Applied Acoustics 86 (2014) 1–16

Contents lists available at ScienceDirect

Applied Acoustics

journal homepage: www.elsevier.com/locate/apacoust

Effects of unattended speech on performance and subjective distraction: The role of acoustic design in open-plan offices

Annu Haapakangas [a,*], Valtteri Hongisto [a], Jukka Hyönä [b], Joonas Kokko [b], Jukka Keränen [a]

[a] *Finnish Institute of Occupational Health, Indoor Environment Laboratory, Lemminkäisenkatu 14-18 B, FI-20520 Turku, Finland*
[b] *University of Turku, Department of Psychology, 20014 Turku, Finland*

ARTICLE INFO

Article history:
Received 25 October 2013
Received in revised form 19 March 2014
Accepted 16 April 2014

Keywords:
Speech Transmission Index
Cognitive performance
Speech intelligibility
Masking
Acoustic satisfaction
Noise effects

ABSTRACT

Unattended background speech is a known source of cognitive and subjective distraction in open-plan offices. This study investigated whether the deleterious effects of background speech can be affected by room acoustic design that decreases speech intelligibility, as measured by the Speech Transmission Index (STI). The experiment was conducted in an open-plan office laboratory (84 m^2) in which four acoustic conditions were physically built. Three conditions contained background speech. A quiet condition was included for comparison. The speech conditions differed in terms of the degree of absorption, screen height, desk isolation, and the level of masking sound. The speech sounds simulated an environment where phone conversations are heard from different locations varying in distance. Ninety-eight volunteers were tested. The presence of background speech had detrimental effects on the subjective perceptions of noise effects and on cognitive performance in short-term memory and working memory tasks. These effects were not attenuated nor amplified within a three-hour working period. The reduction of the STI by room acoustic means decreased subjective disturbance, whereas the effects on cognitive performance were somewhat smaller than expected. The effects of room acoustic design on subjective distraction were stronger among noise-sensitive subjects, suggesting that they benefited more from acoustic improvements than non-sensitive subjects. The results imply that reducing the STI is beneficial for performance and acoustic satisfaction especially regarding speech coming from more distant desks. However, acoustic design does not sufficiently decrease the distraction caused by speech from adjacent desks.

1. Introduction

Acoustic problems of open-plan offices have been widely documented in the literature. These problems are not only manifested as increased noise complaints (e.g., [1]), but have also been associated with a variety of negative outcome variables, such as noise-related stress [2], decreased environmental satisfaction [3], decreased job satisfaction [3], impaired concentration [4], and decreases in self-estimated work performance [5]. However, the environmental problems of open-plan offices are not confined to office noise. Open-plan offices have also been associated with increased complaints about most indoor environmental factors (e.g., [6,1]), lack of privacy (e.g., [7]), decreased satisfaction with the overall environment [5], increased cognitive workload [7], increased prevalence of different symptoms [1], and increased sickness absence [8]. Therefore, it is difficult to separate the effect of acoustic conditions on work performance and well-being from other factors that may confound the perception and the impact of the acoustic environment.

In recent years, a growing number of researchers have adopted an experimental approach to study the effects of open-plan office noise on performance. In this approach, performance effects are tested with different cognitive tests in laboratory settings employing methods from experimental psychology. A few of these studies have tested office noise exposure *per se* using noise that consists of a variety of office sounds [9–11]. However, most studies have focused on the effects of background speech (e.g., [12–17]). The latter approach is motivated by several reasons. Firstly, speech sounds tend to be mentioned as the most distracting noise source by office workers (e.g., [5,18]). Secondly, basic cognitive research has repeatedly demonstrated that background speech impairs cognitive performance, and that these effects are larger than those produced by non-speech noise (for a meta-analysis, see [19]). Thirdly, the practical relevance of this research area has increased since the publication of a new international room acoustic

---

* Corresponding author. Tel.: +358 304741.
E-mail addresses: annu.haapakangas@ttl.fi, armhaa@utu.fi (A. Haapakangas).

http://dx.doi.org/10.1016/j.apacoust.2014.04.018
0003-682X/

measurement standard for open-plan offices [20] in which reducing the distraction of background speech is the most essential idea.

The present study focuses on the role of room acoustic design in decreasing cognitive and subjective distraction that is caused by background speech.

1.1. Speech intelligibility and cognitive performance

Unattended background speech has been shown to affect several cognitive tasks, such as short-term memory [21], mental arithmetic [22], reading comprehension [23], proofreading [24], and writing performance [25]. According to the interference-by-process account, the performance disruption depends on the interplay between the properties of the sound and those of the task (e.g., [26]). More specifically, performance disruption is caused when the processes engaged by the automatic processing of the sound overlap with those needed in the focal task. Thus, the meaningfulness of speech is only seen as relevant in tasks requiring semantic processing [26], whereas the acoustic variation of speech explains the decrements observed in serial memory [27], regardless of whether the speech is comprehensible or not [24].

Unlike with some other noise sources, the performance impairment caused by speech does not depend on the sound pressure level, but rather on the intelligibility of speech [28,29]. While the effects of speech on performance have been extensively researched, few researchers have focused on the role of speech intelligibility. The latter studies show that performance in several cognitive tasks deteriorates with increasing intelligibility (e.g., [13,14,22,29]). Similarly, speech intelligibility predicts a variety of subjective responses, such as acoustic satisfaction [12,30], perceived disturbance [13,22], subjective habituation [17], and subjective workload [12].

The method of determining speech intelligibility has differed between studies, varying from subjective listening tests [15,22] to the signal-to-noise ratio [28,29,31] and the Speech Transmission Index (STI) [12–14,17,32]. While the first two methods provide valuable information about the effect of speech intelligibility in general, the latter approach is more beneficial for applied research because the STI is commonly used in evaluating and designing room acoustics. The STI is also a key quantity in the new international measurement standard [20]. By using the STI, cognitive laboratory experiments can be linked to predicting how acoustic conditions affect performance in office environments.

The STI is an objective descriptor for subjective speech intelligibility (STI 0.00 = not intelligible, STI 1.00 = perfectly intelligible). In practice, the STI of speech depends on absorption, screens, background sound level, and the distance between a speaker and a listener [33]. According to the model proposed by Hongisto [34], cognitive performance deteriorates with increasing STI. Performance is expected to start to decline above STI 0.20 and reach the maximum decrement when STI 0.60 is exceeded. The steepest decline is expected in the range of STI 0.30–0.50. Hongisto [34] concluded that in order to decrease the detriments of background speech, the STI should be below 0.50. This idea is included in the ISO 3382-3 standard [20] as distraction distance (r_D) which defines the distance at which the STI falls below 0.50.

However, the STI-performance model [34] may be somewhat debatable because the model was based on only three experimental studies available at the time. Later studies have given some support to the model by showing that the steepest decline in performance occurs somewhere between STI 0.38 and 0.62 [12,13]. However, recent findings by Jahncke et al. [14] and Keus van de Poll et al. [32] suggest that the maximum deleterious effect on performance might already be reached at STI 0.34. The inconsistency of these studies may be explained by the use of different tasks, as it seems likely that the STI-performance relation is to some degree task-specific [14]. Other methodological differences between the studies may also account for the results. Given the divergent findings and the small number of studies conducted, the relationship between the STI and cognitive performance requires more research. This knowledge would also be beneficial for evaluating whether the acoustic criteria adopted in the ISO 3382-3 standard [20] are sufficient in terms of the desired effects on background speech distraction.

1.2. Limitations in studies on speech intelligibility

Most of the studies on the STI and speech intelligibility have, to a varying degree, attempted to provide practical implications for office environments. However, there are several aspects in which the ecological validity of experiments on the STI-performance relation can still be improved.

As far as speech material is concerned, most speech intelligibility studies have used continuous speech either comprising of very simple successive sentences with no plot [14,22] or a story in the native [32] or a foreign language [29]. A few studies lack the description of the sound materials used [15,31]. However, the meta-analysis of Szalma and Hancock [19] has shown that intermittent background speech causes greater performance impairment than continuous speech. Background speech is also more distracting when it represents half of a dialogue, as in overhearing a phone conversation [35]. Both intermittent speech and one side of a phone conversation are characteristic of many open-plan offices, whereas monologue-like speech is not. In some tasks, constant speech can also be habituated [36], which may lead to imprecise conclusions if used to represent open-plan office noise. Three studies [12,13,17] have used intermittent speech with short pauses of varying length between sentences but the speech was designed to be calm and uninteresting, which may represent some but not the most distracting office discussions.

Another factor that may affect performance effects of speech is the location of the speech source. A few studies suggest that performance is affected most when speech originates from the same direction where visual attention is actively engaged in [37,38], although this may not be true for all tasks [24]. In open-plan offices, the location of the speech source varies, as does the distance to different speakers present in the room. The speech intelligibility studies have usually used speech from one static location, typically in front of the subject, and presented the speech via loudspeakers or using a headphone simulation [14,15,17]. Some studies have used headphones without specifying the perceived speech location [29,32]. Only two studies have used multiple locations for speech sources [12,13]. Schlittmeier and Hellbrück [31] have used open-plan office noise that was recorded with an artificial head, presumably including variation in sound direction and distance, but they reported no details to describe the sound material. However, none of the studies have intentionally varied the STI within a test condition which would simulate open-plan offices where the distance to different speakers varies.

The effect of the exposure time has also been neglected in speech intelligibility experiments. The single speech conditions have typically lasted for less than an hour (e.g., [12,14]), while the longest duration seems to have been 1.5 h at the most [15]. The performance effects of noise in general tend to attenuate with longer exposure times but this may not apply to background speech [19]. In open-plan offices, one working period between breaks might last for three-to-four hours. While it is possible that the effect of background speech could be adapted to over time, it is also conceivable that cognitive or subjective impacts of noise might increase as a result of an emerging stress response or decreasing compensatory resources.

The final criticism on previous studies concerns the creation of different degrees of speech intelligibility. Previous experiments have manipulated speech artificially, either by changing the relative sound pressure levels of speech and masking sound electronically (e.g., [13]) or by using simulation techniques, such as auralization [22]. While this approach is well-suited for charting the general relation between speech intelligibility and performance, it may lead to test conditions and conclusions that are unrealistic from a practical point of view. In practice, acoustic conditions are controlled by changing wall and ceiling materials, furniture, screen height and desk closure, and by using a masking sound. To date, the experimental literature is lacking studies where speech intelligibility has been modified analogously to how it would be done in an authentic open-plan office.

1.3. The aim of the present study

In this paper, we present an experiment that attempted to integrate the findings discussed above in creating as ecologically valid test conditions as possible, both in terms of background speech and the room acoustic manipulations. This study is the first where the acoustic test conditions were created by introducing physical changes in the acoustic design of an office laboratory, instead of relying on the artificial simulation of sounds. We tested three different room acoustic conditions containing background speech and one condition without speech (a quiet condition). The speech sounds simulated an environment where one side of different phone conversations is heard from four locations differing in distance. The STI of the experimental conditions was manipulated by changing the amount of absorption (i.e., screen height and room absorption together) and the level of masking sound while keeping the sound power level of speech constant.

The aim was to investigate whether cognitive performance and subjective perceptions of the acoustic environment can be affected by an acoustic design that reduces speech intelligibility (STI). We chose to test these effects with four memory tasks, focusing mostly on working memory because working memory tasks predict cognitive performance in general (see e.g., [39]). The choice of tasks also enables the comparison of the present results with some of the earlier studies. Noise sensitivity was included in the experimental design because it is strongly associated with subjective responses, such as annoyance, but may also have an effect on performance in noisy conditions (see [40]).

