study, shall be briefly summarized to illustrate the issues involved.

Moran and Loeb [6] had 120 participants perform a listening task (correcting a written text version by comparing it to its headphone-delivered counterpart) and a visual-attention task (checking for target letters) while being exposed to high-level aircraft noise, or working in quiet. They obtained large performance decrements from approximately 80% correct to 30% correct in the listening task, but no changes due to noise in the visual-attention task. Retrospective magnitude estimates of the annoyance produced by the noise indicated it to be significantly greater during the task (listening) with which the noise actually interfered. The authors conclude that 'annoyance may be underestimated if scaled in a situation not involving interference' (p. 725). Note that, while this study clearly demonstrates a link between annoyance and interference, the effects observed are likely to be due to pure acoustical masking, and therefore not indicative of the cognitive interference effects encountered at moderate noise levels.

These were the focus of a recent study by Landström et al. [10] who investigated proofreading, a logical-reasoning task, and physical letter sorting while their participants were exposed to broadband noise or background speech of moderate levels. Surprisingly though, this study showed a performance decrement only in the letter sorting task (under speech), and no systematic relationship between actual performance and annoyance.

The latter study concurs with others aimed at investigating the relationship between annoyance and interference in finding no, or only very weak, performance effects due to the task and noise conditions employed [13,7,11,12]. For the goal of demonstrating *causal* effects of task interference on perceived annoyance, however, it is crucial to demonstrate factual performance decrements in the first place.

1.2. Irrelevant sound paradigm

The lack of a link between task interference and annoyance evident in the few pertinent laboratory studies may simply be due to the fact that the operationalization of interference was not optimal. Note, that the mere presence of background noise, even if of high level and unpleasant character, does not necessarily produce interference (see Ref. [14]). It requires both a potent disruptor (typically: temporally changing, speech, or music-like sound), and a sensitive task (typically one requiring some mental complexity; ideally, some serial storage) in order for interference to reliably occur.

Both requirements are met by the *irrelevant sound paradigm*, in which 'changing-state', and to-be-ignored background sound reliably interferes with a serial-recall task (i.e. memorizing a series of digits or letters in the correct order). The irrelevant sound effect (ISE) has been extensively studied in cognitive psychology (for reviews see Ref. [15]), and the acoustical parameters that give rise to interference in this paradigm are well established [16].

Therefore, the present study was designed to employ the proven ISE paradigm to produce various degrees of task interference, and to study their effect on annoyance judgments. The particular twist of the present investigation is to monitor annoyance at various points in time with respect to the task to be done: (a) before task performance, i.e. in its (potentially) 'unbiased' form, (b) during performance of a difficult memory task, and (c) after having completed the task, in order to check whether experience with the disruptive effect of a sound is enough to change its annoyance score.

Clearly, however, such effects may be a function of exposure duration, the research on which is briefly summarized in the next section.

1.3. Effect of exposure duration

Research on the relationship between the length of time listeners are exposed to noise and its annoyance has produced very contradictory results. Using white noise and a random mix of durations and levels, Hiramatsu et al. [17] found estimated annoyance to increase linearly with the logarithm of duration over a range from 30 ms to 90 s, with no tendency to asymptote, and thereby far exceeding what is known as the 'critical duration' for the temporal integration of loudness (ca. 250 ms). Using much longer durations, Poulsen [18] reported annoyance ratings of recorded traffic and synthetic gunfire noise to remain the same for durations of 1, 5, 15, and 30 min.

By way of interpretation, Poulsen [18] attributes his counterintuitive result to his instructing listeners to rate how annoying the sound may be in a situation at home, rather than to form a judgment based on the stimulus per se. Thus, he argues, participants may well have considered the amount of *disturbance* that they suspected the sound to produce in their activities, while disregarding the accumulation of emotional discomfort over time. To explicitly address this issue, via experimental manipulation of the amount of disruption produced, the present study was designed to investigate how exposure duration and the degree of disruption might interact in generating annoyance by noise.

1.4. Goals of the present investigation

The present study thus addresses two issues:

(1) The effect of task disruption on annoyance (Experiment I).
This shall be done by using an (irrelevant speech) paradigm producing highly reliable cognitive disruption effects. Furthermore, to observe how the experience of disruption might modify annoyance *within a given listener*, annoyance ratings shall be collected before, during and after working on the primary task the noise interferes with.

(2) The effect of exposure duration on the annoyance of noise (Experiment II).
To that effect, by blocking noise conditions, much longer exposures (of up to 10 min) to the *same* background noise will be created. The results are compared to those of Experiment I, in which the sound conditions vary from trial to trial, and last a mere 14 s, thus preventing a continuous build-up of annoyance to a specific auditory stimulation condition.

Taken together, these experiments may contribute to clarifying the role of interference in annoyance, and its interaction with exposure duration.

2. Experiment I: effect of task disruption on annoyance

2.1. Method

2.1.1. Participants

Twenty-four subjects (16 male) participated in the experiment, the majority of whom were students at the University of Aalborg, Denmark. Their median age was 24 years (range: 21–30 years). Subjects were screened for known hearing problems and were paid for their participation. Subjects were naïve to the goals of the experiment.

2.1.2. Apparatus and stimuli

2.1.2.1. Visual stimuli. The visual stimuli to be memorised on each trial were a random permutation of the digits 1–9, displayed sequentially in the centre of a computer monitor. The digits were approximately 3 cm tall, and each digit was displayed for 1 s, with no pause between digits. A red circle was displayed for 1 s before each trial as a warning signal.

