Journal of Environmental Psychology 44 (2015) 16–33

Contents lists available at ScienceDirect

Journal of Environmental Psychology

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

Simultaneous effects of irrelevant speech, temperature and ventilation rate on performance and satisfaction in open-plan offices☆

Johanna Varjo [a,*], Valtteri Hongisto [a], Annu Haapakangas [a], Henna Maula [a], Hannu Koskela [a], Jukka Hyönä [b]

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

ARTICLE INFO

Article history:
Received 19 February 2014
Received in revised form
24 March 2015
Accepted 19 August 2015
Available online 28 August 2015

Keywords:
Cognitive performance
Indoor environment
Noise
Air quality
Room temperature
Open-plan offices

ABSTRACT

The aim of this study was to investigate how irrelevant speech, temperature and ventilation rate together affect cognitive performance and environmental satisfaction in open-plan offices. In Condition A, neutral temperature (23.5 °C), low intelligibility of speech (high absorption and low masking sound level) and high fresh air supply rate (30 l/s per person) were applied. This was contrasted to Condition B with high room temperature (29.5 °C), highly intelligible speech (low absorption and high masking sound level) and a negligible fresh air supply rate (2 l/s per person). Sixty-five participants were tested. In Condition B, performance decrement was observed especially in working memory tasks. Based on subjective assessments, mental workload, cognitive fatigue and symptoms were higher and environmental satisfaction was lower in Condition B. It was concluded that special attention should be paid to the design of whole indoor environment in open-plan offices to increase subjective comfort and improve performance.

1. Introduction

Scientific interest towards subjective satisfaction in open-plan offices has increased because open-plan office has become the most usual office solution, mostly because of its high space efficiency (De Croon, Sluiter, Kuijer, & Frings-Dresen, 2005). Moreover, open-plan offices are also assumed to improve organizational productivity due to the enhanced exchange of information and communication and increased teamwork (Allen & Gerstberger, 1973; Hundert & Greenfield, 1969).

However, many studies have shown that there are disadvantages in open-plan offices if the design of the indoor environment (IE) is inadequate. Increased cognitive workload (De Croon et al., 2005), concentration problems and fatigue (Haapakangas, Helenius, Keskinen, & Hongisto, 2008; Pejtersen, Allermann, Kristensen, & Poulsen, 2006) and the lack of speech privacy (De Croon et al., 2005) have been reported. Open-plan offices have also been associated with decreased job satisfaction (De Croon et al., 2005). Decreased satisfaction with IE has been indicated to have a connection with decreased job satisfaction (Veitch, Charles, Farley, & Newsham, 2007). The amount of annual sick leave has also been shown to be greater in open-plan offices, as assessed by employees' self-ratings (Bodin Danielsson, Chungkham, Wulff, & Westerlund, 2014; Pejtersen, Feveile, Christensen, & Burr, 2011).

One of the most commonly mentioned causes for these problems is poor acoustic conditions, i.e., disturbance caused by colleagues' speech and poor speech privacy (Danielsson, 2005; Frontczak et al., 2012; Haapakangas et al., 2008; Pejtersen et al., 2006). Improper thermal conditions and poor air quality have also been reported as producing discomfort in open-plan offices (Haapakangas et al., 2008; Pejtersen et al., 2006). On the other hand, overall improvement of the IE can significantly increase environmental satisfaction in open-plan offices (Hongisto, Haapakangas, Helenius, Keränen, & Oliva, 2012). That is, differences between open-plan offices can be significant regarding on the quality of IE.

The effects of IE on work performance and various components of environmental satisfaction have been studied in several

☆ This manuscript was submitted for publication to the Journal of Environmental Psychology on **17**th **February**, **2014**. The revision version of the manuscript was send on **8**th **October**, **2014**. The second revision was send on **22**nd **December**, **2014**. The third revision was send on **23**rd **March**, **2015**. This manuscript, or parts of it, has not been submitted to other journals.

* Corresponding author.
E-mail address: johanna.varjo@ttl.fi (J. Varjo).

http://dx.doi.org/10.1016/j.jenvp.2015.08.001
0272-4944/

laboratory experiments. However, most of the previous laboratory studies have focused on the effects of a single factor of IE. In the present study, we simultaneously manipulated three IE factors in order to examine their joint effects on task performance and environmental satisfaction. Fig. 1 depicts how our study was designed as a follow-up study to three previous studies each separately examining the effects of a single factor of IE. We next summarize the evidence for the effects of each factor examined separately.

1.1. Effects of office noise

Office noise, especially irrelevant speech having sufficiently high intelligibility, has been shown to decrease performance in serial recall (e.g., Haapakangas et al., 2011; Haka et al., 2009), information search (Jahncke, Hongisto, & Virjonen, 2013), proofreading (e.g., Venetjoki, Kaarlela-Tuomaala, Keskinen, & Hongisto, 2006) and counting tasks (e.g., Buchner, Steffens, Irmen, & Wender, 1998). Our experiment was preceded by an experiment conducted in the same laboratory, which showed that the room acoustic design, where the intelligibility of irrelevant speech could be reduced, improved work performance (Haapakangas, Hongisto, Hyönä, Kokko, & Keränen, 2014).

Moreover, subjective assessments confirm the negative impact of highly intelligible irrelevant speech; speech and other office activity sounds negatively affect subjective well-being, acoustic satisfaction and self-estimated performance (Evans & Johnson, 2000; Haapakangas et al., 2014; Haapakangas et al., 2011; Haka et al., 2009).

It is important to study how different room acoustic solutions usually applied in open-plan offices can be used to reduce the negative effects of irrelevant speech. The effects seem to mainly depend on speech intelligibility (Ellermeier & Hellbrück, 1998; Hongisto, 2005; Jahncke et al., 2013) and not on the loudness of speech (Colle, 1980). Performance is expected to decrease with increasing Speech Transmission Index, STI (Hongisto, 2005). Subjective speech intelligibility can be objectively evaluated by measuring STI which ranges from 0.00 to 1.00, with large values representing highly intelligible speech (ISO 3382-3). STI can be reduced in open-plan offices by simultaneous application of high room absorption, high screens between desks and the use of masking sound (Bradley, 2003; Keränen & Hongisto, 2013; Virjonen, Keränen, & Hongisto, 2009). By reducing the STI values below 0.30, it can be expected that the negative effects on performance can be significantly reduced (Haka et al., 2009; Jahncke et al., 2013; Keus van de Poll, Ljung, Odelius, & Sörqvist, 2014) compared to a situation where the STI is above 0.50, which is, unfortunately, very typical in open-plan offices (Keränen & Hongisto, 2013; Virjonen et al., 2009).

Our study involved an acoustic manipulation where the two most important factors of acoustic design were considered simultaneously: sound masking and room absorption. Absorption was used to reduce the reflection of sound from room surfaces and to reduce the overall speech level. Sound masking was used to reduce the signal-to-noise ratio of speech. Successful application of masking sounds in the office has been reported by Hongisto (2008) and Hongisto et al. (2012).