Three research questions were addressed: (1) Are cognitive performance and subjective disturbance affected by room acoustic design that reduces the STI? (2) Is there an interaction between the acoustic condition and exposure time? and (3) Are subjects with high or low noise sensitivity differently affected by acoustic design? In terms of the first question, we expected that the presence of speech sounds would have the strongest effect, regardless of the acoustic design. However, within speech conditions, performance and subjective perceptions should improve when the STI is decreased by acoustic design. In terms of exposure time, the purpose was to explore whether the effects of background speech would attenuate or amplify with exposure time. We expected that noise sensitivity would affect the subjective perceptions of noise effects, with noise-sensitive individuals perceiving the speech conditions more negatively. We also wanted to explore whether noise sensitivity would moderate the effect of the acoustic conditions on subjective perceptions.

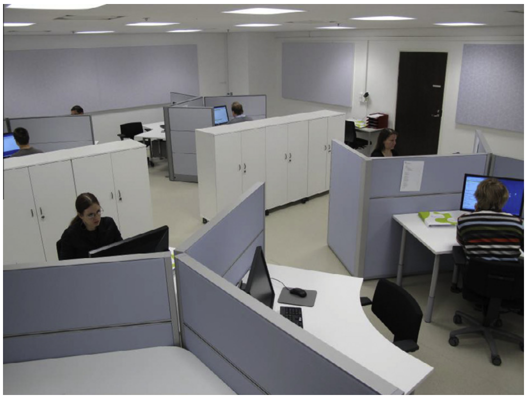

Fig. 1. A photograph of the laboratory (the condition *noAbs_noMask*).

2. Methods

2.1. Subjects

Ninety-eight university students took part in the experiment. One subject was excluded because she reported major sleep deprivation. The remaining 97 subjects (73 female) ranged from 19 to 45 in age (mean = 23.9, SD = 4.0) and were native Finnish speakers. None of the subjects had dyslexia or reported any hearing difficulties. Subjects were paid 40 euro for their participation.

2.2. Test laboratory

The experiment was conducted in an open-plan office laboratory that was 8.9 × 9.4 × 2.55 m in size (Figs. 1 and 2). The room resembled a real open-plan office with no measurement devices or other disturbing artifacts visible in the room, apart from four loudspeakers. The furniture included 12 identical desks, frontal screens (2.4 m per desk, floor-standing), additional side screens (1.6 m per desk, floor-standing) and storage units in the middle of the room. Six desks were used by the subjects and another two were in reserve. The corner desks were reserved for the loudspeakers that were used to play speech sounds. The distance from each subject desk to the nearest and the most remote loudspeaker ranged from 2 to 6 m.

The suspended ceiling consisted of a 600 × 600 mm metal grid where 210 ceiling boards, seven ventilation units, and 16 lighting units were installed. The height of the ceiling was 2.55 m and the suspension depth was 0.3 m. Fourteen loudspeakers were placed above the suspended ceiling for the production of artificial masking sound. Approximately 88% of the ceiling (75 m^2) was reserved for ceiling boards that had sound-absorbing and sound-reflecting alternatives. Approximately 20% (18 m^2) of the total wall area was reserved for sound-absorbing wall boards. The ventilation system was running at a constant power, producing an inherent masking sound level of 33–37 dB L_{Aeq} depending on room absorption.

The indoor air conditions were monitored during each test session. The room temperature was 23.2 ± 0.2 °C, which is optimal given the activity level and the presumed clothing of the subjects [41]. There was no draft at any desk. The air quality was very good with a low CO_2 concentration (580 ± 50 ppm). The vertical illumination level was 630 ± 120 lx on the table surfaces.

2.3. The experimental conditions

Four experimental conditions were investigated. The conditions were created by modifying absorption, screen height, desk isolation, and the level of masking sound (Table 1). Three of the conditions included background speech and differed in terms of the room acoustic design. The fourth condition, *Quiet*, was the

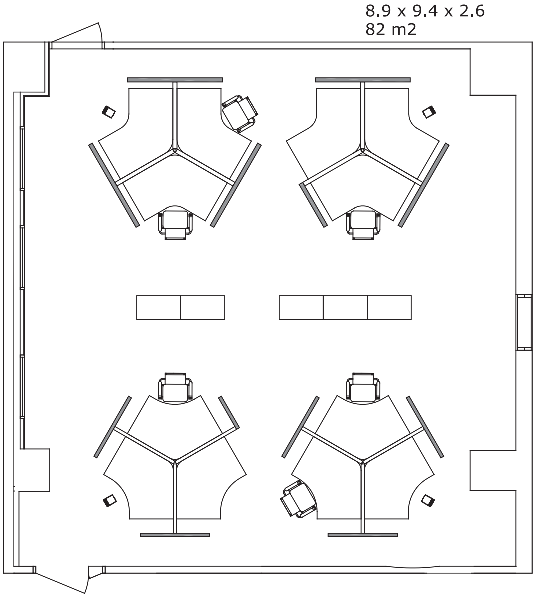

Fig. 2. The layout of the office laboratory. The chairs indicate the desks where the subjects were seated. The additional side-screens are indicated by gray shading. The loudspeakers were placed at the corner desks. Because only one loudspeaker was active at a time, the distance to an active speaker varied over the experiment from 2 m (the nearest loudspeaker) to 6 m (the furthest loudspeaker).

reference condition in which background speech was off. The conditions were:

(1) No absorption nor masking sound (later: *noAbs_noMask*). This condition corresponded to a situation with no special room acoustic treatments. The ceiling, walls, and screens were sound-reflecting and the screen height was low. The background sound level, produced by ventilation, was very low. As a result, background speech was intelligible both at a near and a far distance from the speech source. This is a typical situation when an acoustic designer has not been involved in workplace design.

(2) Absorption used, no masking sound (*Abs_noMask*). This condition corresponded to a situation where all other noise control measures are used, except for a masking sound system. The absorption was maximized on the ceiling, walls, and screens. The isolation of the desks was increased with higher screens and additional side-screens (Fig. 2). The background sound level, produced by ventilation, was very low. Speech intelligibility was still high near the speech source but reduced with increasing distance from the speaker. We did not expect the results in this condition to significantly differ from *noAbs_noMask* as the differences in speech intelligibility were small.

(3) Both absorption and masking sound used (*Abs_Mask*). This condition corresponded to a situation where all possible noise control measures are used. In addition to the means used in *Abs_noMask*, a masking sound system was applied using a recommended level (45 dB L_{Aeq}). As a result, the speech intelligibility was a little reduced near the speech source and significantly reduced with increasing distance. Testing the benefits of this condition in comparison to the conditions *noAbs_noMask* and *Abs_noMask* was an essential aim in the present study.

(4) *Quiet*, which corresponded to a situation where several people work in an open-plan office in silence. It was included as the reference condition. However, the presence of other subjects might have caused minor temporary distraction (e.g., coughing, keyboard tapping) and thus the condition does not correspond to perfect silence which is an often-used control condition in cognitive experiments.

Table 1
The room acoustic components in the four experimental conditions.

	Acoustic condition			
	noAbs_noMask	*Abs_noMask*	*Abs_Mask*	*Quiet*
Background speech	ON	ON	ON	OFF
Expected acoustic privacy	Poor	Better	Good	Perfect
Ceiling absorbents installed:	No	Yes	Yes	
Wall absorbents installed:	No	Yes	Yes	
Screen absorption installed:	No	Yes	Yes	
Screen height (m)	1.3	1.7	1.7	1.3
Side screens:	No	Yes	Yes	
Masking sound level (dBA)	37	33	45	35

The condition *noAbs_noMask* was implemented by installing sound reflecting gypsum boards to the ceiling (EN 11654 [42], unclassified). No sound absorbing boards were installed on the walls. Sound-reflecting screens were used between the desks, with a height of only 1.3 m. Artificial speech masking system was off. The background sound was mainly produced by the ventilation system. However, the spatial variation of ventilation noise was rather large, ranging from 33 to 37 dB L_{Aeq}. To correct for this variation, the masking sound system was played at a very low level (36 dB) to increase the noise levels in silent desks. Thus, the final background level at the desks was 37 ± 1 dB L_{Aeq} (Fig. 3). The speech loudspeakers were on.

The condition *Abs_noMask* was implemented by placing sound absorbing mineral wool to the ceiling (EN 11654 [42], class A, total area 75 m^2) and on three walls (class A, total area 18 m^2). The screens were 1.7 m high and sound-absorbing (EN 11654 [42], class B, one-sided area). The additional side-screens that were used to increase isolation were not sound-absorbing (unclassified). Overall, the total area of sound-absorbing material surfaces in the room was 142 m^2 larger than in the condition *noAbs_noMask*. The spatial variation of the ventilation noise was corrected with the sound masking system similarly as in *noAbs_noMask*. Because the high room absorption also reduced the ventilation noise level, the background sound level was 4 dB lower (i.e., 33 ± 1 dB) than in *noAbs_noMask* (Fig. 3). The speech loudspeakers were on.

The condition *Abs_Mask* was implemented by applying all the measures of *Abs_noMask* with the addition of the speech masking system. The masking sound was filtered pink noise with −5 dB/octave slope within octave bands 125–4000 Hz (Fig. 3), which has also been found adequate in workplaces [43,44]. The masking level was 45 dB conforming to the current target values [45]. The variation of masking level between the desks was very small (±1 dB). The speech loudspeakers were on.

The condition *Quiet* was implemented by having the speech loudspeakers off. The background sound level was set to 35 dB. The room acoustic materials were the same as in *noAbs_noMask* but, due to the absence of speech, the room acoustic conditions were irrelevant.

The original experiment included also a fifth condition that included the masking sound but no absorption. These results have been reported by Haapakangas et al. [46].

2.4. Speech sounds

Creating a natural speech environment was an important methodological aim in this experiment. The presentation of speech

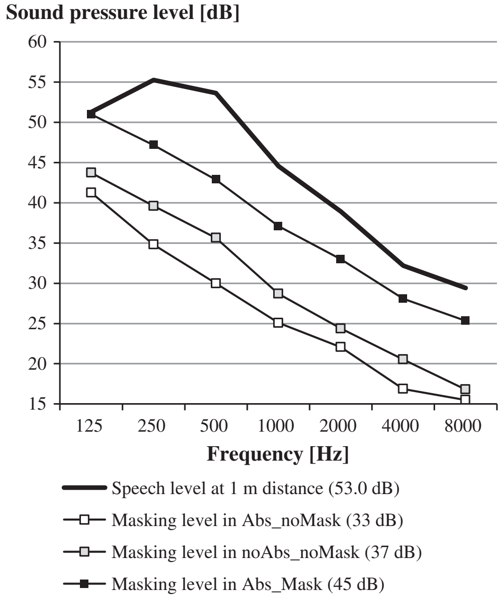

Fig. 3. The spectra of the speech and masking sounds. The speech spectra were measured at a distance of 1 m in free field from the loudspeaker. The speech levels heard by the subjects at the desks were lower and they are described in Table 2 and Fig. 5.

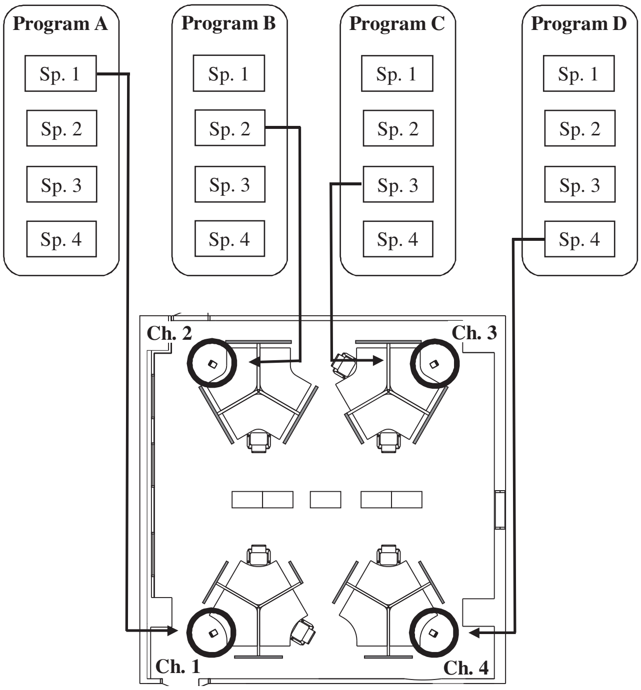

Fig. 4. The distribution of the individual voices in different radio programs to the four speaker channels in the room corners.

sounds was designed to resemble a situation where four people are talking on the phone about different topics one at a time in different corners of the room. Unlike in the previous studies reviewed in the Introduction, more natural, lively, and interesting speech was desired for the present study. Some of the voices were recognizable, as they would also be in real offices.