2.1.2.2. Auditory stimuli. In addition to silence, serving as a baseline condition, four auditory stimuli were used as a background in the memory task: A segment of a Korean poem recited by a male native speaker, white noise, and two frequency modulated (FM) tones, with a 1150-Hz centre frequency, and a modulation depth of 850 Hz. The speech was recorded in anechoic conditions and edited to give a natural-sounding segment of appropriate length, without excess pauses. One FM tone was continuous, and had a modulation rate of 8 Hz (referred to as 8 Hz C in the following), while the other was modulated at a rate of 0.25 Hz, and was interrupted, i.e. contained 300-ms gaps of silence, which alternated with 400 ms 'on' periods to avoid a periodic pattern emerging (0.25 Hz I). All onsets and offsets were smoothed using 10-ms rise and fall times.

In the pre- and post-task annoyance assessments, the same four stimuli were used as well as five additional, distractor stimuli which were included in order to minimize the possibility of subjects remembering their ratings from previous repetitions. It was important that these did not affect subjects' scale-usage. For this reason, distractor stimuli were chosen with similar properties to the existing set, and, from the results of a pilot study, likely to fall within the range of annoyance as specified by the existing set: Two further FM tones, one interrupted tone at 0.5-Hz modulation frequency (0.5 Hz I), and one continuous, 4-Hz modulated tone (4 Hz C). In addition, 10-dB attenuated versions of both the noise and the speech were included, as well as a distorted version of the speech at the same rms level. To add distortion to the speech segment, it was requantized to give a signal to noise ratio of 35 dB.

The stimuli had a total duration of 14 s, and were equated in level to yield rms levels of 63 dB SPL. All stimuli were D/A-converted by a 16-bit soundcard (RME Digi96/8 PST) at a sampling rate of 44.1 kHz, amplified (Rotel RB-976 MkII) and presented diotically to the subjects, sitting in a double-walled sound-attenuated booth, via headphones (Beyerdynamic DT-990).

2.1.3. Procedure

In the experiment, data were collected both on the participants' performance in the serial-recall task, and on the ratings of annoyance at various points in time.

2.1.3.1. Memory task. The memory task was based on the irrelevant sound paradigm. In each trial, a series of nine different digits was displayed for 1 s each. After a 5-s retention interval, the subject's task was to reproduce the digits in the order in which they had been presented. Recall was recorded by means of an on-screen numerical keypad, arranged in a 3×3 array. Once all digits had been entered, the next trial started after a 1 s pause.

Digits were presented, and rehearsed, in one of 5 auditory conditions (quiet, noise, speech, continuous, or intermittent FM tone). The auditory background on each trial was chosen in a pseudo-random fashion whereby every cycle of 5 trials contained a different, random permutation of the 5 auditory conditions. No sound was presented during the recall period, which was self-paced. Subjects were instructed to ignore the background sound and to concentrate on memorizing the number sequence. In each session, 15 repetitions per auditory condition (75 trials) were completed by the subjects, with the trials split into three blocks of 25 trials, lasting approximately 12 min each.

For every participant, data collection for the memory task spanned two 1-h sessions on separate days, one in which serial recall performance was tested exclusively, and one in which annoyance judgments (as described in the next paragraph) were interspersed. The order in which the two types of sessions were run was balanced across subjects (see conditions A and B in Fig. 1). At the start of each day, four practice trials on the memory task were completed prior to the data collection proper.

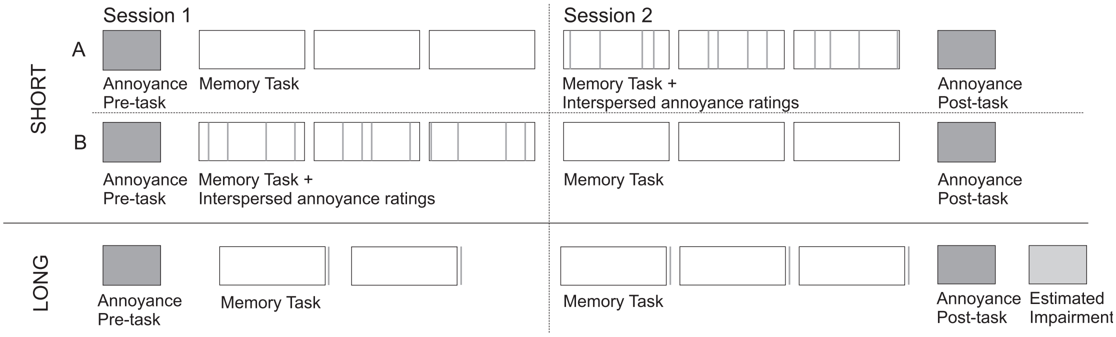

Fig. 1. Experimental design. Experiment I (short duration, $n = 24$) is shown in the upper panel. A and B indicate the groups of participants receiving different orders of sessions ($n = 12$ in each group). Experiment II (long duration, $n = 25$) is shown in the lower panel. All subjects completed a pre- and post-task assessment of annoyance. Vertical grey lines represent interspersed annoyance judgements.

2.1.3.2. Annoyance ratings. Ratings of annoyance were obtained using an expanded 5-category rating scale. Following the recommendations in Refs. [19,20], the five categories were labelled: 'extremely', 'very', 'moderately', 'slightly' and 'not at all' annoying. For the first four categories, a finer grain was added to give three grades of each category; for 'not at all', only one grade was used. The scale was accompanied by the question ''How annoying did you find the background sound?'' and was displayed, with a button for each category, on the computer screen. To make instantly obvious the orientation of the scale, the buttons were shaded—the highest grade of 'Extremely' was black, 'Not at all' white, and the remaining buttons intermediate, graded shades of grey.