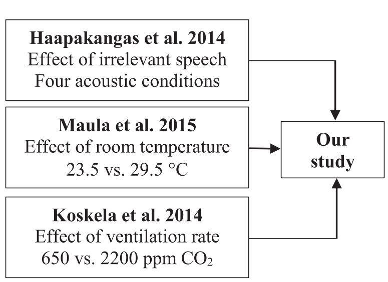

Fig. 1. Our study was preceded by three experimental studies in the same laboratory where the effects of the three IE factors were examined separately.

1.2. Effects of high room temperature

Room temperature can affect cognitive performance (see e.g., reviews of Hancock, Ross, & Szalma, 2007; Pilcher, Nadler, & Busch, 2002). However, the results of these reviews cannot be directly applied to office environments because the examined thermal conditions differed from usual thermal conditions in offices. The desirable room temperature in offices is between 21 °C and 25 °C depending on outside temperature, clothing, activity level and cultural differences. However, much higher temperatures, up to 35 °C, can be found in offices having insufficient cooling capacity or no cooling at all. When neutral temperatures (21–25 °C) have been compared to higher ones (above 26 °C), cognitive performance has been observed to decline at higher temperatures in short-term free recall tasks (Hygge & Knez, 2001), addition and visual tasks (Lan, Wargocki, Wyon, & Lian, 2011) and working memory tasks (Häggblom, Hongisto, Haapakangas, & Koskela, 2011). Maula et al. (2015; Fig. 1) performed an experiment before our study in the same laboratory environment. They found that high temperature (29 °C) affected the performance in working memory demanded N-Back task. However, temperature did not affect psychomotor, attention or long-term memory tasks. These results are consistent with the suggestion of Hancock et al. (2007) that the performance effects of room temperature are task-sensitive.

Subjective assessments yield a more uniform picture of the effects of room temperature. High temperature has been reported to negatively affect mood, energy, motivation, concentration and the assessment of air quality (Lan et al., 2011; Maula et al., 2015). High temperature has also been found to increase self-rated intensity of somatic symptoms compared with neutral temperature (Lan et al., 2011).

1.3. Effects of air quality

Air quality is affected by the ventilation rate and emissions from the building, furniture and occupants (Wargocki, Bakó-Biró, Clausen, & Fanger, 2002). In the majority of laboratory experiments investigating the effects of air quality on performance, the air quality has been reduced by artificial material emissions, such as by installing old and polluting carpets in the room. The combination of artificial material emission and small ventilation rate has marginally decreased performance in typing and negatively affected subjective assessments of air quality and well-being (Wargocki, Wyon, Sundell, Clausen, & Fanger, 2000). Similar results have been found for high material emissions with a constant ventilation rate (Wargocki, Wyon, Baik, Clausen, & Fanger, 1999).

When office buildings are renovated, furniture and decoration are likely to be changed and old material emission sources are usually removed. The emissions from new furniture and surface materials can cause relatively high concentrations of volatile organic compounds (VOCs) for a couple of months. The combination of high material emissions from new materials and small ventilation rate has been found to decrease objectively measured performance in typing, addition and memorization tasks and to reduce the acceptability of perceived air quality (Park & Yoon, 2011). In comparison, a previous study (Koskela, Maula, Haapakangas, Moberg, & Hongisto, 2014, Fig. 1) carried out in the same laboratory as our study investigated the situation where the occupants themselves were the strongest pollution sources of the room. Comparison between high (28 l/s per person, 600 ppm) and low (2 l/s per person, 2200 ppm) ventilation rates with negligible emission from furniture and building did not reveal any remarkable differences in subjective environmental assessment or objective performance despite the fact that the tasks required cognitive processes relevant to many types of office work and the exposure time was reasonably long (3.5 h). In contrast, decision-making performance has been shown to decline by high carbon dioxide (CO_2) concentrations (Satish et al., 2012). Due to methodological differences, more research is needed to confirm the effects of air quality on task performance and subjective assessment of the work environment. It is also worth noting that there is no fundamental theory of what mechanisms, such as fatigue or working memory capacity, would explain the effects of poor air quality on cognitive performance.

1.4. Combined effects of IE

Only a few studies have focused on the effects of several simultaneously modified IE factors on objective performance, despite the fact that occupants are exposed to several factors of IE in open-plan offices. In addition, the results are inconsistent. Some studies have reported expected performance effects (Balazova, Clausen, & Wyon, 2007; Hygge & Knez, 2001). Also interaction effects between IE factors have been reported (Witterseh, Wyon, & Clausen, 2004). However, even major changes in IE conditions have not always affected performance (Balazova, Clausen, Rindel, Poulsen, & Wyon, 2008; Clausen & Wyon, 2008).

On the other hand, subjective ratings have indicated more uniformly how several IE factors together affect human comfort. Simultaneous negative changes in IE factors, such as temperature, irrelevant speech (office noise), traffic sounds and air quality, have increased dissatisfaction (Balazova et al., 2007, 2008) and reduced the perceived ability to concentrate on and perform job-relevant tasks (Balazova et al., 2007; Clausen & Wyon, 2008).

A wide variety of environmental conditions, tasks and procedures have been used in the aforementioned studies. First, performance is affected differently by different noise types (e.g., Banbury & Berry, 1998; Szalma & Hancock, 2011). Studied noises have been originated from office noise (Balazova et al., 2008; Clausen & Wyon, 2008; Witterseh et al., 2004), road traffic (Balazova et al., 2007; Clausen & Wyon, 2008), or ventilation (Hygge & Knez, 2001). It has been shown that intelligible speech interferes with working memory and performance because it is unpredictable and information-rich, while constant noise does not cause interference (Jones, Madden, & Miles, 1992). Second, the exposure times have varied from 20 min (Balazova et al., 2007) to 6 h (Balazova et al., 2008). Third, successful task performance requires several cognitive processes (Sörqvist, 2014); very different tasks relying on varied cognitive processes have been used to measure task performance. Fourth, in some studies participants have been told about the IE conditions before the experiment (e.g., Balazova et al., 2007) or they have been able to personally select the IE conditions (Clausen & Wyon, 2008). Finally, both between-subjects (e.g., Hygge & Knez, 2001) and within-subjects (Balazova et al., 2007, 2008) designs have been employed. All these variations can be assumed to affect the results on performance and subjective assessment of the IE via different paths.

In our study, the environmental conditions and experimental procedures were selected so that both practical questions related to the work environments and the highest possible scientific quality could be met. Speech was used as the noise source since it is the most often complained noise source (e.g., Haapakangas et al., 2008). Several cognitive tasks were applied to cover a variety of different kinds of office work. A moderately long exposure time was applied to mimic a typical working period at the office desk (2 h). The participants were blind with respect to the experimental manipulations. Finally, a within-participants design was used to reduce inter-individual variability in cognitive performance and environmental effects.