Another principal aim was to keep the sound power level of speech emerging from the loudspeakers constant in every acoustic condition. Thus, the speech level that was heard at the subject desks depended solely on the acoustic treatments of the acoustic condition (Table 1).

A four-channel sound file was used to produce speech to four loudspeakers placed in the corner desks (Fig. 4). The speech was obtained from four radio programs bought from YLE (the Finnish Broadcasting Company). In each of them, four participants (politicians or celebrities) were engaged in a friendly debate about a topic of common interest. The speech of each participant was isolated from each program and placed to one channel of the four-channel sound file. The speech material of a single speaker (i.e., each channel) was then cut to separate 5-to-25-s-long sentences. Thereafter, the four-channel speech file was expanded and arranged to remove any overlap between the channels. The sound levels of the sentences were individually adjusted to the same A-weighted level to correct for variations in speech effort. In addition, the spectrum of each speaker was modified so that the octave band levels were deviating from the speech spectrum shape of ISO 3382-3 [20] by less than 3 dB. A 1-to-8-s break was placed between the speech sentences when a switch to another corner (i.e., channel) took place. The order of the channels was pseudo-randomized so that one speaker was not active twice in succession and the total amount of speech from every channel was equal. Adobe Audition 3.0 software was used for editing. The length of the final four-channel playback recording was 181 min.

The speech was produced by four loudspeakers having mouth-like directivity (Genelec 6010). Only one loudspeaker was active at a time. The height of the loudspeakers was 1.2 m from the floor. To ensure that the speech level was constant the sound power level of each channel was measured in a reverberant room according to ISO 3741 [47].

The speech level was selected to be between a normal and casual speech effort because conversational speech levels in open-plan offices may be slightly lower than the standardized normal speech effort [48]. The A-weighted mean level over all directions was 53.0 dB at a 1 m distance in a free field. (For comparison, the standardized normal effort speech is 57.4 dB according to ISO 3382-3 [20]). The measured sound pressure levels were 51.3, 55.3, 53.6, 44.6, 38.9, 32.2, and 29.4 in octave bands 125, 250, 500, 1000, 2000, 4000, and 8000 Hz, respectively (Fig. 3).

2.5. Acoustic measurements of the experimental conditions

The room acoustic conditions were measured in two different ways. To determine the exact values of the STI as experienced by the subjects, the speech level measurements were made for the speech levels prevailing at the desks in different acoustic conditions. These measurements were conducted with the speech loudspeakers used in the experiment. Secondly, the standard method ISO 3382-3:2012 [20] was applied in order to objectively describe the room acoustic conditions. This information is important for the acoustic characterization of offices. These measurements were done using an omni-directional loudspeaker, and the results are shown in Appendix A.

The results of the actual speech levels and the STI measurements at the desks are shown in Table 2 and Fig. 5. The measurements confirmed that the implementation of the acoustic conditions was successful in terms of producing distinct differences between conditions. It should, however, be noted that the STIs in *noAbs_noMask* and *Abs_noMask* did not differ considerably even though the level of speech is reduced in the latter condition. Differences could only be found at larger distances from the speaker. The measurements have been reported in more detail by Keränen et al. [49,50].

2.6. Cognitive tasks

2.6.1. Serial recall

Serial recall is a classic task in investigating the cognitive effects of speech. In the task, the subjects had to recall visually presented

Table 2
The A-weighted sound levels, $L_{S,A}$, and the Speech Transmission Indices, STI, of the acoustic conditions.

	Acoustic condition			
	noAbs_noMask	*Abs_noMask*	*Abs_Mask*	*Quiet*
STI of speech from the nearby desk (2 m away)	0.70	0.80	0.51	0.00
STI of speech from the most remote desk (6 m away)	0.60	0.42	0.11	0.00
$L_{S,A}$ (dB) from nearby desk (2 m away)	53	50	50	0
$L_{S,A}$ (dB) from the most remote desk (6 m away)	50	38	38	0

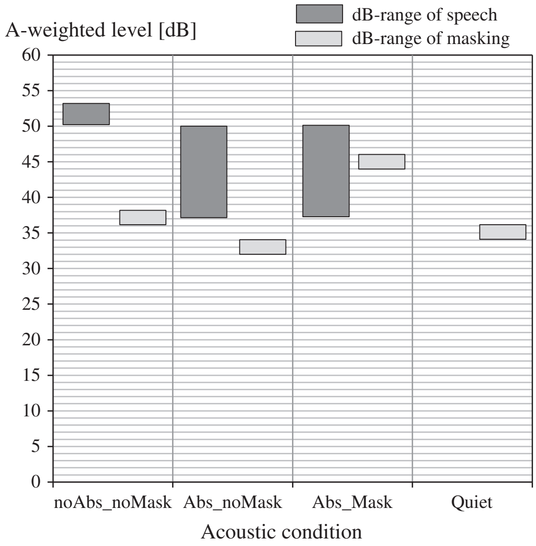

Fig. 5. The range of the speech and masking levels at the desks in the four acoustic conditions. The variation of the masking level was always negligible because of the 14 evenly distributed masking loudspeakers in the ceiling. However, the speech levels varied more because the distance to the presently active loudspeaker varied with time.

digits from 1 to 9 in the order of presentation. The digits were presented in a random order, each digit appearing once in a sequence. Digits were presented on the computer screen at the rate of 1 per second, with an inter-digit interval of 1.5 s, i.e. the onset-to-onset of digits was 2.5 s. After a sequence, the subjects recalled the digits by clicking numbers in the same order on a 3 × 3 array on the screen. The subjects had the possibility to click "empty" if they could not remember a digit in a certain serial position. The subjects were instructed to be as accurate as possible, but performance speed was not mentioned nor analyzed. The task contained 12 sequences of which the first was not analyzed. The percentage of digits recalled in their correct positions was calculated for each sequence and averaged across all 11 sequences for the total score.

2.6.2. Working memory capacity

Working memory (WM) capacity was included in the experimental design as a covariate because individual differences in cognitive abilities were anticipated to form a possible confounder in the experiment. The measure for WM capacity was modified from the original operation span task developed by Turner and Engle [51]. The task consists of equations and words presented by turns. The idea of the task is to measure the functional capacity of WM by including a memory storage component (word recall) and a secondary processing component (equations) that interferes with the memory task.

First, an equation appeared on the computer screen (for example, $9 \times 6 - 3 = 52$) and the subject had a maximum of 10 s to decide whether it was true or false by clicking the appropriate option on the computer screen. After this, a word to be memorized was presented for 2 s. This was followed by another equation, followed by another word. At the end of each set, the subject typed in all the words. A correct order for words within lists was not required. Subjects were instructed to aim at a minimum of 85% accuracy on equations, and received feedback on this after typing each word list. The purpose of this was to prevent them from easing the word recall by ignoring the secondary task.

The length of equation-word sets varied between three and seven pairs, and the sets were presented in a random order. The task consisted of ten sets. A score for the subject's WM capacity was calculated by the percentage of all correctly recalled words during the task (partial-credit load scoring, [52]. Minor misspellings were accepted, when unambiguous.

2.6.3. Operation span task

The operation span task was similar to the WM capacity task with the following changes. The set length varied from three to eight and the sets were presented in a growing order. The materials used for words and equations differed from the WM capacity task and are described in detail by Haka et al. [13]. The task contained 11 sets. Three matching versions were used. The equations and words were presented randomly within each version. The scoring procedure was the same as in the WM capacity task. The task lasted approximately 10 min.

2.6.4. N-back task

The N-back task is another type of working memory task (e.g., [53]). It is assumed to engage some of the key processes of working memory, such as online monitoring and updating information in memory. The task consisted of sequences of letters presented on a computer screen. Each letter was shown for 500 ms that was then followed by a 2500 ms pause before the next letter appeared. Three levels of difficulty were used: 0-back, 1-back and 2-back. In 0-back, the subject's task was to identify the target letter 'X' by pressing YES or NO for each stimulus. In 1-back, the subject had to decide whether the presented letter was identical to the one immediately preceding it (i.e. one trial back), again responding YES or NO. In 2-back, the task was to respond whether the presented letter was identical to the one presented two trials back. The subjects were instructed to respond quickly, but accurately. Each set contained $30 + n$ letters ($n = 0$, 1 or 2). 30% of the letters were targets requiring a positive response. Upper case and lower case letters were varied to prevent subjects from relying on visual recognition only. The task included three blocks containing one set of each level, i.e. each level of difficulty appeared three times during the task. As the task was presented three times during the experiment, there were materials for three versions. The materials and the order of the difficulty levels were counterbalanced across the test blocks and subjects. Response accuracy (%) and reaction times (RTs) for the correct responses (ms) were measured. The responses deviating over 3 SD from the subject's mean were excluded from the average RTs. The task lasted approximately 18 min.

2.6.5. Text memory task

The task consisted of a text developed by Kaakinen et al. [54]. The text discussed four diseases and was three pages long. The diseases were generally unfamiliar to the subjects in order to prevent the subjects from relying on their prior knowledge. The subject's task was to memorize facts concerning one of the diseases (typhus). Subjects were given 6 min to read the text. After this, the subjects performed other tasks (N-back and operation span task) that created an approximately 30-min delay between reading and recall. This was done to prevent active maintenance of the text in short-term memory between reading and recall. In recall, the subjects had 4 min to write as much as they could remember about the target disease using the WordPad program. The task was scored by giving one point for every detail that was remembered (max. 45 points) and calculating the percentage of correct responses.

2.7. Questionnaires

Three questionnaires were used. A questionnaire on background information included questions about age, gender, sleep during the preceding night, and noise sensitivity. The latter was assessed with a four-item version [55] of the subscale 'work' in NoiSeQ [56].

The NASA Task Load Index, NASA-TLX [57], was used for the assessment of subjective workload on six scales (mental, physical, and temporal demand, performance, effort, and frustration). The response was given on a slider bar. The poles of the scales were marked with *low* and *high*, except for the performance scale that was marked with *good* and *poor*. The scale ranged from one to 100 but the values were not shown to the subjects.

The final questionnaire included 17 statements about the perceived acoustic environment, its effect on performance and the overall environment. These were rated on a five-point scale (completely disagree – completely agree). Originally, most of the items were intended to be included in a sum score of acoustic satisfaction. However, this resulted in unacceptably low Cronbach's alphas in *Quiet*. The most reliable sum score was obtained from four questions and was labeled as *perceived disturbance* (Cronbach's alphas 0.65–0.83 in the different conditions). The variable included the following statements: It was easy to habituate to the sound environment (SE); The SE was pleasant; The SE impeded my ability to concentrate; My attention was drawn to the SE. In addition, the disturbance of 13 indoor environmental factors was rated on a five-point scale (not at all, slightly, to some extent, quite a lot, very much). Three of these items were essential for the research questions, namely the distraction caused by speech from nearby desks and desks further away and the distraction caused by background hum (i.e., ventilation or masking sound). The remaining items were included in order to confirm that the conditions were optimal in terms of other environmental factors.

2.8. Design

A mixed 4 × 3 × 2 design was used, with acoustic condition as a between-subjects factor (4 levels), test block (3 levels) as a within-subjects factor and noise sensitivity (2 levels) as another between-subjects factor. The test block refers to the exposure time, i.e. tasks were repeated in three blocks to assess the possible interactions between time and performance. There were 24 subjects in each group, except for the *Quiet* group, which had 25. Each subject only took part in one of the conditions. WM capacity was included as a covariate to increase the statistical power of the design. The different versions of the tasks were counterbalanced across test blocks and subjects.