Annoyance ratings were collected at three points in time: before task exposure, while performing the serial-recall task, and upon its completion after two sessions. On each of these occasions, three annoyance assessments were obtained for each of the four auditory stimuli, as well as for the five distractor stimuli used in the pre- and post-task conditions only. In these pre- and post-task conditions, the sounds were presented in a pseudo-random order whereby every cycle of nine trials contained a different, random permutation of all nine stimuli, and the rating scale and question were presented immediately after the sound had finished.

In one of the two sessions—the second for group A, the first for group B (see Fig. 1)—annoyance assessments were interspersed at approximately every fifth trial of the memory task in which a background sound had been presented (concomitant condition). In each of the three blocks of 25 trials, annoyance judgements were collected once on each of the four auditory stimuli, at random intervals, and immediately after the subject had finished reproducing the digit sequence. Subjects were not informed about the occurrence of these interspersed assessments beforehand.

2.2. Results and discussion

To analyse performance in the memory task, an error was scored whenever the participant failed to report the correct number in the correct position. The mean number of errors over 30 trials for a given auditory condition (ranging from 0 to 9) served as the dependent variable. To analyse the annoyance ratings, means were taken across the three repetitions in each condition.

2.2.1. Performance

Mean errors per trial ranged between 2.3 (for quiet, and noise) and 3.5 (for speech). Fig. 2 (lower panel) shows these means for all auditory conditions along with the corresponding standard errors.

An analysis of variance of the error rate yielded a highly significant effect of sound on performance: $F(4, 92) = 28.24$, $p < 0.001$. Post hoc analyses [21] showed that significantly more errors were made in speech than when the intermittent, 0.25-Hz FM tone was played. Both of these sounds led to significantly higher error rates than either the continuous 8-Hz FM tone, noise, or quiet. With a 52% increase in error rate for speech as compared to silence, the effect size is comparable to the results reported in Ref. [22] on a larger subject sample, and somewhat larger than the one obtained when auditory conditions are presented in a blocked fashion (e.g. Ref. [23]). When looking at the number of errors in speech vs. quiet for subjects individually, it turned out that for 22 of the 24 participants, an irrelevant speech effect was factually found. The effects of the FM tones are also consistent with what has been found in the literature, in that only interrupted FM produces disruption [24]. It may therefore be concluded that the background sounds used in this experiment induced a satisfactory variation in the degree of task disruption, from not disrupting at all (noise, continuous 8-Hz FM tone), to mildly disrupting (the intermittent 0.25-Hz FM tone, yielding a performance decrement of 17.4%), to clearly disrupting (speech, effect size: 52.2%).

As typically obtained [21], the number of errors decreased from the first to the second session, $F(1,23)=28.24$; $p<0.001$, but there was no interaction of session number with the sound conditions, $F(4,92)=0.62$; $p=0.648$, indicating that despite an overall improvement, the pattern of noise effects remained the same. More importantly, the two types of sessions (with or without annoyance ratings, s. Fig. 1, two top rows) were statistically indistinguishable, meaning that the concomitant assessment of annoyance did not make a difference with respect to the number of errors produced, ($F(1,23)=0.74$; $p=0.4$).

2.2.2. Annoyance

Mean annoyance ratings obtained from the 24 listeners before, during and after the memory task are given in Fig. 2, upper panel, together with the respective standard errors. It can be seen that the ratings spanned a much larger range both in the pre-task, and the post-task assessments, than in the judgments obtained while performing the memory task. Most strikingly, with a mean rating of 3.44, the speech sample was considered as ‘slightly’ annoying at the outset, while it was judged to be ‘very’ annoying during the memory task (mean rating: 8.24). After task completion, annoyance dropped to an intermediate value of 5.07.

An analysis of variance with the factors *sound* (4 stimuli) and *rating condition* (before, during and after the memory task) yielded highly significant effects of *sound*, $F(3,69)=27.77$; $p<0.001$, and of the interaction,

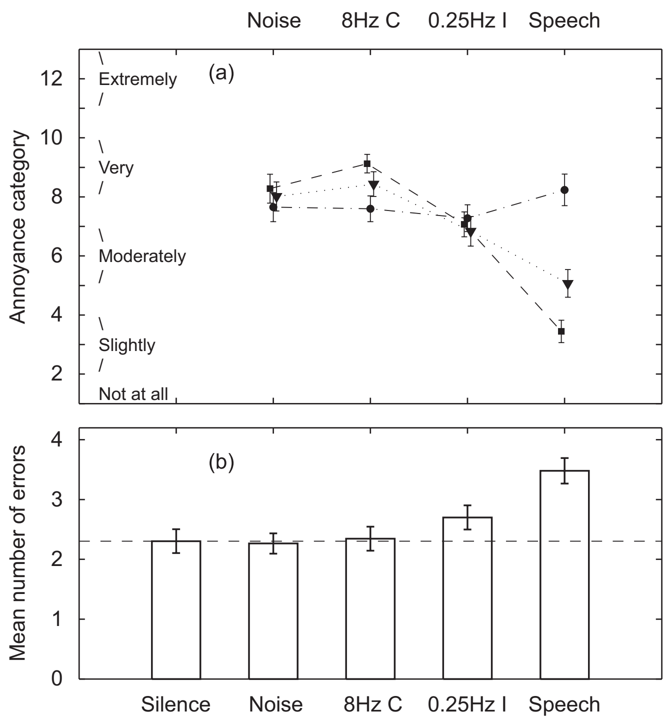

Fig. 2. Mean annoyance ratings (a) in Experiment I with standard errors: ■ – – ■ pre-task; ● · · ● concomitant; ▼ – · – ▼ post-task, including annoyance category labels. In (b), mean number of errors are given for all auditory conditions, together with their standard errors.