1.5. The aim of the study

Our study represents the final experiment of a larger research programme (Fig. 1). The experiment, which combines three factors of IE (intelligibility of irrelevant speech, temperature and air quality), was preceded by three laboratory experiments that focused on each single factor. Out of these three experiments, performance effects were found in the experiment concerning the intelligibility of irrelevant speech (Haapakangas et al., 2014) and room temperature (Maula et al., 2015). The effects of these factors were also evident on subjective responses. However, the impact of low ventilation rate (2200 ppm CO_2 concentration; Koskela et al., 2014) on subjective ratings was small and no effect on task performance was observed. Theoretically, it is interesting to study how the effects of these three factors add up when the conditions are presented jointly. For example, it is possible that the detrimental effects of the individual conditions are exacerbated when combined. Because IE problems are often multifaceted in workplaces, the simultaneous evaluation of these factors is also important for its practical significance. As far as we know, no experimental studies exist that have investigated the simultaneous exposure to the intelligibility of irrelevant speech, room temperature and air quality (ventilation rate) and the related effects on cognitive performance and subjective experience in an open-plan office. A variety of tasks was used to tap into the job-related cognitive processes required in typical office work (non-communicative tasks). Because subjective assessment has proved to be sensitive to capture the effects of IE, an extensive battery of questionnaires was also employed.

The aim of this experiment was to investigate the simultaneous exposure to highly intelligible irrelevant speech, high room temperature and low ventilation rate and the effects of exposure on cognitive performance and environmental satisfaction. The study was carried out in open-plan office which was built in a laboratory environment. Two experimental conditions were investigated. Condition A was a combination of neutral temperature (23.5 °C), high outdoor air supply rate (30 l/s per person, 580 ppm CO_2) and low intelligibility of surrounding irrelevant speech (high absorption and adequate masking in the room). Condition B was a combination of high room temperature (29.5 °C), highly intelligible irrelevant speech (no absorption or masking in the room) and a very low outdoor air supply rate (2 l/s per person, 1470 ppm CO_2). The hypothesis was that Condition B would be inferior to Condition A with respect to the various subjective and objective measures used.

2. Materials and methods

2.1. Participants

Sixty-five students (49 women and 16 men) from six faculties of the University of Turku took part in the laboratory experiment. The participants were 19–29 years old ($M = 23$, $SD = 2$) and were all native Finnish speakers. Each participant took part in the two conditions on two successive weeks. None of the participants reported any hearing difficulties, dyslexia or an attention deficit disorder and all had normal or corrected-to-normal vision. Participants were recruited via university email lists and were paid 50 euros (minus taxes) for their participation. Prior to the experiment, the participants were told that the aim of the experiment was to investigate work performance in an open-plan office environment. The participants were not informed beforehand about the experimental conditions.

2.2. Design

The experiment was carried out using a within-participants design, i.e., all participants were tested in both experimental conditions (two sessions), thus acting as their own controls. The within-participants variable was an indoor environment (IE) which had two conditions. The order of exposure to these experimental conditions was counterbalanced across participants in altogether thirteen groups. Half of the participants were first exposed to Condition A (32 participants) and the other half to Condition B (33 participants).

2.3. Test environment

The experiment was carried out in an open-plan office specially built for the purpose (Fig. 2). The room was carefully furnished and finished to resemble a normal work environment. The room was free from measurement apparatus usually found in laboratories.

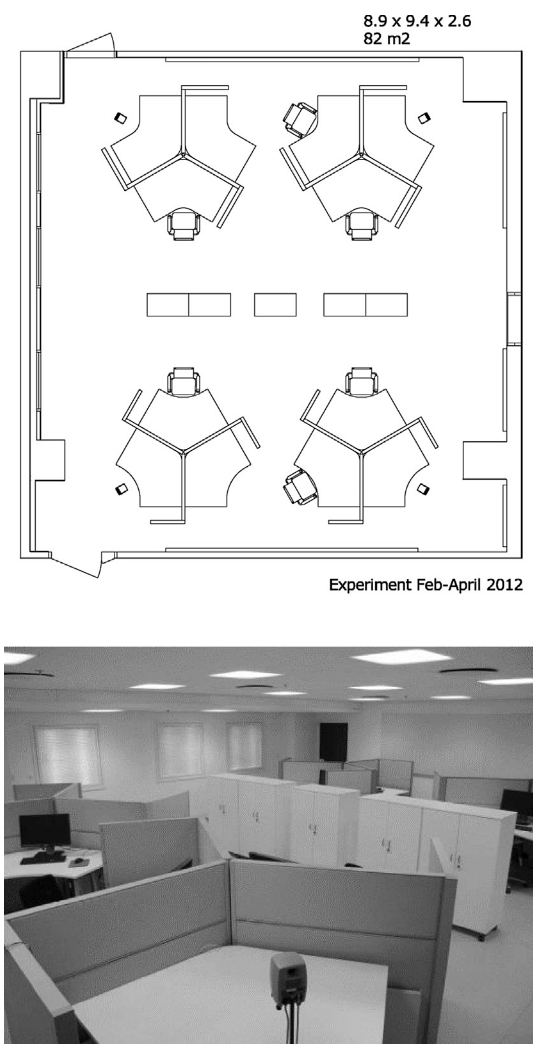

Fig. 2. Top: The layout of the office laboratory. The four speech loudspeakers are shown at the corner desks. Bottom: A photograph of the desk area showing one of the loudspeakers.

There were no measurement devices or other artificial artefacts visible in the room. The furniture and materials visible in the room were modern and commercially available. The furniture consisted of 12 identical desks, 1.3 m high screens between the desks, chairs, computers and storage units in the middle of the room. Six desks were reserved for the participants. The corner desks were reserved for the loudspeakers from which the background speech was played. Two desks were empty during the experiment.

The height of the suspended ceiling was 2.55 m and suspension depth was 0.30 m. The suspended ceiling was made of 600×600 mm metal grid where 210 ceiling boards, six ventilation inlets, a ventilation outlet and 16 lighting units were installed. Approximately 88% of the ceiling (75 m^2) was reserved for ceiling boards where either sound-absorbing (Condition A) or the non-sound-absorbing (Condition B) boards could be installed. Both boards had the same visual appearance so that the participants could not detect any differences between them. The walls were made of double drywall to provide good sound insulation to the surrounding office premises. Approximately 20% of the total wall area (18 m^2) was reserved for porous linen pictures behind which sound-absorbing boards could be installed in Condition A. Artificial masking sound was produced from 14 loudspeakers placed evenly above the suspended ceiling so that the participants could not see them and the masking sound was most probably experienced as ventilation noise. The temperature and ventilation of the room could be controlled by an independent air-conditioning system. The air leakages in the room and ventilation ducts were minimized. Natural daylight had no access to the room. Two artificial windows were installed on one wall behind which lighting units were installed to resemble daylight. The illumination level of this artificial daylight was negligible in the desk area.

2.4. The experimental conditions

The two experimental conditions were designed on the basis of our previous laboratory experiments focussing on single IE factors (Haapakangas et al., 2014; Koskela et al., 2014; Maula et al., 2015). The experimental conditions are described in Table 1. In Condition A, the IE was designed according to the most stringent target values used in open-plan offices in Finland (Class S2 of LVI 05–10440 en,

Table 1
Room temperature, air quality and speech characteristics of the two experimental conditions. Acoustic conditions were not constant since the participant's distance to the active speaker varied with time since the distance to the active loudspeaker changed with time.