The conditions were implemented in the following order: *noAbs_noMask*, *Quiet*, *Abs_noMask*, and *Abs_Mask*. The completion of the experiment took three months because considerable retrofitting and acoustic measurements were required between conditions. Due to this time frame, subjects could not be randomly assigned to different conditions but were acquired for one condition at a time. Finding male subjects proved difficult toward the end of the study, which resulted in an unbalanced gender distribution across conditions (13, seven, two and two males, respectively).

2.9. Procedure

The experiment started with general information about the study including an informed consent form. First, the subjects performed the measurement of WM capacity. After this, the subjects completed a questionnaire on background information and practiced all the tasks in silence. A short break was given after the practice session, if necessary. The masking sound was on from the beginning but the speech sounds were turned on at this point. The subjects were instructed to ignore the background sounds and concentrate on the tasks.

The tasks were presented in three blocks that lasted 40, 60, and 60 min, respectively. The serial recall, N-back, and operation span tasks were done in every block, in this order. The text memory task was performed during the second block so that the text was read after the serial recall task and the text recall took place at the end, after the operation span task. The last block had an additional 15-min filler task at the end of the block. Each task was followed by a NASA-TLX rating done on the computer. A ten-minute break was given after the first block. During this, the subjects were offered a sandwich outside the room but were instructed not to discuss the experiment with each other. At the end of the third block, the speech sounds were turned off. The subjects then completed a questionnaire on acoustic satisfaction and perceived sound environment in silence.

The experiment lasted 4 h and all the groups were tested at the same time of the day, starting at 9:30 a.m. Two to six subjects were tested at a time. A researcher was present in the laboratory at all times. After the experiment, the subjects received more information about the study and were given the opportunity to ask questions.

2.10. Statistical analysis

The data were analyzed with IBM SPSS Statistics for Windows, Version 20.0. (Armonk, NY: IBM Corp). The subjects were divided into two groups on noise sensitivity (sensitive vs. non-sensitive) by a median split on the summed noise sensitivity scores. Serial recall and operation span task were analyzed with 4 (acoustic condition) × 3 (test block) × 2 (noise sensitivity) ANCOVAs with WM capacity as a continuous covariate. Acoustic condition and noise sensitivity were between-subjects factors, whereas the test block was a within-subject factor. The text memory task was analyzed with a 4 (acoustic condition) × 2 (noise sensitivity) ANCOVA with WM capacity as a continuous covariate. The ANCOVAs were followed up with paired comparisons of the adjusted means. The assumptions of the linearity and homogeneity of regression were met in all ANCOVAs. The RTs in the N-back task were analyzed with a 4 (condition) × 3 (test block) × 3 (N-back level) × 3 (task block) × 2 (noise sensitivity) ANOVA, leaving out WM capacity as it did not have a linear relation with the RTs. When Mauchly's test indicated a violation of sphericity for the within-subject variables, the Greenhouse–Geisser correction was applied and the corresponding p-values are reported. The Benjamini–Hochberg procedure was used for alpha-error adjustments in all paired comparisons.

The sum score of perceived disturbance and the subjective distraction due to speech were analyzed with a non-parametric Kruskal–Wallis test, followed by Mann–Whitney *U* tests for paired comparisons. The same analyses were then repeated separately for noise-sensitive and non-noise-sensitive subjects.

The total scores and the subscales of the NASA-TLX were analyzed with a 4 (acoustic condition) × 3 (test block) × 2 (noise sensitivity) ANOVA in the short-term and working memory tasks and with a 4 (acoustic condition) × 2 (noise sensitivity) ANOVA in the text memory task. As the NASA-TLX has been associated with gender effects in some studies (e.g., [58]), the analysis was only performed for female subjects due to the unbalanced gender distribution across acoustic conditions.

One subject was excluded from the serial recall, operation span, and text memory tasks due to missing data on the covariate. Another subject was excluded from serial recall due to a technical error in data recording. Unlike in typical studies of the operation span task, we did not exclude subjects who failed to achieve an 85% level on equation accuracy because the independent variable (speech intelligibility) could also have affected arithmetic performance [14,22]. Instead, subjects with equation/word recall trade-offs were identified by examining multivariate outliers using the Mahalanobis distance. One such outlier was found and excluded from the operation span task. Two extreme outliers performing consistently over 3 SDs above group means were excluded from the N-back task. One subject was excluded from the analyses of the perceived disturbance due to a misunderstanding of the questions.

3. Results

The main assumptions of this experiment were that the presence of background speech would have detrimental effects on cognitive performance and subjective distraction, and that good room acoustic design would decrease these negative effects. Accordingly, we hypothesized that *Quiet* would yield the best results compared to all the other conditions, and that the condition with the lowest STI (*Abs_Mask*) would be better than the other two conditions containing speech (*noAbs_noMask* and *Abs_noMask*). Therefore, one-tailed tests are reported in the corresponding paired comparisons. The latter two conditions were not expected to differ from each other and were, therefore, compared with two-tailed tests.

3.1. Performance effects in cognitive tasks

3.1.1. Serial recall

The ANCOVA revealed a significant main effect of acoustic condition on performance in serial recall ($F_{3,86} = 3.32$, $p = .024$, partial $\eta^2 = 0.10$). As shown in Fig. 6, the results seem to follow the expected pattern, except for *noAbs_noMask* in which the performance was better than anticipated. Pairwise comparisons revealed that performance was better in *Quiet* than in *Abs_noMask* ($p = .01$, 1-tailed) and *Abs_Mask* ($p = .04$, 1-tailed), as expected. However, performance in *Quiet* was not better than performance in *noAbs_noMask* ($p > .05$, 1-tailed). In addition, *Abs_Mask* did not yield better results than the other two conditions containing speech (p values > .05, 1-tailed). Thus, the expectation that acoustic design would decrease the distraction caused by background speech was not supported by the serial recall data.

The main effect of test block was significant ($F_{2,172} = 4.11$, $p = .018$, partial $\eta^2 = 0.05$), indicating a small learning effect toward the end of the experiment. This effect did not interact with acoustic condition ($p > .05$), indicating that the effect of acoustic condition was not moderated by the exposure time.

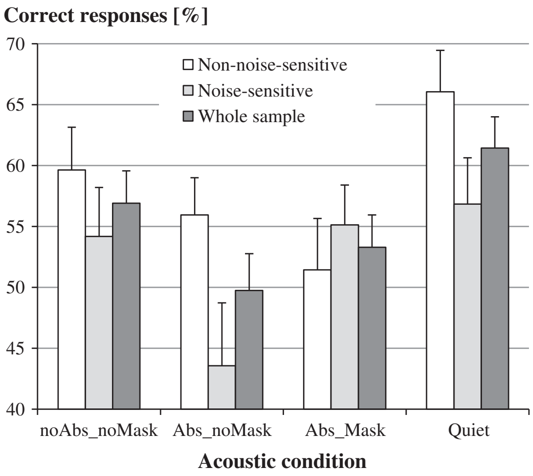

Fig. 6. Performance in serial recall in the whole sample and the noise sensitivity groups. Adjusted means and standard errors.

The non-noise-sensitive subjects performed better than the sensitive subjects ($F_{1,86} = 4.58$, $p = .035$, partial $\eta^2 = 0.05$; Fig. 6), but this effect did not interact with acoustic condition.

3.1.2. Operation span

The adjusted means in Fig. 7 show a pattern that is in line with the expectations. Accordingly, the ANCOVA showed a significant main effect of acoustic condition on performance in the operation span task ($F_{3,86} = 3.58$, $p = .017$, partial $\eta^2 = 0.11$). Pairwise comparisons confirmed that performance was better in *Quiet* than in *noAbs_noMask* ($p = .01$, 1-tailed) and *Abs_noMask* ($p = .03$, 1-tailed). There was also a marginal improvement in *Abs_Mask* compared to *noAbs_noMask* ($p = .07$, 1-tailed) suggesting a tendency toward an effect of acoustic design within the speech conditions. The main effect of test block indicated a small learning effect ($F_{2,172} = 6.11$, $p = .003$, partial $\eta^2 = 0.07$) but test block did not interact with acoustic condition ($p > .05$). Noise sensitivity had a small effect on task performance ($F_{1,86} = 4.28$, $p = .042$, partial $\eta^2 = 0.05$), with the non-sensitive subjects performing better than the sensitive subjects.

3.1.3. N-back

The initial analyses of the N-back task indicated a possible ceiling effect in accuracy as it exceeded 90% in all acoustic conditions on all N-back levels. Thus, only the analyses on RTs are reported.

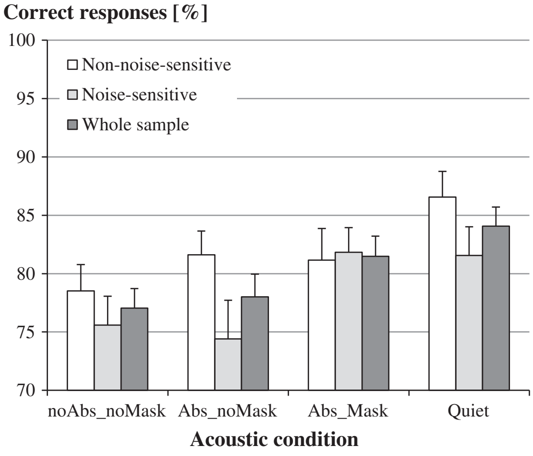

Fig. 7. Performance in the operation span task in the whole sample and the noise sensitivity groups. Adjusted means and standard errors.

The pattern of RTs in Fig. 8 suggests an increase in the cognitive load with increasing STI, although the RTs in *Quiet* appear slightly higher than expected. The effect of acoustic condition on the RTs was indeed significant ($F_{3,87} = 4.42$, $p = .006$, partial $\eta^2 = 0.13$). Paired comparisons revealed that RTs were significantly longer in *noAbs_noMask* than in *Abs_Mask* ($p = .002$, 1-tailed) and *Quiet* ($p = .03$, 1-tailed), indicating a higher cognitive load in *noAbs_noMask*. These results are in line with the predictions. However, no other expected differences emerged between *Quiet* and the speech conditions or within the speech conditions.

There was also an interaction between acoustic condition and the difficulty level of the task ($F_{6,174} = 4.07$, $p = .007$, partial $\eta^2 = 0.12$). As shown in Fig. 9, the acoustic condition seemed to have a larger effect with increasing difficulty. This was confirmed with separate ANOVAs for each N-back level. There was no effect on 0-back ($p > .05$), whereas a statistically significant effect emerged for both 1-back ($F_{3,87} = 4.00$, $p = .01$, partial $\eta^2 = 0.12$) and 2-back ($F_{3,87} = 4.79$, $p = .004$, partial $\eta^2 = 0.14$). The RTs were again significantly shorter in *Quiet* than in *noAbs_noMask* in both 1-back ($p = .03$, 1-tailed) and 2-back ($p = .02$, 1-tailed). Similarly, *noAbs_noMask* resulted in significantly longer RTs compared to *Abs_Mask* in both 1-back ($p = .004$, 1-tailed) and 2-back ($p = .001$, 1-tailed). The RTs in 2-back were also marginally longer in *noAbs_noMask* than in *Abs_noMask* ($p = .056$, 2-tailed), which was not expected.

There was a strong learning effect, as revealed by a main effect of test block ($F_{2,174} = 65.57$, $p < .001$, partial $\eta^2 = 0.43$) but again no interaction with acoustic condition ($p > .05$). Noise sensitivity did not affect RTs in the N-back task ($p > .05$).

Text memory. The text memory task was not affected by acoustic condition nor noise sensitivity, as shown by the ANCOVA yielding no significant effects (all p values > .05, Fig. 10).