rating $\times$ *sound*, $F(6,138) = 31.73$; $p < 0.001$. The main effect of sound implies that, ignoring the differences between rating conditions, speech was the least annoying sound, while the continuous 8-Hz FM tone was the most annoying, with assessments for the other two sounds falling in between. The *rating* $\times$ *sound* interaction implies that the pattern of annoyance ratings is significantly different depending on the point in time the rating is obtained. Post hoc analyses [21] of this interaction effect showed that assessments did not change across rating conditions for the intermittent 0.25-Hz FM tone or for noise. Annoyance ratings were, however, significantly higher before than during the memory task for the continuous, 8-Hz FM tone. For the speech sample, this pattern was reversed, in that annoyance was significantly lower prior to the task than thereafter, which was in turn lower than the rating during serial recall.

It thus appears that both in their concomitant, and in their post-task annoyance ratings, the listeners took into account the strong disruption produced by the speech sound, while no variation in annoyance judgments was found for the mildly disruptive intermittent FM tone. As the ratings were based on rather short sound samples (14 s), however, it is conceivable that a longer continuous exposure duration would further sensitize the listeners towards the interference caused by the sounds, and that factual task disruption would then play a larger role in annoyance assessments. The second experiment addresses this question.

3. Experiment II: exposure duration

3.1. Method

3.1.1. Participants

Twenty-five paid subjects (15 male), none of whom had participated in Experiment I, participated in the experiment. The majority were students at the University of Aalborg, Denmark, and their median age was 24 years (range: 21–31 years). All subjects had a hearing threshold no higher than 20 dB above the normal reference according to Ref. [25], and no knowledge of the Korean language.

3.1.2. Apparatus and stimuli

3.1.2.1. Auditory stimuli. The auditory stimuli used as a background in the memory task were longer versions of the same white noise and frequency-modulated tones (8 Hz C and 0.25 Hz I) as in Experiment I, and a new recording of the same male speaker. The apparatus and overall rms levels were the same as in Experiment I. The sounds had a duration of approximately 10 min.

To enable direct comparison between experiments, the pre- and post-task annoyance assessments of this experiment were kept as similar as possible to those of Experiment I. The stimuli presented for annoyance ratings (including the 5 distractor sounds) were therefore identical, with the exception of the speech, which was instead a 14-s section of the newly-recorded 10-min speech, chosen to have similar speaking rate and pauses as the original 14-s sample from Experiment I.

3.1.3. Procedure

Data were collected on the participants' performance in the serial-recall task, as well as their reaction time from the moment the keypad is shown to the first button press, and the duration of the self-paced recall section. As in experiment I subjects also rated annoyance at various points in time. At the end of the experiment, data was collected on how the subjects estimated they were impaired by the 4 background sounds. The experimental flow is shown in the lower panel of Fig. 1.

3.1.3.1. Memory task. In Experiment II the implementation of the memory task remained identical to that in Experiment I. However, the organization of the sound conditions was modified: Trials were arranged in blocks of approximately 10 min duration during which one of the five auditory conditions served as the irrelevant sound, and was played continuously throughout the whole block. A block was terminated when the total elapsed time at the end of a trial was greater than 585 s (10 min–15 s), resulting in an average block duration of 599 s (range 585–630 s). The number of trials in each block also varied (from 14 to 27 trials, mean = 23). This uncertainty arose from the self-paced nature of the paradigm and this implementation was chosen to ensure foremost that the auditory exposure was kept as constant as possible for all subjects while keeping the task itself identical to that in Experiment I. The order in which the auditory stimuli were presented to the 25 subjects was balanced using a 5×5 latin square repeated 5 times. Each subject completed one 10-min block per auditory condition, with two blocks on the first day, and three on the second.

3.1.3.2. Annoyance ratings. Pre- and post-task ratings of annoyance were collected as in Experiment I. The concomitant assessment of annoyance, however, was made directly after each 10-min block of the memory task (not including the silence condition), and hence only one assessment per stimulus was collected.

3.1.3.3. Ratings of estimated impairment. At the end of the second session, after completing the post-task annoyance assessment, a further task was given to the subjects: estimating the impairment to their performance in the memory task caused by each of the four auditory conditions (silence was not rated). Five-second samples of the stimuli were played, after which rating was done by means of a seven-point bipolar scale taken from Ref. [22]. The scale ranged from −3, ‘‘It severely hurt my performance’’ through 0, ‘‘No effect’’ to 3, ‘‘It helped considerably’’. The accompanying question was ‘‘How much did this sound affect your performance?’’. It should be noted that the instruction to this part was given only when the subject had completed the rest of the experiment.

3.2. Results and discussion

3.2.1. Performance: effect of exposure duration

In the second experiment, the memory task was terminated after the listener had been exposed to a given sound for approximately 10 min. Like in the first experiment, the task was self-paced. Therefore, the number of trials with a given background sound was not fixed—varying between 22.6 for speech, and 23.4 for noise, on average,—but it did not vary across sound conditions in a statistically significant fashion, $F(4, 96) = 1.96$; $p = 0.106$.