	Experimental condition	
	A	**B**
Room temperature		
Average over all test days [°C]	23.6	29.5
Range [°C]	23.2–23.9	29.3–29.8
Air quality		
Fresh air flow rate [l/s per person], at least	30	2
CO_2-concentration: average over test days [ppm]	580	1470
CO_2-concentration: range [ppm]	530–630	1370–1600
TVOC concentration [μg/m3]	70–140	290–550
A-weighted level of speech [dB]		
Equivalent level over the 1½ hour session	45	51
Minimum level (speaker 6 m away), $L_{Aeq,15s}$	42	49
Maximum level (speaker 2 m away), $L_{Aeq,15s}$	48	53
A-weighted level of masking sound [dB]		
Equivalent level over the session	45	36
Speech Transmission Index STI		
Mean value over the 1½-hour session	0.37	0.67
Minimum value (speaker 6 m away)	0.22	0.62
Maximum level (speaker 2 m away)	0.52	0.72

2008). In Condition B, the IE was designed to violate even the least demanding target values (Class S3 of LVI 05–10440 en, 2008).

A neutral temperature has been found to be 23.5 °C in a study (Maula et al., 2015) similar to our study. This is very typical in office workplaces throughout the year in buildings equipped with modern air conditioning systems. Elevated temperatures up to 30 °C are, however, found during the summer season in many workplaces situated in buildings where cooling has not been installed. The temperatures in Conditions A and B were 23.5 and 29.5 °C, respectively, as used by Maula et al. (2015).

Air quality in office-type buildings, where the occupants are the main pollutant sources, is usually determined by measuring the CO_2 concentration. Concentrations up to 2500 ppm have been reported in office buildings (Seppänen, Fisk, & Mendell, 1999). Concentrations exceeding 3000 ppm and even up to 5000 ppm have been found, for example, in schools and meeting rooms with an insufficient ventilation rate and high crowdedness (Bakó-Biró, Clements-Croome, Kochhar, Awbi, & Williams, 2012; Seppänen et al., 1999). In large modern office buildings, the concentrations are normally below 1000 ppm (Apte, Fisk, & Daisey, 2000). The recommended upper limit for the CO_2 concentration in offices is 1200 ppm in the lowest quality class S3 (LVI 05–10440 en, 2008). In quality classes S1 and S2, the recommendations are 750 ppm and 900 ppm, respectively. Values below 650 ppm are usual in modern Finnish offices. Condition A (580 ppm) conformed to this recommendation. Ventilation rate (30 l/s per person) was selected to certainly meet recommendations but not cause noise or the risk of draught. For Condition B, a CO_2 concentration of 1470 ppm was applied. It should be noted that in the desk areas of open-plan offices, values exceeding 1500 ppm are rare. Ventilation rate for Condition B (2 l/s per person) was determined by ASHRAE standard 62.1 where the minimum ventilation rate is 2.5 l/s per person based on adapted persons.

Room acoustics in Condition A corresponded to the most stringent requirements (Bradley, 2003; Hardy, 1957; LVI 05–10440 en, 2008; Virjonen et al., 2009). Speech intelligibility is reasonably high close to a speaker to enable normal face-to-face conversation but it declines fast with increasing distance resulting in low intelligibility of distant speech. Condition B corresponded to a situation where the room acoustic design does not conform even to the least demanding target values. This is nevertheless very typical in workplaces even though the room acoustic design principles have been published 57 years ago (Hardy, 1957). Speech intelligibility is very high both at short and large distances from the speaker so speech is expected to be disturbing independent of the distance of the speaker. The sound levels of background speech (the supposedly disturbing sound) and masking sound (the supposedly non-disturbing sound) in the desks are depicted more closely in Fig. 3.

The lighting level at the table was constant, approximately 500 lux, in both conditions. The air velocity was less than 0.1 m/s in all desks and in both conditions.

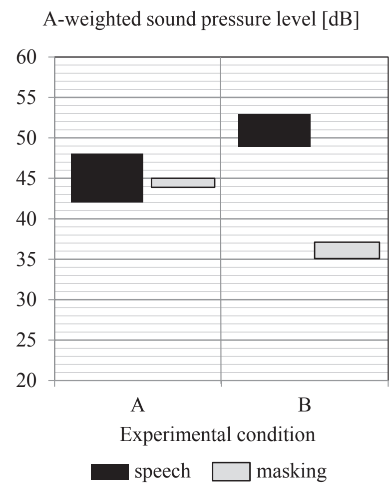

Fig. 3. The range of speech and masking levels in both experimental conditions. In Condition A, the signal-to-noise ratio was significantly smaller than in Condition B. The differences of masking levels between desks and with time were negligible because of 14 ceiling loudspeakers. Speech levels varied more with time and between desks because the distance to the active speaker varied during the experiment (see Production of irrelevant speech and masking sounds).

2.5. Implementation of thermal control and ventilation

The room temperature was controlled by six commercially available chilled beams in the ceiling and six electric radiators (total thermal power 5.5 kW) hidden below the tables of the unoccupied desks. The intake air was filtered with level F7 filters. Radiators were on during Condition B and chilled beams were used in Condition A. The room temperature was measured continuously in the corner desks (not visible to the participants) at a height of 1.1 m. Prior to the experiment, the implementation of thermal conditions in the desks (temperatures and draught) was also tested by installing six dummy bodies (60 W thermal load) at the desks. Based on this, it was deemed sufficient to monitor the temperatures during the experiment only at the corner desks.

During Condition A, the CO_2 concentration was planned to be low by means of high outdoor air supply. During Condition B, the desired CO_2 concentration was planned to be high. Therefore, the rate of outdoor air supply was minimized by using a circulation duct in the technical room. To implement the poor air quality, eight employees worked in the office with this reduced ventilation rate for two hours to increase the CO_2 concentration up to 1470 ppm before the participants arrived. By doing so, the human-based pollution rate (both CO_2 and other human-based pollutants) was at the intended level right at the beginning of the experiment and the participants maintained the same pollution rate thereafter. There were typically six participants and the experimenter in the room at a time but the number of persons varied from four to seven because of occasional cancellations. The thermal load was compensated accordingly and temperature variation between sessions and within a session was negligible. However, the ventilation rate was not changed according to the number of persons present in the room resulting in some variation of CO_2 concentration between sessions (see Table 1).

The participants were instructed to wear trousers, long-sleeve shirt, t-shirt, socks and angle-length shoes. The estimated clothing insulation including office chair was 0.83 clo (ISO 7730) in both conditions. The participants' main activity was typing and the estimated activity level was 1.1 met. According to the PMV-model, the temperatures 23.5 °C and 29.5 °C are estimated to be neutral and slightly warm, respectively (Maula et al., 2015).