The individual differences in WM capacity had a significant effect on performance in all tasks (all p values $\leqslant .003$, partial $\eta^2 = 0.07$–0.52), except the N-back task in which it was not included. There were no interactions between WM capacity and acoustic condition in any of the tasks.

3.2. Effects on subjective variables

3.2.1. Perceived disturbance

The sum score of perceived disturbance was affected by acoustic condition (χ^2 (3) = 57.83, $p < .001$, Fig. 11). The paired comparisons showed that the presence of speech sounds increased perceived disturbance as all speech conditions differed significantly from *Quiet* (all U values $\leqslant 16.00$, p values < .001, 1-tailed). Within the speech conditions, *Abs_Mask* resulted in the lowest disturbance with a significant difference to *noAbs_noMask* ($U = 130.00$, $p = .001$, 1-tailed) and a marginal difference to *Abs_noMask* ($U = 209.50$, $p = .062$, 1-tailed). The latter two did not differ from each other ($p > .05$, 2-tailed). In sum, the results were in line with the predictions, as the absence of background speech resulted in the lowest perceived disturbance, and the decrease in STI reduced perceived disturbance between the speech conditions.

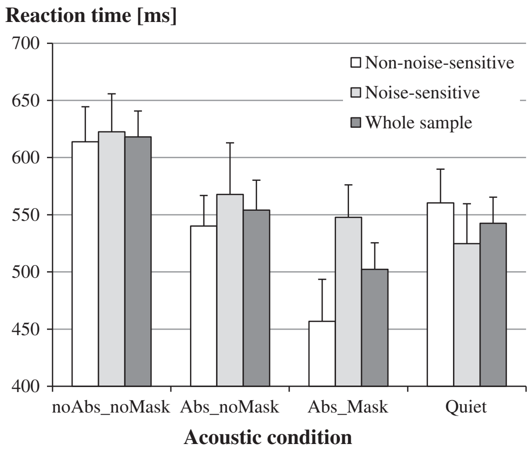

Fig. 8. Performance in the N-back task in the whole sample and the noise sensitivity groups. Means and standard errors.

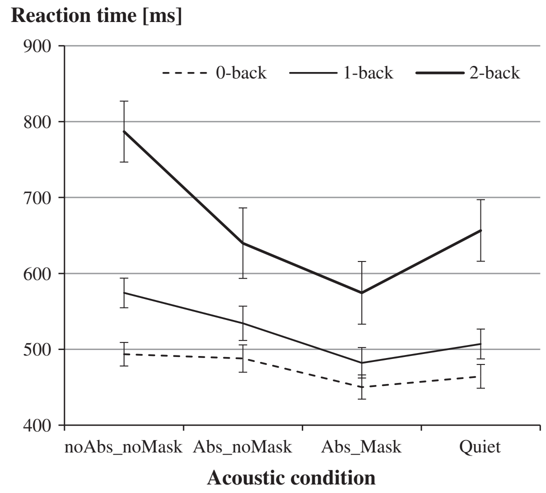

Fig. 9. The interaction between the memory load (0–2) and acoustic condition in the reaction times of the N-back task. Means and standard errors.

A separate analysis of the noise sensitivity groups indicated that acoustic condition had an effect on perceived disturbance in both groups (non-sensitive: χ^2 (3) = 31.13, $p < .001$; sensitive: χ^2 (3) = 28.68, $p < .001$). Paired comparisons showed similar differences as observed with the whole sample, except that *Abs_Mask* resulted in lower perceived disturbance than *Abs_noMask* only in the noise-sensitive group ($U = 170.50$, $p = .019$, 1-tailed).

3.2.2. Perceived distraction

The subjective distraction due to speech sounds was analyzed with the three speech conditions (Figs. 12–14). There was a main effect of acoustic condition for the perceived distraction caused by speech sounds heard from desks further away (χ^2 (2) = 14.30, $p = .001$) but not for speech sounds heard from nearby desks ($p > .05$). That is, the acoustic design did not seem to affect

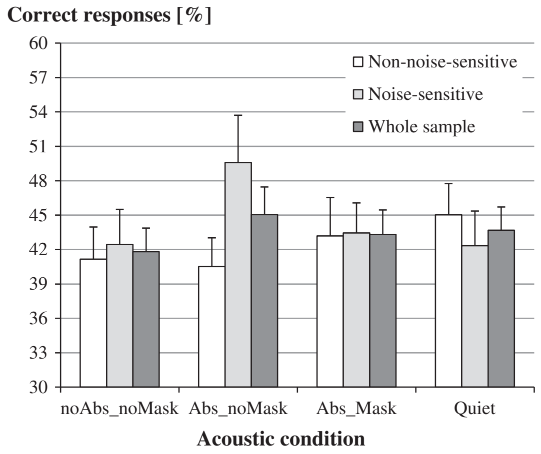

Fig. 10. Performance in the text memory task in the whole sample and the noise sensitivity groups. Adjusted means and standard errors. No statistically significant effects were observed.

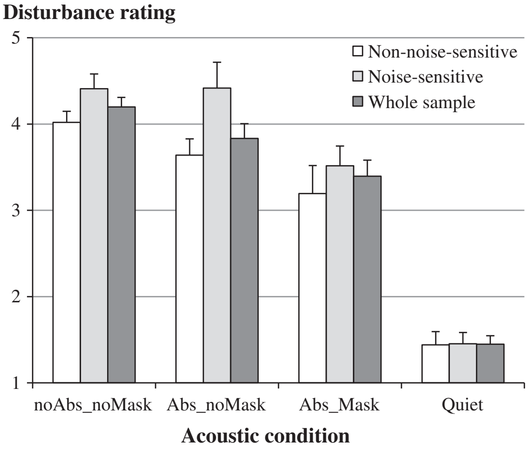

Fig. 11. The sum scores for subjective disturbance in the whole sample and the noise sensitivity groups. Means and standard errors. Scale 1–5 (not at all – very much).

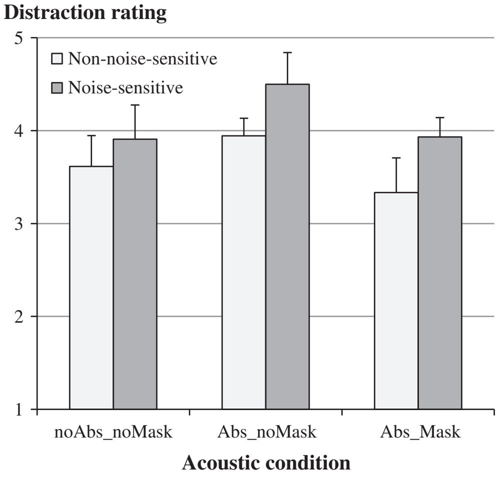

Fig. 13. Perceived distraction caused by speech from nearby desks in the noise sensitivity groups. Means and standard errors. Scale 1–5 (not at all – very much).

subjective distraction for speech that originated from the adjacent desk. Paired comparisons showed that speech heard from desks further away was perceived as less distracting in *Abs_Mask* compared to both *Abs_noMask* ($U = 131.50$, $p = .001$, 1-tailed) and *noAbs_noMask* ($U = 126.00$, $p = .001$, 1-tailed), which supports the assumption of the effect of acoustic design. The latter two conditions did not differ from each other ($p > .05$, 2-tailed).

Separate analyses of the noise sensitivity groups revealed that the effect of speech from further away was highly significant for the noise-sensitive subjects (χ^2 (2) = 13.93, $p = .001$, Fig. 14). Accordingly, the noise-sensitive subjects were less distracted by speech in *Abs_Mask* as opposed to *Abs_noMask* ($U = 4.00$, $p = .002$, 1-tailed) and *noAbs_noMask* ($U = 30.50$, $p = .004$, 1-tailed). For the non-noise-sensitive subjects, the pattern of results was weaker with only a tendency towards a main effect of acoustic condition (χ^2 (2) = 5.17, $p = .076$). The differences between conditions were similar to those observed with noise-sensitive subjects, with the least distraction perceived in *Abs_Mask* as opposed to *Abs_noMask* ($U = 38.00$, $p = .06$, 1-tailed) and *noAbs_noMask* ($U = 25.00$, $p = .03$, 1-tailed).

The distraction caused by the background hum (i.e., ventilation or masking sound) was analyzed with all four conditions. There was no effect of acoustic condition on the perceived distraction due to background hum ($p > .05$, Fig. 12). The results were similarly non-significant when noise-sensitive and non-sensitive groups were separately analyzed.

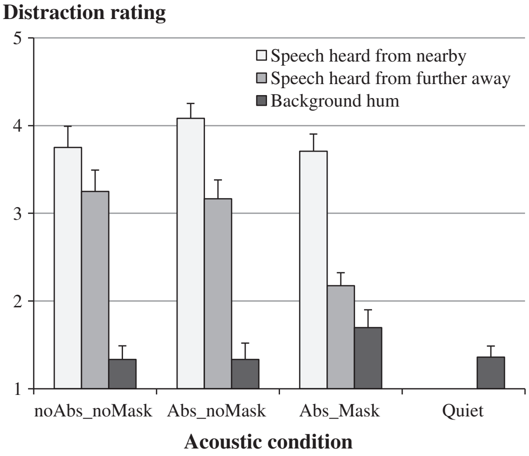

Fig. 12. Perceived distraction caused by speech from different distances for the whole sample. Means and standard errors. Scale 1–5 (not at all – very much).

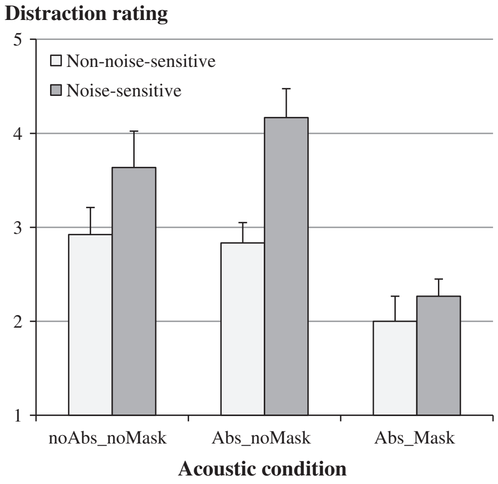

Fig. 14. Perceived distraction caused by speech from desks further away in the noise sensitivity groups. Means and standard errors. Scale 1–5 (not at all – very much).

3.2.3. NASA-TLX

Subjective workload was measured with the NASA-TLX (five items) after every task. Descriptive statistics for the statistically significant findings are shown in Table 3. There were no main effects of acoustic condition on the total NASA-TLX scores for any of the tasks. However, a significant interaction between test block and acoustic condition was observed in the N-back task ($F_{6,130} = 2.54$, $p = .032$, partial $\eta^2 = 0.11$). Separate repeated measures ANOVAs for each acoustic condition showed that the subjective workload decreased with increasing exposure time in *Quiet* ($F_{2,42} = 3.28$, $p = .048$, partial $\eta^2 = 0.14$), whereas no effects of test block were observed in any of the speech conditions. Noise sensitivity also had a small effect on the NASA-TLX scores of the N-back task ($F_{1,65} = 4.21$, $p = .044$, partial $\eta^2 = 0.06$), with noise-sensitive subjects demonstrating higher overall workload than non-sensitive subjects.

Separate ANOVAs for the six subscales in each task yielded only a few significant effects. The subscale for perceived performance

Table 3
Means and standard deviations for the NASA-TLX ratings in different tasks. Scale 0–100; higher ratings indicate higher subjective workload. Only the data for the total scores and subscales that yielded statistically significant main effects of acoustic condition, or interactions involving acoustic condition, are shown.