Mean errors per trial ranged between 2.3 for the continuous 8 Hz FM tone, and 3.7 for speech. Task disruption was thus comparable in magnitude to that obtained in the first experiment (see Section 2.2.1). A two-factorial analysis of variance with factors *sound* (silence, intermittent 0.25 Hz FM tone, continuous 8 Hz FM tone, noise, speech) and *exposure duration* (14 s in Experiment I vs. 10 min in Experiment II) ascertained that the error rates did not differ across experiments. There was no main effect of exposure duration: $F(1, 47) = 0.03$; $p = 0.858$; and no *sound* × *exposure* interaction, $F(4, 188) = 1.64$; $p = 0.166$, indicating that the pattern of errors was quite similar to what is seen in the bottom portion of Fig. 2.

3.2.2. Annoyance: effect of exposure duration

Fig. 3 illustrates the mean annoyance ratings collected in the second experiment, and the respective standard errors. As in the first experiment, ratings for the speech sound differed most across the rating conditions. While speech was judged to be the least annoying sound before and after completing the memory task (mean ratings: 4.4 and 6.6, respectively, corresponding to judgments of ‘‘slightly annoying’’ and ‘‘moderately annoying’’), it was given the highest annoyance rating, namely 11.1 on average (‘‘extremely annoying’’), in the concomitant assessment.

The influence of exposure duration on annoyance was investigated by way of a three-factor mixed analysis of variance with the factors *sound* (4 stimuli), *exposure duration* (14 s vs. 10 min), and *rating condition* (before, during, and after the memory task). As the pattern of disruption did not change across experiments, any differences found in concomitant and post-task annoyance judgments may be attributed to the variation in exposure time, as long as the pre-task assessments of annoyance in the two experiments are the same.

The analysis of variance yielded statistically significant main effects for all three factors (at $\alpha = 0.05$). Moreover, the interactions of *sound* × *rating condition* ($F(6, 282) = 58.13$; $p < 0.001$), and of *rating condition* × *exposure duration* ($F(2, 94) = 4.93$; $p = 0.009$), turned out to be significant.

Concomitant annoyance ratings were higher than either pre-, or post-task ratings, $F(2, 94) = 19.53$, $p < 0.001$. Further analysis of the *sound* × *rating* interaction showed that the point in time listeners were queried mattered only for speech ($F(2, 96) = 99.10$; $p < 0.001$), while the ratings did not change significantly before, during, and after the memory task for the other sounds. Furthermore, annoyance was generally higher

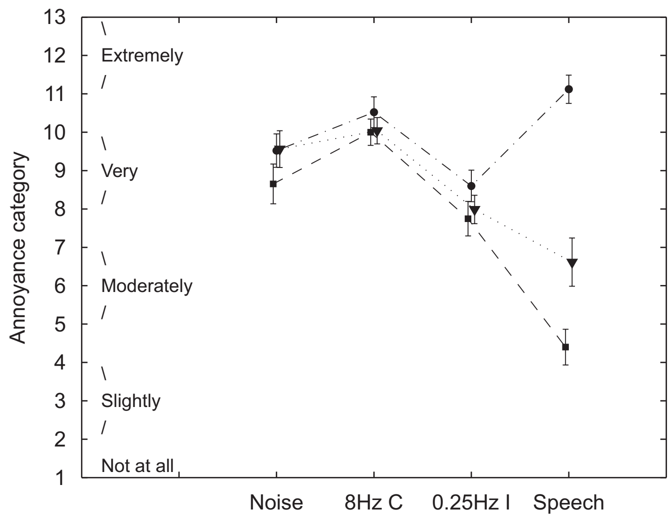

Fig. 3. Mean annoyance ratings in Experiment II with standard errors. ■ – – ■ pre-task; ● · · ● concomitant; ▼ – · – ▼ post-task, including annoyance category labels.

in the second (10 min. continuous exposure) than in the first (14 s exposure) experiment, $F(1,47) = 14.41$; $p<0.001$. Post hoc analysis of the *rating condition* × *exposure duration* interaction revealed that this result holds both for the concomitant, and the post-task annoyance assessments: $F(1,47) = 19.46$, $p<0.001$, and $F(1,47) = 8.94$, $p = 0.004$, respectively. Pre-task ratings, however, did not differ significantly in the two experiments ($F(1,47) = 2.70$, $p = 0.107$). It may thus be concluded that annoyance increased with exposure duration. As the three-way interaction (*sound* × *exposure time* × *rating condition*) turned out to be insignificant ($F(6,282) = 0.99$; $p = 0.429$), this conclusion holds for all sounds.

Fig. 4 illustrates the outcome: In the top panel, pre-task annoyance assessments of the sounds, as obtained in the two experiments, are given. It is evident, that the listeners in the second experiment gave marginally (but not significantly) higher ratings than the participants in the first experiment, when being presented with the (short) sound samples, for annoyance ratings. Subsequent annoyance assessments collected while working on the memory task, however, differed considerably across experiments, that is with exposure duration, for all sounds (Fig. 4, middle panel). Annoyance ratings were significantly lower, when the sound to be judged had been playing in the background for 14 s at a time, compared to 10 min continuously. After working on the memory task for a total of 2 h, post-task annoyance ratings were also higher in the second, compared to the first experiment. As is depicted in the bottom panel of Fig. 4, the difference in assessments somewhat decreases when compared to that obtained from the concomitant ratings.

3.2.3. Subjective impairment

At the end of the second experimental session, after giving their post-task assessments of annoyance, participants were asked to rate to which degree each of the background sounds affected their performance, on a seven-point scale ranging from "helped considerably" (+3) to "severely hurt" (−3). In result, average ratings ranged from −0.84 for the intermittent 0.25 Hz FM tone to −1.96 for speech. A repeated-measures analysis of variance revealed that the ratings differed across sounds: $F(3,72) = 4.26$; $p = 0.008$. From post hoc testing it turned out that speech was judged as more impairing than the intermittent FM tone, with ratings for the other sounds lying in between.