2.6. Production of irrelevant speech and masking sounds

Speech was used as the sound source in both Condition A and B. Four loudspeakers in the four corner desks were used to simulate a situation where four persons would be talking on the phone one after the other about different topics so that a different plot was ongoing in each corner of the room (Figs. 2 and 4). In most previous experiments investigating the effects of speech on task performance, the speech levels have been manipulated electronically. Our experiment did not contain electronic manipulation of speech levels to demonstrate the effects of two extremely different room

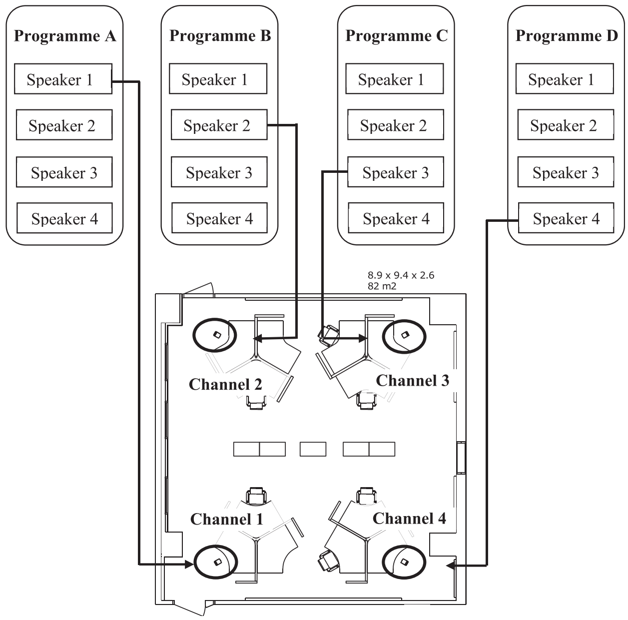

Fig. 4. The capturing of the individual speakers from four different radio programmes to the four loudspeakers (channels).

acoustic conditions in such way which could also be implemented in real open-plan offices.

The sound power level of speech emerging from the loudspeakers was constant in both conditions. The speech heard by the participants depended solely on the room acoustic treatment of the room, i.e., the changes in the amount of room absorption materials and the level of masking sound (see Chapter 2.7).

A four-channel sound file (wav file, 44.1 kHz) was used to produce the speech to four loudspeakers (Fig. 4). Only one corner speaker was active at a time. The speech was obtained from four radio programmes bought from YLE (The Finnish Broadcasting Company). The speech for each loudspeaker originated from a different radio programme so that the topic in each corner was different. In each programme, four participants (politicians, celebrities or specialists) were discussing a topic of common interest. The speech of each participant in each programme was isolated from the programme and placed to one channel of the four-channel sound file. The speech material of the speaker in each channel was thereafter edited so that the speech consisted of separate 5-to-25-s-long tokens. The four-channel speech file was then expanded and arranged so that there was no overlap between the channels. The order of the speakers was pseudo-randomized but the total amount of speech from every speaker was equalized. A silent period of 1–8 s was inserted between speech tokens. The sound level of each speech token was adjusted to the same A-weighted level. The editing was done using audio editing software (Adobe Audition 3.0). In addition, the spectrum of each speaker was modified so that the octave band spectrum shape deviated from the speech spectrum of ISO 3382-3 by less than 3 dB. The lengths of the two four-channel recordings used in the experiment were approximately 180 min. Two versions of speech recordings were created and counterbalanced across conditions. The versions included different topics and speakers.

The loudspeakers having mouth-like directivity were used (Genelec 6010). The height was 1.2 m from the floor. The speakers were directed to the centre of the room. The participants could see the loudspeakers when they entered the room but the speakers were not visible while working. The sound power level of the speech was constant and equal between every channel. This was checked out by full-time sound power level measurements of each channel in a reverberant room according to ISO 3741. The speech level (effort) was set between normal and casual. The A-weighted mean level over all directions was 53.0 dB at 1 m distance in a free field (normal effort is 57.4 dB). The linear sound power levels (L_W *re* 1 pW) 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.

The overall level of masking sound was 9 dB higher in Condition A than Condition B. However, the spectrum shape of the masking was constant and close to Brownian noise in octave bands 125–4000 Hz (Fig. 5) both in Condition A and B. This spectrum has been found adequate both in workplaces and in laboratory conditions (Hongisto, 2008; Hongisto et al., 2012, 2014). The difference in masking levels between the desks was negligible (±1 dB) because of the smooth distribution of masking loudspeakers.

2.7. Implementation of room acoustic conditions

In Condition A, sound-absorbing boards (EN 11654:1997; class A, 20 mm mineral wool) were installed to the whole ceiling (75 m^2). Sound-absorbing absorbers (EN 11654:1997; class A) were installed behind the porous linen wall pictures (18 m^2). The materials had the highest available M1 classification so that the emission levels of various compounds were very low. The overall room absorption area including the furniture was very large, altogether 123 m^2. The artificial masking system was set on so that the mean level in the desks was 45 dB (L_{Aeq}). As a result of efficient sound masking and high room absorption area, the intelligibility of the speech was moderate when the speech was originating from the nearest desk (2 m away) and reasonably low when speech was originating from a desk further away (6 m, see Fig. 2 and Appendix). It must be noted that the speaker producing the speech was changed every 6–33 s. Therefore, the STI of speech varied with time and the variation is indicated in Table 1.

In Condition B, sound reflecting ceiling boards were installed (EN 11654; unclassified, plasterboard). The sound-absorbing boards

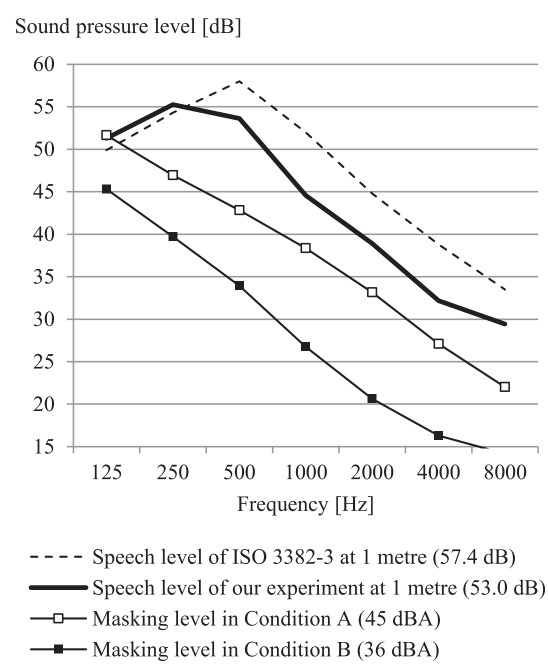

Fig. 5. The nominal spectra of speech and masking sounds at a distance of 1 m from the speaker in a free field. The actual speech levels heard by the participants were significantly lower (see Appendix 1). The masking spectra correspond to the actual levels.

were removed behind the linen wall pictures (EN 11654; unclassified). The overall room absorption area was very small mainly caused by furniture, altogether 30 m^2. Ventilation sound was planned to be the only masking sound. However, the artificial masking sound was played at a level of 33 dB (L_{Aeq}) to reach the same masking level in all desks since the ventilation level varied from 33 to 36 dB between desks. The final masking level was reasonably low (36 dB) and the resulting masking efficiency was negligible as desired. As a result of the inefficient sound masking level and negligible room absorption area, the STI value of speech was very high and almost independent on the distance to the speaker (see Appendix).

Neither the screen height nor the screen absorption was modified between Conditions A and B. The screen height was 130 cm and the screens were not sound-absorbing (EN 11654; unclassified).

The room acoustic conditions were measured according to ISO 3382-3 (See Appendix).