		Acoustic condition			
		noAbs_noMask	*Abs_noMask*	*Abs_Mask*	*Quiet*
N-back task					
Total score	Test block 1	44.4 (12.3)	42.2 (12.5)	48.6 (12.1)	49.9 (13.1)
	Test block 2	40.3 (14.4)	48.6 (14.6)	44.5 (15.4)	44.7 (13.8)
	Test block 3	47.0 (13.0)	49.5 (15.7)	47.7 (19.3)	43.6 (15.0)
	Noise-sensitive	48.1 (14.6)	55.6 (14.4)	48.9 (16.1)	45.9 (11.4)
	Non-noise-sensitive	40.5 (12.0)	43.1 (12.5)	42.7 (14.6)	46.3 (16.2)
Frustration	Noise-sensitive	60.5 (24.3)	76.3 (18.3)	59.0 (26.9)	56.6 (23.7)
	Non-noise-sensitive	38.9 (18.0)	37.1 (25.3)	51.0 (23.6)	53.2 (26.8)
Operation span task					
Performance	Overall	60.0 (20.1)	63.8 (20.2)	51.6 (21.8)	51.4 (20.8)
Serial recall					
Performance	Overall	66.3 (18.9)	80.9 (13.7)	68.7 (23.0)	64.7 (24.8)
	Test block 1	72.5 (14.9)	89.2 (12.6)	78.6 (21.5)	70.9 (25.3)
	Test block 2	58.0 (21.4)	78.6 (15.8)	66.5 (22.7)	66.0 (26.4)
	Test block 3	68.3 (20.4)	74.9 (12.7)	61.1 (24.7)	57.3 (22.9)
Frustration	Test block 1	68.5 (18.1)	79.4 (27.0)	79.4 (22.9)	79.2 (15.6)
	Test block 2	68.1 (18.0)	70.1 (28.0)	65.5 (24.0)	63.5 (27.3)
	Test block 3	74.7 (14.4)	69.9 (21.2)	58.4 (30.7)	58.7 (17.6)
Text memory, reading					
Frustration	Overall	46.5 (24.9)	27.2 (21.2)	28.3 (23.4)	20.7 (16.0)
Text memory, recall					
Temporal demand	Overall	63.5 (19.1)	53.8 (25.6)	64.0 (24.3)	71.3 (17.1)

was affected by acoustic condition in the operation span task ($F_{3,65} = 3.48$, $p = .021$, partial $\eta^2 = 0.14$) and serial recall ($F_{3,65} = 3.19$, $p = .029$, partial $\eta^2 = 0.13$). In both tasks, the subjects in *Quiet* and *Abs_Mask* perceived their performance as better than subjects in *Abs_noMask* (both *ps* < .05, 1-tailed). In the text memory task, there was an effect of acoustic condition on frustration during reading ($F_{3,65} = 4.05$, $p = .011$, partial $\eta^2 = 0.16$). Again, the lowest frustration was observed in *Quiet* and *Abs_Mask* compared to *Abs_noMask* (both *p* values < .05, 1-tailed). There was also an effect of acoustic condition on the temporal demand during text recall ($F_{3,65} = 3.23$, $p = .028$, partial $\eta^2 = 0.13$) but the effect was in the opposite direction to the prediction with higher demand perceived in *Quiet* as opposed to *Abs_noMask* ($p = .033$, 1-tailed).

There was a marginal interaction between test block and acoustic condition in frustration in serial recall ($F_{6,130} = 2.11$, $p = .056$, partial $\eta^2 = 0.09$). Separate repeated measures ANOVAs for each acoustic condition showed that frustration decreased with increasing exposure time in *Quiet* ($F_{2,46} = 10.65$, $p = .001$, partial $\eta^2 = 0.32$) and *Abs_Mask* ($F_{2,44} = 7.29$, $p = .004$, partial $\eta^2 = 0.25$), whereas no effects of test block were observed in the other two conditions.

There were few main effects of noise sensitivity on the subscales. The only significant interaction between noise sensitivity and acoustic condition was observed for frustration in the N-back task ($F_{3,65} = 2.83$, $p = .045$, partial $\eta^2 = 0.12$). Paired comparisons showed that noise-sensitive subjects experienced higher frustration than the non-sensitive subjects in *Abs_noMask* ($p < .001$, 1-tailed) with a similar trend also emerging in *noAbs_noMask* ($p = .085$, 1-tailed). Noise sensitivity did not affect frustration in *Abs_Mask* or *Quiet.*

4. Discussion

The primary aim of the present study was to test whether the detrimental effects of unattended background speech (hereinafter referred to as "background speech") could be reduced by room acoustic design that decreases the STI of speech. In line with previous research, the results showed that the presence of background speech has detrimental effects on short-term memory and working memory performance and on the subjective perceptions of noise effects. These effects were neither attenuated nor amplified within a three-hour working period. The results also showed that a reduction of STI by room acoustic means decreases subjective disturbance. However, the support for the effect of acoustic design on cognitive performance was somewhat weaker than expected. The results also suggest that noise-sensitive individuals may subjectively benefit more from acoustic improvements than less-sensitive individuals as the effects of the room acoustic design on subjective distraction were stronger in the noise-sensitive group than in the non-sensitive group.

4.1. The effect of room acoustic design on cognitive performance

In terms of cognitive performance effects, the results largely replicate the previous findings that the presence of background speech is detrimental to cognitive performance (e.g., [13,21,22]). This effect was observed in serial recall and two working memory tasks. However, the effect was mainly observed between the *Quiet* condition and either or both of the conditions with the highest STIs (*noAbs_noMask* and *Abs_noMask*) but not between *Quiet* and the condition with the lowest STI (*Abs_Mask*). The present results are also compatible with the literature showing that the effects of speech are task-specific (e.g., [14]). This was shown in the N-back task, which demonstrated that the deleterious effect of background speech on performance increased as a function of the working memory load.

The presence of background speech did not affect performance in the text memory task that required a more complex set of cognitive processes, such as reading comprehension and long-term memory. The explanation for the differences between this and the other tasks probably derives from different task demands. The text memory task is essentially a semantic task that requires the processing of meaning, whereas the other tasks rely more on attention, memory for order, and working memory processes. Other studies have similarly reported no or only weak effects of speech intelligibility on semantic distraction (e.g., [13–15,17]). It is currently not clear why semantic tasks would be less affected by speech intelligibility than other tasks, particularly as the existence of semantic interference by background speech as such has been demonstrated by several studies in recent years (e.g., [25,26]). This issue is clearly an important topic for future studies.

The most important findings of the present study concern the effects of acoustic design on the interference caused by background speech. In terms of cognitive performance, these effects were somewhat weaker than expected. One would assume that changes in STI, introduced by room acoustic design, would have improved performance in the three tasks that were affected by speech sounds *per se*. However, the acoustic conditions only affected performance in the N-back task in which the RTs indicated decreased cognitive load in the optimal acoustic condition (*Abs_Mask*) compared to the condition where the acoustic design was inadequate (*noAbs_noMask*). A marginal tendency toward the effect of acoustic design was observed in the operation span task and no effect in serial recall. These findings are somewhat surprising, as previous research has demonstrated a relation between STI and performance in both the serial recall [12,13] and the operation span tasks [13].

There are a few possible explanations for this inconsistency. Firstly, the present results may reflect a more accurate effect of the STI on performance in terms of ecological validity because the experimental conditions and speech materials were carefully designed from a practically motivated, rather than a theoretical, point of view. To be more specific, the present study incorporated variation in STI, and more importantly, included the physical limitations that are associated with reducing STI at short distances between a speaker and a listener. Along with earlier research [33], the STI measurements of the present study (see Table 2 and the spatial decay curves in Appendix A) demonstrate that the effect of acoustic design becomes more pronounced with increasing distance from the speech source. This effect was also reflected in the subjective responses, which showed that the acoustic condition had an effect on perceived distraction of speech from further away but not for speech from nearby. In fact, the STI of speech from the adjacent desk remained relatively high exceeding 0.50 in all conditions, placing all the subjects within a distraction distance from at least one speech source. According to previous studies, STI should be below 0.50 in order for performance to improve [12–14,32,34]. The present results demonstrate that open-plan offices similar to the condition *Abs_Mask* are not accurately characterized by a low STI but rather with a smaller probability of exposure to STIs exceeding 0.50. Therefore, the effect of acoustic improvements on performance will, in practice, be smaller than has been suggested by previous studies where the limitations related to distance have not been considered. To further emphasize this point, it should be noted that the optimal acoustic condition (*Abs_Mask*) of the present study was very effective in reducing STI because a very low STI was obtained at only a 6 m distance from a speech source (Table 2). Thus, the somewhat limited findings on the effect of acoustic design on performance were not due to inadequacies in the execution of acoustic conditions, but they rather represent the practical limitations of room acoustic design.

A related explanation for the rather small performance effects is based on the possibility of increased attentional capture by intelligible speech in the optimal acoustic condition (*Abs_Mask*). The idea of change, or variation, in acoustic perception is seen as an essential factor in several accounts of noise-related performance effects (e.g., [19,27,59]). For instance, *a duplex-mechanism account* [60] holds that performance impairment is either caused by a conflict of overlapping processes or by attentional capture caused by an unexpected acoustic event. Adopting this framework, one may speculate that intelligible speech had more detrimental impact in the best acoustic condition (*Abs_Mask*) than in the other conditions because the nearby speech was likely to be perceptually salient in an otherwise reasonably peaceful acoustic environment. In the other speech conditions, the variation in the acoustic perception was probably more predictable as all speech sounds were at least moderately intelligible.

As a second line of explanation, we need to consider methodological issues as a possible source for the limited number of findings on cognitive performance. The practical execution of the conditions required the acoustic conditions to be manipulated as a between-groups variable which introduced individual differences as a source of experimental error. Individual differences in intelligence *per se* will be reflected in any test of cognitive ability (see [39]), while the susceptibility to background speech also has considerable individual variation, at least in serial recall [61]. This individual variation might mask the effects of acoustic conditions particularly if the effects of acoustic manipulation are expected to be small in relation to those produced by individual differences. While WM capacity and noise sensitivity were included in the analyses to account for some of the variance, it is important to note that this procedure does not correct for any pre-existing differences between groups nor remove the effect of individual differences on the results. Thus, some statistical power to detect differences was probably lost by the necessity of manipulating the acoustic conditions between groups. This might explain why only the differences between the extreme acoustic conditions were most likely to reach statistical significance.

To further clarify the effect of the between-groups design on the findings, it is worth comparing the present results with a related study by Varjo et al. [62]. The study was conducted in the same laboratory with a within-subject manipulation of two conditions that had acoustic properties similar to those of the conditions *Abs_Mask* and *noAbs_noMask* of the present study. The acoustic conditions were combined with changes in the temperature and ventilation rate but, as separate experiments showed that the temperature and ventilation rate did not affect the cognitive tasks that were used [41,63], the effects obtained by Varjo et al. [62] can be mainly attributed to the decrease in STI. They demonstrated significant effects in the serial recall, operation span, and N-back task. These findings support the assumption that an optimal acoustic design similar to the present study can decrease the cognitive disruptions caused by background speech. The differences in the experimental design are also consistent with the view that the lack of similar findings in the present study resulted, at least to some degree, from the between-groups manipulation of the acoustic conditions.

4.2. The effect of room acoustic design on subjective perceptions

The data on subjective perceptions provided stronger evidence in support of the assumed effects of acoustic design on distraction by speech. Especially perceived disturbance in general and distraction caused by speech from further away were sensitive to the differences between the acoustic conditions. The results showed that perceived distraction decreased in the optimal acoustic design (*Abs_Mask*) as opposed to the conditions with the highest STIs (*noAbs_noMask* and *Abs_noMask*), while best subjective ratings were obtained when background speech was absent.

The pattern of results is in line with other studies that have demonstrated that subjective measures tend to be more sensitive to differences between acoustic conditions than measures of objective performance (e.g., [12,13,22]). An often-mentioned explanation for the differences between performance and questionnaire measures is offered by the enhanced effort hypothesis [12,13,22]. According to this view, subjects recognize the acoustic distraction as harmful for their performance and compensate for the anticipated deleterious effects by investing more effort in the task. This is assumed to diminish the observable effects on performance while amplifying the negative experience of the acoustic environment. The role of compensatory resources and effort has also seen as an essential mediator between noise and performance effects by other researchers [19].