Fig. 5 illustrates the pattern of results with respect to estimated impairment, as well as concomitant, and post-task, annoyance.

The rating pattern for estimated impairment seems to match the pattern for concomitant annoyance quite well. On an individual level, the average correlation between estimated impairment and concomitant annoyance is $\rho = -0.415$ (based on the Fisher-Z transform of each individual's correlation over the four

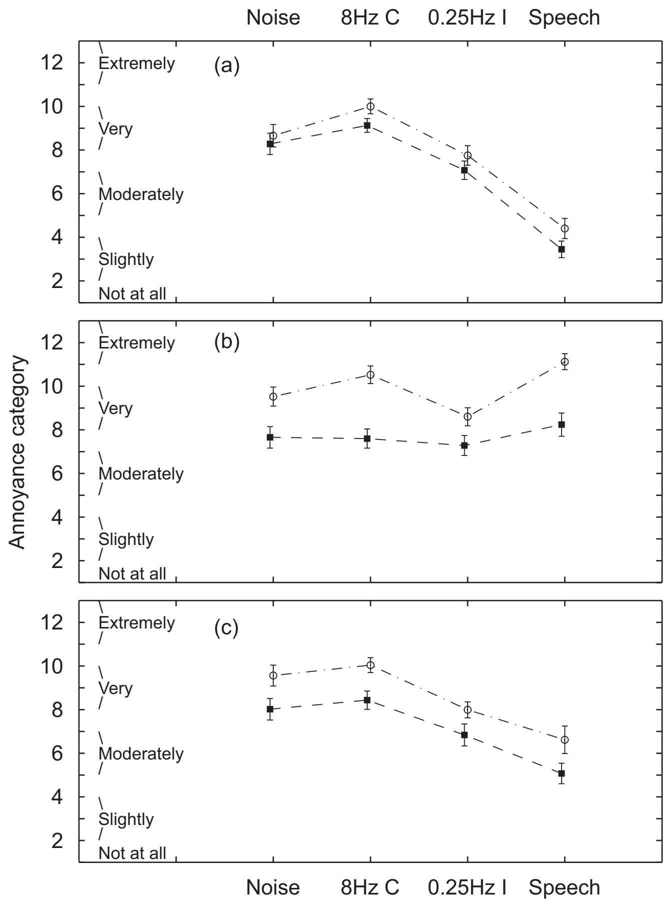

Fig. 4. Mean annoyance ratings in both experiments with standard errors: (a) pre-task, (b) concomitant, and (c) post-task assessments, including annoyance category labels. ■ –– ■ short duration; ○ · ·○ long duration.

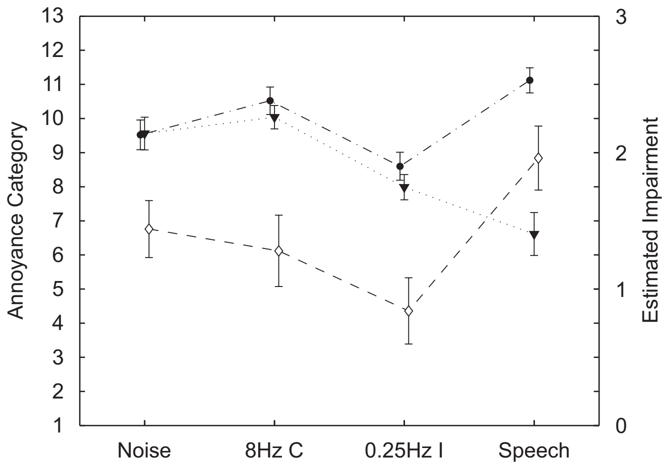

Fig. 5. Ratings of ◇–◇ estimated impairment; ● · ·● concomitant annoyance and ▼ – · – ▼ post-task annoyance.

sounds). When the estimated impairment is related to factual task disruption, i.e. the number of errors produced, the average rank-order correlation is $\rho = -0.421$, and thus in the same range as reported previously ($|r| = 0.44$; [22]). It thus appears that subjective impairment relates to objective disruption and concomitant annoyance to the same degree. By contrast, the association of estimated impairment with post-task annoyance is clearly lower, $\rho = -0.249$, and not significantly different from nil.

4. General discussion

4.1. Summary

The present laboratory study is the first to show in an experimental, within-subjects design, in what way task disruption affects annoyance. In previous investigations (e.g. Refs. [13,7,11,12]), task performance was either not disrupted by the sounds chosen, or, if a sound interfered with the task, no relationship between factual impairment and annoyance could be established. By contrast, different degrees of factual task disruption were reliably induced in the two experiments presented here, and annoyance ratings were obtained before, during and after working on a memory task. Both experiments show that factual task disruption does indeed alter noise annoyance ratings of sounds presented at moderate sound pressure levels: For the highly disruptive speech sound, annoyance ratings collected while working on the task are higher than ratings collected either before or after. Annoyance judgments for the mildly disruptive, or the completely ineffective sounds, on the other hand, generally do not change (an exception being the annoyance ratings—pre-task and concomitant—for the 8 Hz FM tone in the first experiment).

Note that the experimental nature of the study (by virtue of being able to introduce or remove the sound-sensitive task at will) helps to disambiguate the causal dependencies. It may be ruled out that the sound properties simultaneously affect annoyance and performance, or that annoyance is causal for the performance differences observed: In both cases, one would not expect a difference in annoyance between pre-task and concurrent assessments. Furthermore, other work has shown (e.g. Ref. [26]) that factors that greatly affect annoyance (such as overall loudness) have little or no effect on task disruption in the irrelevant speech paradigm. Rather, task interference as created in the present experiment appears to modify the annoyance response to the sounds over and above the operation of known acoustic factors (such as loudness, spectral content, and modulation rate).