2.8. Cognitive tasks

Six cognitive tasks were used: a serial recall task, an operation span task, an N-back task, an information search task, a typing task and a story-writing task. The tasks mimicked the cognitive processes required in many types of office work. All tasks rely on several cognitive processes, but only the main processes that are typically related to these tasks in the literature are mentioned in the description of the tasks. This does not exclude the possibility that cognitive processes that are not typically associated with the particular task type may also have some effect on the performance of these tasks (Sörqvist, 2014). The serial recall task, the operation span task and the information search task were programmed with Visual Basic 6 (Microsoft) and the N-back task with E-Prime 2.0 (Psychology Software Tools Inc.).

The *serial recall task* is a classic short-term memory task where participants have to recall randomly presented digits from 1 to 9 in the correct presentation order. Numbers were presented on the screen one by one at the rate of 1 per second with an inter-digit interval of 1.5 s. During recall, the numbers 1 to 9 appeared in a 3×3 array on the screen and participants responded by clicking with the mouse the numbers in the recalled order. Participants were instructed to guess or press 'empty' in case they did not know the correct answer in a certain serial position. After each trial, participants had 15 s to respond, after which a new trial began. A total of 12 trials were presented but the first two were excluded from the analysis. The main dependent variable was the percentage of digits recalled in the correct serial position. The task took about 7 min to complete.

The *operation span task* is another highly used working memory task. The version used in our study is based on the original operation span task developed by Turner and Engle (1989). The task consisted of mathematical equation-word pairs, such as $3 \times 3 + 7 = 22$ and 'BOOK'. First, an equation appeared on the computer screen. Participants had 10 s to decide whether the presented equation was true or false by clicking the appropriate option on the computer screen. After each equation, a word appeared on the screen and participants had 2 s to memorize the word. Then the next equation and word were presented. After a predetermined number of equation-word pairs was presented (the number of presented equation-word pairs increased gradually from three to eight), participants typed in all the words they remembered. Small misspellings were allowed and the precise order of recalled words was not required. In this task, the function of the arithmetic task is to interfere with memorizing the words. To ensure that participants focused on both equations and words, participants were instructed to get at least 85% of responses to equations correct. Participants received feedback on this after typing each word list. The materials used for words and equations are described in detail in Haka et al. (2009). Two matched versions of the task were constructed and counterbalanced between the conditions and sessions. Within each version, equations and words were presented randomly. The main dependent variable was the percentage of correctly remembered words. The task took about 12 min to complete.

The *N-back task* (Gray, Chabris, & Braver, 2003) requires working memory and the ability to sustain attention. In this task, sequences of letters were presented on the screen one by one, each for 500 ms with a 2500 ms inter-stimulus-interval. Three difficulty levels were used (0-back, 1-back and 2-back). In 0-back, the participant's task was to press YES every time the letter 'X' appeared on the screen and press NO for all other letters. This is the baseline condition not taxing working memory. In 1-back, participants were required to respond whether the presented letter was identical to the one immediately preceding it. In 2-back, participants were required to response whether the presented letter was identical to the one presented two trials back. Participants were instructed to respond quickly, but accurately. Answers were given by pressing a key on the keyboard labelled YES (leftwards arrow) or NO (downwards arrow).

Each set (0-, 1- and 2-back) included 30 + n letters (n = 0, 1 or 2) in a pseudorandom order. One set included 9 letters requiring a 'YES' response (30%). Upper and lower case letters were varied requiring participants to use abstract letter codes and preventing them from relying only on visual letter feature. The whole task included three blocks containing one set of each difficulty level, i.e., each difficulty level appeared three times during the task. The materials and the order of the difficulty levels were counterbalanced across the test blocks and participants. Six matching versions of the task were constructed and counterbalanced across the conditions and sessions. The main dependent variables were response accuracy (%) and reaction time of the correct responses (in ms). Reaction times deviating over 2.5 SD from the participant's mean were excluded. The reliability of the reaction time measurement was based on Maula et al. (2015). The task took about 20 min to complete.

The *information search task* (Jahncke & Halin, 2012) requires working memory, attention and strategic executive functions. The task bears similarity to a problem-solving task. In this task, a table of 20 rows and seven columns was presented on the screen. In each row, an object was presented (e.g., a country) and each column described one aspect of the object (population, multilinguality, area, highest point in metres, major religion, etc.). In the descriptions, both categorical and numerical values were given. Participants were required to search for the object that met the presented criterion (see below for examples). Participants had a maximum of 60 s to respond, after which the next question appeared. The questions had the same grammatical structure. Altogether, four tables were used, two for both sessions. The task included 20 questions, 10 questions concerning the first table and 10 questions concerning the second table. Half of the questions were simple and half of the questions were difficult. In the case of difficult questions, the object could be nominated by investigating three columns, one with categorical and two with numeric values (e.g., "Which country is multilingual, its highest point over sea level exceeds 1000 m, and has the highest population?"). In the case of simple questions, participants had to follow two columns, one with a categorical value and one with a numeric value (e.g., "Which employee is a salesperson and has the highest annual income?"). Two matching task versions were constructed and counterbalancing across the conditions and sessions was made in relation to the task difficulty and the features of the objects and columns including categorical and numeric values across the tables. The main dependent variables were the percentage of total number of correct responses and the percentage of correct responses in simple and difficult questions separately. The number of exceeded response time was also analysed. The task took about 20 min to complete.

The *typing task* is a routine task that requires different subprocesses of perception, attention and motor functions (Rumelhart & Norman, 1982). In the task, participants had to copy a text presented on paper by typing it using the computer keyboard. Participants were required to type quickly but avoid making mistakes. Participants were instructed to correct any typing errors they made. The text was a story (Juvonen, 2012a, 2012b) with a rich vocabulary and plot. The story printed on paper was placed vertically on a slightly slanted rack for easy reading; the participants were allowed to freely place the rack on the desk. Participants had 10 min to copy the text. Two stories with similar writing style from the same writer were selected and counterbalanced between the conditions and sessions. The number characters in the final text, writing fluency, the number of corrected and uncorrected errors, and the total time of pauses exceeding 2 s were analysed. Corrected errors refer to typing errors that a participant made but corrected, whereas the errors that a participant did not correct were labelled as uncorrected errors. There were no spelling errors in the original text that was given to the participants. Writing fluency was operationalized as the total number of characters produced during the given time (i.e., the sum of the number characters of the final text and the total number of deleted characters; Sörqvist, Nöstl, & Halin, 2012). The task was carried out and the data were collected using the ScriptLog program (Strömqvist & Karlsson, 2002).

The *story-writing task* requires psychomotor performance, creativity and the ability to produce text in writing (Sörqvist et al., 2012). In the task, a photograph was presented on the screen. Participants were required to write a story about the scene depicted in the photograph. Participants were instructed to write whatever came to mind from that picture. The presented photographs displayed two different nature scenes (a forest road surrounded by green trees and a snowy mountain view with a small cottage in the middle). Participants had 5 min to write the story and they were instructed to write as much as possible but correct their typing mistakes. The presented pictures were counterbalanced between the conditions and sessions. The ScriptLog program (Strömqvist & Karlsson, 2002) was used to collect and analyse the data. The stories were analysed both quantitatively and qualitatively. In the quantitative analysis, the same variables as in the typing task were analysed. In the qualitative analysis, the quality of the stories (atmosphere, the existence of negative turning point, the invocation of the photograph and concreteness) was analysed by two raters. These factors were found adequate for most stories written by the participants. Inter-rater reliability was assessed using the overall percentage of agreement. Agreement varied between 53 and 94%.