At first glance, the current results may not seem compatible with this line of explanation because the subjective workload (as measured by the NASA-TLX) did not considerably increase in the speech conditions. However, the NASA-TLX may not have been an appropriate measure for subjective workload in the present study. Firstly, the NASA-TLX questionnaire was completed immediately after each task and the instructions focused on assessing the perceived workload in the task, not the effect of noise on task performance. Thus, the ratings probably mostly reflect the demands of different task characteristics, particularly as the presence of several cognitive tasks may have led subjects to use other tasks as a point of comparison. Secondly, the large individual variation observed in the ratings (Table 3) suggests that individual differences in response style probably decreased the sensitivity of the NASA-TLX due to the between-groups design. In other studies of speech intelligibility, the only other study that has used the NASA-TLX for assessing the subjective workload also reported a few small effects [15], whereas studies that have specifically focused on assessing the perceived performance effect of background speech (e.g., [12]) have typically reported larger effects. Thus, the sum scores of perceived disturbance seem more appropriate for describing the perceived workload also in the present study because the included questions were mainly related to performance, such as effects on concentration and adaptation. The pattern of results on subjective responses and cognitive performance is, therefore, compatible with the explanation that the noise effects may have been compensated by subjects, leading to fewer observable performance differences between the acoustic conditions. However, while this explanation has been proposed by several researchers [12,13,22], it has not yet been specifically tested in relation to speech intelligibility and performance. Including the moderating effect of compensatory effort and adaptation more explicitly in future studies would help to clarify this issue.

4.3. The effect of the exposure time

A secondary aim of the present study was to assess whether the effect of acoustic condition would be moderated by the exposure time, which has not been studied in speech intelligibility research before. Whereas learning effects were observed across experimental blocks in all tasks, there were no interactions between acoustic conditions and the exposure time in the cognitive measures. In other words, the effects of background speech on performance were not attenuated nor amplified with increasing exposure time. The lack of attenuation is in line with the findings that there does not seem to be habituation to background speech, at least in serial recall [64]. While the present study was not designed to specifically test habituation, the results suggest that an absence of habituation might extend to more complex working memory tasks.

The few interactions observed in the NASA-TLX provide modest support for the alternative hypothesis that some effects of noise might be amplified with increasing exposure time. This was seen as a decrease in the subjective workload in *Quiet*, whereas no change in relation to the exposure time was observed in the conditions containing background speech. In *Quiet*, the decline in the workload most likely reflects decreasing task demand resulting from task repetition and learning. The absence of this effect in speech conditions does not, however, indicate impaired learning because the performance data demonstrate the opposite. The effect may rather represent an increase in compensatory effort or in stress response resulting from continuing exposure to background speech.

The effect of the exposure time is an important topic for future studies, both regarding the short-term and long-term effects. This issue is also related to the moderating effects of compensatory effort discussed in the previous section, as the degree to which noise effects can be adapted to would influence the long-term effects on work performance in a wider sense.

4.4. Noise sensitivity and acoustic conditions

The present results also contribute to the research on noise sensitivity by showing that noise-sensitive subjects are more affected by background speech than non-sensitive individuals, both in terms of cognitive performance and the subjective effects of noise. Whereas the relation between noise sensitivity and subjective responses, such as annoyance, has been well-documented (e.g., [65]), there has been less research conducted on the relation between noise sensitivity and performance (see e.g., [40]). Within the speech intelligibility research, the present study is the third to include an investigation of noise sensitivity and the first to suggest some moderating effects in relation to acoustic conditions. The difference between the present findings and those of the previous studies including noise sensitivity [13,17] is probably accounted for by the choice of noise sensitivity measures. The previous studies used items from Weinstein's [66] noise sensitivity scale, which has been considered to be too general to be substantially related to specific measures of performance [61]. In the present study, we employed a more specific measure from NoiSeQ [56] that focuses on perceived susceptibility to performance effects due to noise. Apparently, this measure was more sensitive and appropriate for the present research design and could be also utilized in similar studies in the future.

The present results also suggest that noise-sensitive individuals might benefit more from acoustic improvements, at least in terms of subjectively perceived distraction caused by speech. This was demonstrated by the stronger effects of acoustic condition among the noise-sensitive than the non-sensitive individuals. However, the effect of noise sensitivity was a secondary question in the present study and, to simplify the interpretation, it was described with a rough categorical variable. Future studies could explore this issue further using more detailed analyses of the effect of individual noise sensitivity.

4.5. Limitations of the present study

Both the strengths and weaknesses of the present study are related to the practically motivated research design. As a result of the physically-built experimental conditions, two methodological compromises occurred, namely the between-groups manipulation of the acoustic conditions and the lack of random assignment of subjects to the experimental conditions. As already discussed above, the between-groups manipulation may have weakened, to some extent, the statistical power of the experiment to detect expected differences between the acoustic conditions.

Due to the long time frame of the study, subjects were acquired for one condition at a time and, therefore, could not be randomly assigned to experimental conditions. This resulted in an unfortunate unbalance in gender distribution across the conditions as there were fewer male volunteers available toward the end of the study. Although this issue weakened the analysis of subjective workload (the analyses were only performed for females), we have no reason to believe that there were gender effects on other measures used in the study. Even if that had been the case, this would have only affected the results obtained in the condition with no acoustic treatment (*noAbs_noMask*), as the proportion of males was not substantially different in the other three conditions. The lack of random assignment may have, in theory, introduced other unknown biases but none of the collected background data (i.e. age, noise sensitivity, WM capacity) suggest that.

The choice of presenting speech from only one speaker at a time might seem to limit the practical validity of the study design as in most open-plan offices occupants are more likely to hear several simultaneous voices. There is indeed some data to suggest that the babble caused by simultaneous speakers has less effect on performance as individual voices become undistinguishable [67]. However, this effect is only observed if the voices come from one location, whereas the disruption is restored when speakers have different locations [67], as they would have in an open-plan office. Thus, simultaneous speech sounds would only be beneficial in exceptional cases, such as in very large open-plan offices where a large number of workers are continuously engaged in conversation or phone-calls (e.g., call centers). Hence, the use of only one speaker at a time seems justified in the present study.

4.6. Practical implications and conclusions

The present results show that, with an optimal acoustic design, the intelligibility of background speech and the associated distraction can be effectively reduced when the speaker is at least four-to-six meters away from the listener. Achieving such conditions requires the simultaneous use of high room absorption, high screens, and artificial speech masking sound. According to the present results, a recommended masking sound level of 45 dB L_{Aeq} is not in itself perceived as a distraction, even among the more noise-sensitive individuals. The present results also clearly demonstrate that the mere addition of high room absorption and high screens does not affect subjective distraction or cognitive performance if the masking sound level is insufficient. However, the STI cannot be reduced enough by room acoustic means to decrease distraction caused by speech from the nearest desks (two-to-three meters away). This conclusion also applies to a situation where occupants are already slightly lowering their voice levels as this effect was taken into account in the present experimental design.

The limitations of reducing STI at close distances indicate that the acoustic problems resulting from unwanted speech in open-plan offices cannot be solved by room acoustic design alone. There are several additional ways to reduce distraction and enhance possibilities for concentration. Attention should be paid to desk density as higher density will increase the number of speech sources nearby, i.e., within the distraction distance, thus weakening the effectiveness of the existing room acoustic means. For comparison, the average seat-to-seat distance was 2.5 m in the present study. Seating arrangements may also be important as team members or workers with similar tasks are more likely not to disturb each other than workers with unrelated work contents. Another commonly applied practice is the use of a behavioral etiquette to reduce the degree of distraction and interruption in an open-plan office. Most importantly, the possibilities for concentration can be improved by providing anonymous rooms within easy reach where working phases with higher concentration demands, telemeetings, phone calls, and *ad hoc* meetings can be carried out. A peaceful working space should be provided particularly for individuals whose work involves high demands on short-term and working memory and for individuals who are noise-sensitive.

The standardized room acoustic features of the three acoustic conditions in the present study were also measured according to the new standard ISO 3382-3 [20] (see Appendix A). The results of the present study support the use of the standard in the future. Especially, the use of the distraction distance as a primary indicator of room acoustic quality is encouraged as it is closely associated with speech intelligibility and the distraction caused by background speech.

Table A1
The single-number quantities of ISO 3382-3 in the three acoustic conditions that included speech in the background.

Acoustic condition	$D_{2,S}$ (dB)	r_D (m)
noAbs_noMask	1.4	38
Abs_noMask	8.5	11
Abs_Mask	8.5	3.4

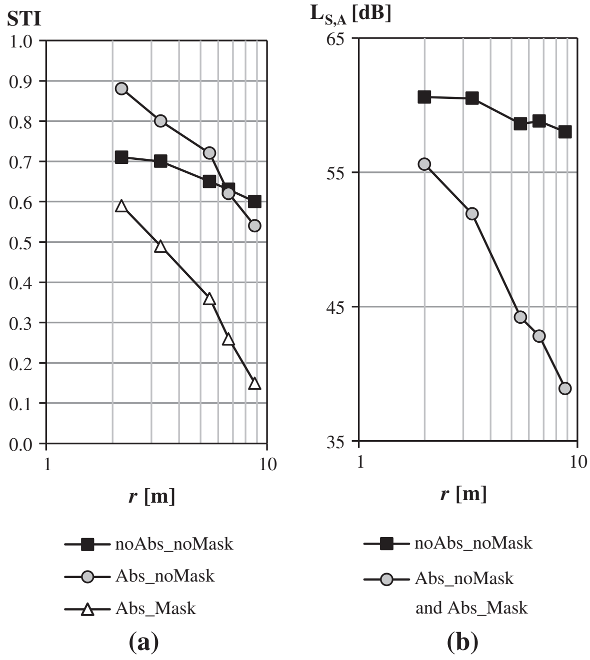

Fig. A1. Spatial decay of (a) Speech Transmission Index, STI, and (b) A-weighted equivalent speech level, $L_{S,A}$, in the three acoustic conditions containing speech. The distance to the speaker is r.

In sum, the present results show that reducing the distraction distance of an office improves the conditions particularly in terms of subjective perceptions and distraction, but also in terms of cognitive performance. However, the best conditions for these outcomes are achieved when background speech is completely absent, which supports the use of private rooms when the cognitive demands of work tasks are high.

Acknowledgements

This study was a part of the TOTI research program (User-Oriented Office Spaces 2009–2012) funded by Tekes and 15 companies. We would like to thank David Oliva and Jarkko Hakala for their help in the acoustic measurements and in the building of the acoustic conditions.

Appendix A. ISO 3382-3 measurements

The single-number quantities, spatial decay rate of speech, $D_{2,S}$ and distraction distance, r_D, that resulted from the ISO 3382-3:2012 [20] measurements are shown in Table A1 and Fig. A1. The spatial decay rate of A-weighted speech, $D_{2,S}$, describes how many decibels the speech level is reduced when the distance to a speaker is doubled. The distraction distance r_D describes the distance within which the value of STI remains above 0.50. In order to have good speech privacy, r_D should be small. $D_{2,S}$ alone does not indicate speech privacy because it does not take masking sound into account. However, this quantity describes how well the room itself (including the furniture) attenuates sound. It must be emphasized that these measurements were made using standard speech effort and an omni-directional speaker which were not used in the experiment. The perceived conditions of the experiment are described in Table 2 and Fig. 5. Therefore, the results of this Appendix A present supportive data to understand how the experimental conditions related to typical conditions found in workplaces (see e.g., the measurements by Virjonen et al. [33]), and how the STI and speech levels behaved at different distances from a speaker.

References

[1] Pejtersen J, Allermann L, Kristensen TS, Poulsen OM. Indoor climate, psychosocial work environment and symptoms in open-plan offices. Indoor Air 2006;16:392–401.