Furthermore, exposure duration affected the annoyance judgments of *all* sounds. Experiment II demonstrated that the annoyance produced by a sound increased when the continuous exposure time rose from 14 s to 10 min, i.e. in the time frame bridging the very short exposure durations used in Ref. [17] and the longer durations in Ref. [18]. From the non-significant three-way interaction (*sound* × *rating condition* × *exposure duration*), however, it may be concluded that the influence of task disruption on annoyance did not vary with exposure duration. It thus appears that exposure duration increases (concomitant) annoyance directly, and not indirectly, such as by means of factual task disruption becoming more evident.

4.2. The changing concept of annoyance

The annoyance pattern for the highly disruptive speech sound was the same in both experiments in that the post-task assessments lay between the low pre-task ratings and the high concomitant ratings. Thus, after finishing the task, the effect of performance disruption on the annoyance judgment diminished rapidly (while still leaving a residue effect). From the impairment estimates obtained at the very end of Experiment II, it becomes clear that respondents were still able to recollect the disruption they experienced, but they were not basing their post-task annoyance assessment on it to the same degree as before, while involved in the cognitive task. One might speculate whether the post-task assessments would return to the pre-task level after some time has elapsed between performing the task in noise and assessing noise annoyance, and, if so, how long this would take.

Moreover, it seems as if acoustical properties strongly contribute to ratings of noise annoyance whenever a sound is judged in isolation, i.e. without reference to a context or scenario [5]. This holds (to some extent) even if the sound has disrupted performance immediately prior to the assessment. During the memory task, the criteria used to judge noise annoyance are different, however: in that situation, task disruption actually seems to play a major role. Taken together, these results argue against the notion of *one*, strictly context-independent, concept of annoyance. Researchers should be aware of the different meanings the term 'annoyance' may adopt, and should consider asking specifically for that aspect of annoyance which is in the focus of their interest—either the disruptive potential of the sound, or the annoyance experienced when disruption is disregarded. The present investigation shows paradigmatically that both aspects of annoyance can be addressed, and quantified.

4.3. Field assessments of annoyance

Assessments of noise annoyance in the field take a different approach. The standardized formulation in survey questionnaires is "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?" [20], thus trying to encompass all possible aspects of this negative emotional reaction. Given that the intuitive annoyance reaction is highly context-specific, and based on potentially quickly-shifting criteria of appraisal, this all-encompassing assessment seems unnatural, and hard to interpret. Depending on how much weight the respondents give to concurrent or recently experienced task disruption, their answers might rely on different concepts of annoyance, or rely on those concepts to differing degrees. One cannot be sure a priori, however, that the amount of disturbance a sound produces, will, in any case, enter into the annoyance assessment obtained at a later point in time.

4.4. Informational content of speech

There is an on-going discussion in the literature on whether or not speech plays a special role in the ISE [27], yet in any case speech is known to be a reliable disruptor of serial-recall tasks. In the present study, not only did speech produce this factual disruption, but in addition annoyance ratings unique to speech were observed, with listeners changing their rating dramatically depending on the context. In Experiment II, the somewhat less effective non-speech disruptor (the 0.25 Hz, interrupted FM tone) did not produce similar changes in rated annoyance, rather it was rated just as annoying before and after the task as during it. One explanation for the comparitively large effect of speech could be that in the concomitant condition the informational content of the speech is disregarded and it is perceived purely as an acoustical signal with a particular loudness, roughness and fluctuating character, which, like a similarly fluctuating machine noise, should be perceived as very annoying. One might speculate that listeners make allowances pre- and post-task for what should be an annoying acoustical signal because they are giving it enough attention to realise it is, or could be, conveying information. This question shall be addressed in future experiments.

4.5. Further research perspectives

Assuming that annoyance assessed during laboratory exposure, and (retrospective or concurrent) annoyance ratings collected in field surveys share a common basis in the participants' responses, it might be interesting to establish whether the relationships found (between disruption and annoyance, and between stimulus duration and annoyance, respectively) also hold for real-life noise exposures. That might be difficult to determine, though, since field research does not offer experimental control, e.g. for varying the disruption potential of the sound exposures (as done in the present study), for varying task properties in order to make the disruption more or less evident, or for manipulating exposure duration. To encounter the appropriate levels of these variables 'naturally' is highly unlikely, unless a very costly 'field experiment' is performed. Furthermore, field surveys typically have no way of measuring disruption objectively (as done in the present study), and to query participants about their impression of the disruption experienced—as is frequently done—carries the risk of not being able to distinguish it from annoyance proper. Nevertheless, it might be promising to take the hypotheses produced by this laboratory study one by one, and to check them against relevant results emanating from field research.

Acknowledgements

This research was carried out while all three authors were at the "Sound Quality Research Unit" (SQRU) at Aalborg University. The unit was funded and partially staffed by Bang & Olufsen, Brüel & Kjær, and DELTA

Acoustics & Vibration. Further financial support came from the Ministry for Science, Technology, and Development (VTU), and from the Danish Research Council for Technology and Production (FTP). The authors would like to thank Søren Legarth and Signe Kristensen for collecting portions of the data, Sylvain Choisel for help with stimuli and programming, Florian Wickelmaier for advice regarding statistics, Wookeun Song for the use of his voice and all test subjects for their time.