2.9. Questionnaires measuring subjective experience

Five different questionnaires were created. Participants responded to the questionnaires at the baseline phase (questionnaire A), at the beginning of the acclimatization phase (questionnaire B, the first session or questionnaire E, the second session), before a short break during the experimental phase (questionnaire C) and at the end of the session (questionnaire D) (Fig. 6). In questionnaire A, background information including gender, age, sleep during the preceding night, ability to function normally, noise sensitivity, bodily symptoms and experienced level of tiredness and motivation were measured. Noise sensitivity was measured with three items from Weinstein (1978) and with four items from Noise-Q (Schutte, Marks, Wenning, & Griefahn, 2007). The intensity of the six symptoms was measured (Table 3). Experienced cognitive fatigue were measured with a modified version of the Swedish Occupational Fatigue Inventory (SOFI; Åhsberg, Gamberale, & Gustafsson, 1998; Åhsberg, Gamberale, & Kjellberg, 1995) including three factors (tiredness, lack of energy and motivation). Every factor included three items (tiredness: sleepy, yawning, drowsy; lack of energy: worn out, exhausted, drained; lack of motivation: uninterested, indifferent, passive).

Questionnaire B included questions about experienced tiredness and motivation, and symptoms, but also questions about overall thermal comfort, satisfaction with temperature, draught and emotional reactions. Overall thermal comfort was measured using a 7-point scale (1 = cold, 7 = hot). Emotional reactions were measured with the modified version of the Zuckerman Inventory of Personal Reactions and Feelings (ZIPERS; Zuckerman, 1977), including four factors (*Fear/arousal*, *Positive affect*, *Anger/aggression* and *Attentiveness*) and one item measuring sadness (*I feel sad*). *Fear/arousal* was measured with three items (*My heart is beating fast*; *I feel fearful*; *I am breathing fast*), *Positive affect* with four items (*I feel carefree*; *I feel gentle*; *I feel happy*; *I feel like I could act friendly to someone*), *Anger/aggression* with two items (*I feel like hurting or telling off someone*; *I feel angry*), and *Attentiveness* with two items (Cronbach's alphas 0.53–0.89). Attentiveness is not reported in the results, as it overlaps in content with one symptom item (*Difficulties in concentration*).

Questionnaire C contained questions about overall thermal comfort, satisfaction with temperature, draught, emotional reactions, experienced tiredness and motivation, and symptoms but also included questions of subjective workload. Subjective workload was measured with four items (mental demand, frustration, effort and performance) which were modified from the NASA Task Load Index, NASA-TLX (Hart & Staveland, 1988; Moroney, Biers, & Eggemeier, 1995). Workload questions were included because subjective awareness of distraction may lead participants to invest more effort in performance (Schlittmeier, Hellbrück, Thaden, & Vorländer, 2008).

Questionnaire D contained questions about overall thermal comfort, satisfaction with temperature, draught, emotional reactions, experienced tiredness and motivation, symptoms and subjective workload. In addition, this questionnaire included questions of local thermal comfort (ten items), work effectiveness in a similar temperature (one item, Table 3), acoustic satisfaction, acoustic and visual privacy (three items, Table 3), experienced disturbance due to different factors present in the work environment (17 items) and subjective experience of the work environment (eight items). Acoustic satisfaction was measured with five items (habituation, disturbance, pleasantness, attention capture, and work efficiency in a similar sound environment). Items were combined to form a sum score (henceforth called *Acoustic satisfaction*; Cronbach's alphas 0.82–0.90). The disturbance caused by lightning, screen brightness and ergonomic conditions were also rated to rule out possible effects due to these factors.

Questionnaire E contained questions about overall thermal comfort, satisfaction with temperature, draught, emotional reactions, experienced tiredness and motivation, and symptoms. In addition, this questionnaire included background information questions: the amount of sleep during the preceding night and ability to function normally.

All questionnaires were presented with an Internet-based software (QuestBack, Finland), except for the thermal comfort questionnaire, which was also presented on paper.

2.10. Procedure

The experiment was conducted between February and April 2012. Participants took part in the experimental sessions during

Session 1	Session 2
Baseline phase (90 min) • Questionnaire A • Practice versions of the three tasks	
Break (20 min)	
Acclimatization phase (30 min) • Questionnaire B • Practice version of the rest tasks	Acclimatization phase (30 min) • Questionnaire E • Revision versions of all tasks (excluding the typing task)
Experimental phase (90 min) • Serial recall task (7 min) • Operation span task (12 min) • Story-writing task (5 min) • Questionnaire C • Break (a few minutes) • Information search task (20 min) • Typing task (10 min) • N-back task (20 min) • Questionnaire of thermal comfort on paper • Questionnaire D	Experimental phase (90 min) • Serial recall task (7 min) • Operation span task (12 min) • Story-writing task (5 min) • Questionnaire C • Break (a few minutes) • Information search task (20 min) • Typing task (10 min) • N-back task (20 min) • Questionnaire of thermal comfort on paper • Questionnaire D

Fig. 6. The procedure of the experiment. The participants were exposed to the two experimental conditions on two separate days which are called sessions 1 and 2.

two mornings on consecutive weeks. On the first morning, participants were asked to arrive at the laboratory at 8.30 a.m. for the baseline phase. The acclimatization and the experimental phase was carried out between 10.30 a.m. and 12.30 p.m. on both days. At the beginning of the first session, participants were given written information about the experiment and they signed the informed consent form. Participants were informed of the progress and the content of the test sessions, payment, confidentiality and their right to interrupt their participation at any time.

The procedure is described in Fig. 6. The first session started with a baseline phase in silence in a room near the experimental room. During the break after the baseline phase, sandwiches and soft drinks were served in the hallway to ensure that the participants had eaten breakfast prior to the rather long experimental session.

During the acclimatization phase, participants were acclimatized to the experimental condition and they practised the cognitive tasks (30 min). Performance was not measured during the acclimatization phase. Participants were unaware of the purpose of the acclimatization phase. The speech sounds were off.

When the experimental phase started, the experimenter switched on the speech. Participants were instructed to ignore the sounds and to concentrate on the tasks. Before every task, the experimenter shortly repeated the task instructions, after which the participants performed the task at their own pace. A short break was given in the middle of the experimental phase to give the participants an opportunity to visit the restroom and drink water. During the break the speech was switched off. The temperature in the hallway to the restroom and in the restroom was similar to that in the laboratory. Conversation was not allowed during the experiment or the break.

Prior to the beginning of the second session in the next week, sandwiches and soft drinks were served in the hallway. The acclimatization phase included questionnaire E and the revised versions of each task excluding the typing task. The experimental phase was identical to the first session. Participants were informed in detail about the aim of the study at the end of the second session.