[2] Haapakangas A, Helenius R, Keskinen E, Hongisto V. Perceived acoustic environment, work performance and well-being–survey results from Finnish offices. In: 9th International congress on noise as a public health problem (ICBEN); July 2008. p. 21–5.

[3] Sundstrom E, Town JP, Rice RW, Osborn DP, Brill M. Office noise, satisfaction, and performance. Environ Behav 1994;26:195–222.

[4] Banbury SP, Berry DC. Office noise and employee concentration: identifying causes of disruption and potential improvements. Ergonomics 2005;48:25–37.

[5] Kaarlela-Tuomaala A, Helenius R, Keskinen E, Hongisto V. Effects of acoustic environment on work in private office rooms and open-plan offices–longitudinal study during relocation. Ergonomics 2009;52:1423–44.

[6] Kim J, de Dear R. Workspace satisfaction: the privacy-communication trade-off in open-plan offices. J Environ Psychol 2013;36:18–26.

[7] de Croon E, Sluiter J, Kuijer PP, Frings-Dresen M. The effect of office concepts on worker health and performance: a systematic review of the literature. Ergonomics 2005;48:119–34.

[8] Pejtersen JH, Feveile H, Christensen KB, Burr H. Sickness absence associated with shared and open-plan offices—a national cross sectional questionnaire survey. Scand J Work Environ Health 2011;37:376–82.

[9] Evans QW, Johnson D. Stress and open-office noise. J Appl Psychol 2000;85:779–83.

[10] Jahncke H, Hygge S, Halin N, Green AM, Dimberg K. Open-plan office noise: cognitive performance and restoration. J Environ Psychol 2011;31:373–82.

[11] Witterseh T, Wyon DP, Clausen G. The effects of moderate heat stress and open-plan office noise distraction on SBS symptoms and on the performance of office work. Indoor Air 2004;14:30–40.

[12] Haapakangas A, Kankkunen E, Hongisto V, Virjonen P, Oliva D, Keskinen E. Effects of five speech masking sounds on performance and acoustic satisfaction. Implications for open-plan offices. Acta Acust United Acust 2011;97:641–55.

[13] Haka M, Haapakangas A, Keränen J, Hakala J, Keskinen E, Hongisto V. Performance effects and subjective disturbance of speech in acoustically different office types – a laboratory experiment. Indoor Air 2009;19:454–67.

[14] Jahncke H, Hongisto V, Virjonen P. Cognitive performance during irrelevant speech: effects of speech intelligibility and office-task characteristics. Appl Acoust 2013;74:307–16.

[15] Liebl A, Haller J, Jodicke B, Baumgartner H, Schlittmeier S, Hellbrück J. Combined effects of acoustic and visual distraction on cognitive performance and well-being. Appl Ergon 2012;43:424–34.

[16] Smith-Jackson TL, Klein KW. Open-plan offices: task performance and mental workload. J Environ Psychol 2009;29:279–89.

[17] Venetjoki N, Kaarlela-Tuomaala A, Keskinen E, Hongisto V. The effect of speech and speech intelligibility on task performance. Ergonomics 2006;49:1068–91.

[18] Tang SK, Wong CT. Performance of noise indices in office environment dominated by noise from human speech. Appl Acoust 1998;55:293–305.

[19] Szalma JL, Hancock PA. Noise effects on human performance: a meta-analytic synthesis. Psychol Bull 2011;137:682–707.

[20] ISO 3382-3:2012. Acoustics – measurement of room acoustic parameters – Part 3: open-plan offices. Geneva, Switzerland: International Organization for Standardization; 2012.

[21] Salamé P, Baddeley A. Noise, unattended speech and short-term memory. Ergonomics 1987;30:1185–94.

[22] Schlittmeier SJ, Hellbrück J, Thaden R, Vorländer M. The impact of background speech varying in intelligibility: effects on cognitive performance and perceived disturbance. Ergonomics 2008;51:719–36.

[23] Martin RC, Wogalter MS, Forlano JG. Reading comprehension in the presence of unattended speech and music. J Memory Lang 1988;27:382–98.

[24] Jones DM, Miles C, Page J. Disruption of proofreading by irrelevant speech: effects of attention, arousal or memory? Appl Cogn Psychol 1990;4:89–108.

[25] Sörqvist P, Nöstl A, Halin N. Disruption of writing processes by the semanticity of background speech. Scand J Psychol 2012;53:97–102.

[26] Marsh JE, Hughes RW, Jones DM. Interference by process, not content, determines semantic auditory distraction. Cognition 2009;110:23–38.

[27] Jones D, Madden C, Miles C. Privileged access by irrelevant speech to short-term memory: the role of changing state. Q J Exp Psychol 1992;44:645–69.

[28] Colle HA. Auditory encoding in visual short-term recall: effects of noise intensity and spatial location. J Verb Learn Verb Behav 1980;19:722–35.

[29] Ellermeier W, Hellbrück J. Is level irrelevant in” irrelevant speech”? and binaural unmasking. J Exp Psychol: Hum Percept Perform 1998;24:1406–14.

[30] Veitch JA, Bradley JS, Legault LM, Norcross S, Svec JM. Masking speech in open-plan offices with simulated ventilation noise: noise level and spectral composition effects on acoustic satisfaction. In: Institute for research in construction, Internal, Report IRC-IR-846; 2002.

[31] Schlittmeier SJ, Hellbrück J. Background music as noise abatement in open-plan offices: a laboratory study on performance effects and subjective preferences. Appl Cogn Psychol 2009;23:684–97.

[32] Keus van de Poll M, Ljung R, Odelius J, Sörqvist P. Disruption of writing by background speech: the role of speech transmission index. Appl Acoust 2014;81:15–8.

[33] Virjonen P, Keränen J, Hongisto V. Determination of acoustical conditions in open-plan offices – proposal for new measurement method and target values. Acta Acust United Acust 2009;95:279–90.

[34] Hongisto V. A model predicting the effect of speech of varying intelligibility on work performance. Indoor Air 2005;15:458–68.

[35] Emberson LL, Lupyan G, Goldstein MH, Spivey MJ. Overheard cell-phone conversations when less speech is more distracting. Psychol Sci 2010;21:1383–8.

[36] Banbury S, Berry DC. Habituation and dishabituation to speech and office noise. J Exp Psychol: Appl 1997;3:181–95.

[37] Buchner A, Bell R, Rothermund K, Wentura D. Sound source location modulates the irrelevant-sound effect. Memory Cogn 2008;36:617–28.

[38] Spence C, Ranson J, Driver J. Cross-modal selective attention: on the difficulty of ignoring sounds at the locus of visual attention. Percept Psychophys 2000;62:410–24.

[39] Colom R, Rebollo I, Palacios A, Juan-Espinosa M, Kyllonen PC. Working memory is (almost) perfectly predicted by g. Intelligence 2004;32:277–96.

[40] Belojevic G, Jakovljevic B, Slepcevic V. Noise and mental performance: personality attributes and noise sensitivity. Noise and Health 2003;6:77–89.

[41] Maula H, Haapakangas A, Östman L, Hongisto V, Koskela H, Hyönä J. The effect of 29 °C room air temperature on work performance and comfort in office environment – laboratory experiment with wider range of cognitive tasks. Healthy Buildings, Brisbane; 8–12 July, 2012.

[42] EN ISO 11654:1997. Acoustics – sound absorbers for use in buildings – rating of sound absorption; 1997.

[43] Hongisto V. Effect of sound masking on workers - a longitudinal study in an open office. In: Acoustics ‘08, paper 1178, Paris, France; June 29–July 4, 2008.

[44] Hongisto V, Haapakangas A, Helenius R, Keränen J, Oliva D. Acoustic satisfaction in an open-plan office before and after the renovation. In: Euronoise, Prague, Czech Republic; June 10–13, 2012. p. 654–9.

[45] SFS 5907:2004. Acoustic classification of spaces in buildings. Helsinki, Finland: Finnish Standards Association SFS; 2004 [Available in English].

[46] Haapakangas A, Hongisto V, Kokko J, Oliva D, Keränen J, Hakala J, et al. Room acoustics and work performance – experimental study in a full-scale open-plan office laboratory. In: 10th International congress on noise as a public health problem (ICBEN), London, UK; 24–28 July, 2011.

[47] ISO 3741:2010. Acoustics – determination of sound power levels and sound energy levels of noise sources using sound pressure – precision methods for reverberation test rooms. Geneva, Switzerland: International Organization for Standardization; 2010.

[48] Warnock ACC, Chu WT. Voice and background noise levels measured in open offices. In: IRC-IR-837, National Research Council Canada, Institute for Research in Construction, Ottawa, Canada; January 2002.

[49] Keränen J, Hongisto V, Hakala J, Oliva D. Speech privacy in an open-plan office with different room acoustic conditions. In: Forum Acusticum 2011, paper 135, Aalborg, Denmark; 27 June–1 July, 2011. p. 1897–900.

[50] Keränen J, Hongisto V, Oliva D, Hakala J. The effect of different room acoustic elements on spatial decay of speech – a laboratory experiment. In: Euronoise, Prague, Czech Republic; June 10–13, 2012; p. 624–9.

[51] Turner ML, Engle RW. Is working memory capacity task dependent? J Memory Lang 1989;28:127–54.

[52] Conway AR, Kane MJ, Bunting MF, Hambrick DZ, Wilhelm O, Engle RW. Working memory span tasks: A methodological review and user’s guide. Psychon B Rev 2005;12:769–86.

[53] Owen AM, McMillan KM, Laird AR, Bullmore E. N-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies. Hum Brain Mapp 2005;25:46–59.

[54] Kaakinen JK, Hyönä J, Keenan JM. How prior knowledge, expository text. J Exp Psychol: Learn Memory Cogn 2003;29:447–57.

[55] Griefahn B, Marks A, Schreckenberg D. Design, reliability and validity of a short version of the noise sensitivity questionnaire. Unpublished manuscript.

[56] Schutte M, Marks A, Wenning E, Griefahn B. The development of the noise sensitivity questionnaire. Noise and Health 2007;9:15–24.

[57] Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. Hum Mental Workload 1988;1:139–83.

[58] Hancock PA, Vercruyssen M, Rodenburg GJ. The effect of gender and time-of-day on time perception and mental workload. Curr Psychol 1992;11:203–25.

[59] Cowan N. Attention and memory. Oxford University Press; 1995.

[60] Hughes RW, Vachon F, Jones DM. Disruption of short-term memory by changing and deviant sounds: support for a duplex-mechanism account of auditory distraction. J Exp Psychol: Learn Memory Cogn 2007;33:1050–61.

[61] Ellermeier W, Zimmer K. Individual differences in susceptibility to the “irrelevant speech effect”. J Acoust Soc Am 1997;102:2191–9.

[62] Varjo J, Hongisto V, Haapakangas A, Maula H, Koskela H, Hyönä J. The effect of irrelevant speech, high temperature and low ventilation rate on cognitive performance and environmental satisfaction in open-plan offices. J Environ Psychol February 2014; Submitted for publication.

[63] Koskela H, Maula H, Haapakangas A, Moberg V, Hongisto V. Effect of low ventilation rate on office work performance and perception of air quality – a laboratory study. In: Extended abstract, Submitted to Indoor Air 2014, Hong Kong; July 7–12, 2014.

[64] Banbury SP, Macken WJ, Tremblay S, Jones DM. Auditory distraction and short-term memory: phenomena and practical implications. Hum Factors: J Hum Factors Ergon Soc 2001;43:12–29.

[65] van Kamp I, Job RS, Hatfield J, Haines M, Stellato RK, Stansfeld SA. The role of noise sensitivity in the noise–response relation: a comparison of three international airport studies. J Acoust Soc Am 2004;116:3471–9.

[66] Weinstein ND. Individual differences in reactions to noise: a longitudinal study in a college dormitory. J Appl Psychol 1978;63:458–66.

[67] Jones DM, Macken WJ. Auditory babble and cognitive efficiency: role of number of voices and their location. J Exp Psychol: Appl 1995;1:216–26.