References

[1] World Health Organization, in: B. Berglund, T. Lindvall, D.H. Schwelam, *Guidelines for Community Noise*, World Health Organization, Geneva, 2000.

[2] R. Guski, U. Felscher-Suhr, The concept of noise annoyance: how international experts see it, *Journal of Sound and Vibration* 223 (4) (1999) 513–527.

[3] C. Harris, *Handbook of Noise Control*, second ed., McGraw-Hill, New York, 1979.

[4] C. Marquis-Favre, E. Premat, D. Aubrée, Noise and its effects—a review on qualitative aspects of sound. Part II: noise and annoyance, *Acta Acustica United with Acustica* 91 (2005) 626–642.

[5] E. Zwicker, A proposal for defining and calculating the unbiased annoyance, in: A. Schick (Ed.), Contributions to Psychological Acoustics, Oldenburg, BIS, 1991, pp. 187–202.

[6] S.L.V. Moran, M. Loeb, Annoyance and behavioral aftereffects following interfering and noninterfering aircraft noise, *Journal of Applied Psychology* 62 (6) (1977) 719–726.

[7] A. Kjellberg, B. Sköldström, Noise annoyance during the performance of different nonauditory tasks, *Perceptual and Motor Skills* 73 (1) (1991) 39–49.

[8] R. Guski, H.-G. Bosshardt, Gibt es eine "unbeinflußte" Lästigkeit? [Is there an 'unbiased' annoyance?], *Zeitschrift für Lärmbekämpfung* 39 (1992) 67–74.

[9] U. Landström, A. Kjellberg, M. Byström, Acceptable levels of tonal and broad-band repetitive and continuous sounds during the performance of nonauditory tasks, *Perceptual and Motor Skills* 81 (3 Part 1) (1995) 803–816.

[10] U. Landström, L. Söderberg, A. Kjellberg, B. Nordström, Annoyance and performance effects of nearby speech, *Acta Acustica United with Acustica* 88 (4) (2002) 549–553.

[11] K. Persson Waye, J. Bengtsson, A. Kjellberg, S. Benton, Low frequency noise 'pollution' interferes with performance, *Noise and Health* 4 (2001) 33–49.

[12] J. Bengtsson, K. Persson Waye, A. Kjellberg, Evaluations of effects due to low-frequency noise in a low demanding work situation, *Journal of Sound and Vibration* 278 (1–2) (2004) 83–99.

[13] J.F. Wohlwill, J.L. Nasar, D.M. DeJoy, H.H. Foruzani, Behavioral effects of a noisy environment: task involvement versus passive exposure, *Journal of Applied Psychology* 61 (1) (1976) 67–74.

[14] A. Smith, D. Jones, Noise and performance, in: D. Jones, A. Smith (Eds.), *Handbook of Human Performance*, Academic Press, London, 1992, pp. 1–28.

[15] R. Hughes, D.M. Jones, The intrusiveness of sounds: laboratory findings and their implications for noise abatement, *Noise and Health* 4 (13) (2001) 51–70.

[16] S. Banbury, W. Macken, S. Tremblay, D. Jones, Auditory distraction and short-term memory: phenomena and practical implications, *Human Factors* 43 (1) (2001) 12–29.

[17] K. Hiramatsu, K. Takagi, T. Yamamoto, J. Ikeno, The effect of sound duration on annoyance, *Journal of Sound and Vibration* 59 (4) (1978) 511–520.

[18] T. Poulsen, Influence of session length on judged annoyance, *Journal of Sound and Vibration* 145 (2) (1978) 217–224.

[19] J. Fields, R. DeJong, T. Gjestland, I. Flindell, R. Job, S. Kurra, P. Lercher, M. Vallet, T. Yano, R. Guski, U. Felscher-Suhr, R. Schuemer, Standardized general-purpose noise reaction questions for community noise surveys: research and a recommendation, *Journal of Sound and Vibration* 242 (4) (2001) 641–679.

[20] ISO/TS 15666:2003, *Acoustics—assessment of noise annoyance by means of social and socio-acoustic surveys*, International Organization for Standardization, Genève, 2003.

[21] S. Maxwell, H. Delaney, *Designing experiments and analyzing data: A model comparison perspective*, Pacific Grove, CA, 1990.

[22] W. Ellermeier, K. Zimmer, Individual differences in susceptibility to the "irrelevant speech effect," *Journal of the Acoustical Society of America* 102 (4) (1997) 2191–2199.

[23] D. Jones, P. Beaman, W. Macken, The object-oriented episodic record model, in: S. Gathercole (Ed.), *Models of Short-Term Memory*, Psychology Press, Hove, UK, 1996, pp. 209–237.

[24] D. Jones, W. Macken, A. Murray, Disruption of visual short-term memory by changing-state auditory stimuli: the role of segmentation, *Memory and Cognition* 21 (3) (1993) 318–328.

[25] ISO 389, Part 1, *Standard reference zero for the calibration of pure tone air conduction audiometers, part 1: reference equivalent threshold sound pressure levels for pure tones and supra-aural earphones*, International Organization for Standardization, Genève, 1998.

[26] W. Ellermeier, J. Hellbrück, Is level irrelevant in 'irrelevant speech'? Effects of loudness, signal-to-noise-ratio and binaural unmasking, *Journal of Experimental Psychology: Human Perception and Performance* 24 (5) (1998) 1406–1414.

[27] S. Tremblay, A.P. Nicholls, D. Alford, D. Jones, The irrelevant sound effect: does speech play a special role?, *Journal of Experimental Psychology: Learning, Memory and Cognition* 26 (6) (2000) 1750–1754.