2.11. Statistical methods

SPSS (IBM SPSS Statistics for Windows 20, IBM Corp, Armonk, NY) was used for the statistical analyses. The normality of data was tested with the Kolmogorov–Smirnov test. The serial recall, the operation span, the typing and the story-writing tasks were analysed using repeated measures ANOVA with experimental condition (later: condition) as the within-participants variable. The quality of the stories in the story-writing task was analysed using Pearson's chi-squared test for independence. The N-back and information search tasks were analysed with a 2 (condition) $\times$ 3 (difficulty level) repeated measures ANOVA. A repeated measures ANOVA was also used for the questionnaire items that were normally distributed or when distributions were similarly skewed. Subjective workload was analysed with a 2 (condition) $\times$ 2 (exposure time: after one hour vs. after two hours) repeated measures ANOVA. Experienced tiredness and motivation, and emotional reactions were analysed with a 2 (condition) x 3 (exposure time: beginning, after one hour, after two hours) repeated measures ANOVA. Whenever needed, the homogeneity of variance was estimated with Mauchly's test of sphericity. When Mauchly's test indicated a violation of sphericity, the Greenhouse-Geisser correction was applied and the corresponding p-values are reported. When an interaction was found, paired comparisons between conditions were performed using t-tests for the variables that were normally distributed or when distributions were similarly skewed; otherwise the Wilcoxon signed-rank test was used. An alpha level of 0.05 was used in all analyses. The Benjamini-Hochberg procedure (Benjamini & Hochberg, 1995) was used for alpha-error adjustment in paired comparisons. In the N-back task, the data were combined over the two blocks, because the main effects of the block remained non-significant.

One participant was excluded from the N-back analyses because of a misunderstanding of the instructions. Unlike in typical studies of the operation span task, participants who failed to achieve an 85% level on equation accuracy were not excluded, because the independent variable (speech intelligibility) could also have affected arithmetic performance (Jachnke et al., 2013; Schlittmeier et al., 2008). Instead, multivariate outliers were checked using Mahalanobis distance for identifying possible changes in performance strategy. As a result, one participant was excluded from the analysis. In the story-writing task, one participant was excluded from the analysis because he did not follow the instructions.

The decrement in performance, D [%], was defined as $D = 100 \cdot (1 - P_B/P_A)$ where P_B and P_A are the performance levels in Condition B and A, respectively.

The results for the four items (mental demand, frustration, effort and performance) measuring subjective workload yielded a main effect of condition for all items and there were no interactions between exposure time and condition. Thus, the items were combined to form a sum score (henceforth called *Subjective workload*; Cronbach's alphas 0.66–0.77).

During the experiment, the visual appearance of the office was also investigated using a between-groups design with two groups. Thirty-one participants experienced Conditions A and B in visual environment 1 and thirty-four participants in visual environment 2. Because the visual appearance did not significantly affect the objective or subjective results concerning the effect of condition, the groups were combined. There was one exception: the self-rated impairment of performance due to thermal conditions and air quality (heat, odours and stuffiness of the indoor air) was affected by the visual appearance. Thus, only the other group was included in the analyses of these variables (n = 31). For the remaining analyses, the participants from both groups were included.

3. Results

3.1. Performance measures

In the serial recall task, a significant main effect of condition was found for the percentage of digits recalled in the correct serial position ($F_{1,64} = 5.86$, $p = .018$, $\eta^2 = 0.08$; Fig. 7a). Performance was significantly better in Condition A ($D = 6.7\%$).

In the operation span task, a significant main effect of condition was found for the percentage of correctly recalled words ($F_{1,63} = 10.84$, $p = .002$, $\eta^2 = 0.15$; Fig. 7b), with better performance observed in Condition A ($D = 4.0\%$).

In the N-back task, a significant main effect of condition was found for response accuracy ($F_{1,63} = 4.01$, $p = .049$, $\eta^2 = 0.06$; Fig. 7c). Significantly more correct responses were given in Condition A. However, the performance decrement was negligible ($D = 0.5\%$). There was no significant interaction between difficulty level and condition. That is, the effect of condition on accuracy was not affected by the difficulty level. Reaction times were not affected by condition nor was there an interaction between difficulty level and condition.

The performance in the information search task was not affected by condition, neither by difficulty level ($p > .05$). Moreover, there was no significant interaction between difficulty level and condition ($p > .05$).

In the typing task, a significant main effect of condition was found for the total number of errors ($F_{(1,64)} = 4.03$, $p = .49$, $\eta^2 = 0.06$; Fig. 7d). Significantly fewer errors were made in Condition A ($D = 5.2\%$). This resulted primarily from the number of uncorrected errors: there was a marginally significant difference in the number of uncorrected errors ($p = .05$), but no difference in the number of corrected errors ($p > .05$). None of the other variables measuring typing fluency were affected by condition.

The performance in the story-writing task was not affected by condition for any measure ($p > .05$).

The mean values with standard deviations for the task variables where significant differences were not found (N-back, information search, typing and story-writing) are presented in Table 2.

3.2. Subjective responses

The results for the subjective assessments of the working conditions are presented in Fig. 8. The ratings for the conditions for working as a whole were significantly more positive in Condition A than B ($Z = -5.97$, $p < .001$). Similarly, the ratings of the possibility to work effectively were significantly higher ($Z = -4.66$, $p < .001$) and ratings of the riveting of the tasks were significantly more positive ($Z = -2.18$, $p = .029$) in Condition A.

There was a significant main effect of condition for the sum score of *Subjective workload* ($F_{1,64} = 15.02$, $p < .001$, $\eta^2 = 0.19$; Fig. 9a). *Subjective workload* was rated to be significantly lower in

Table 2
Means and standard deviations (in brackets) for the dependent variables of the performance tasks where performance did not differ between IE conditions.

Task	Variable	Experimental condition	
		A	B
N-back task			
- Accuracy (%)	0-back	98.9 (1.3)	98.4 (1.8)$^{n.s.}$
	1-back	98.1 (2.0)	97.9 (1.9)$^{n.s.}$
	2-back	96.7 (2.8)	96.0 (3.6)$^{n.s.}$
- Reaction time (ms)	0-back	447 (62)	447 (69)$^{n.s.}$
	1-back	479 (92)	481 (98)$^{n.s.}$
	2-back	551 (146)	563 (160)$^{n.s.}$
	Combined mean from the all levels of difficulty	492 (100)	497 (109)$^{n.s.}$
Information search task			
	Total correct responses (%)	81.9 (11.6)	79.6 (12.2)$^{n.s.}$
	Correct responses in simple questions (%)	92.9 (9.5)	91.1 (8.9)$^{n.s.}$
	Correct responses in difficult questions (%)	70.8 (17.6)	68.2 (20.5)$^{n.s.}$
	Number of exceeded response time in simple questions	0.0 (0.0)	0.0 (0.0)$^{n.s.}$
	Number of exceeded response time in difficult questions	1.4 (1.3)	1.1 (1.2)$^{n.s.}$
Typing task			
- Quantitative aspects	Characters of final text	2431 (683)	2434 (608)$^{n.s.}$
	Writing fluency	1282 (357)	1286 (322)$^{n.s}$
- Corrections and errors	number of corrected errors	48.2 (24.7)	49.9 (26.1)$^{n.s.}$
	Number of uncorrected errors	6.4 (4.0)	7.7 (5.9)$^{p=.05}$
	Total time of pauses exceeding 2 s	28.6 (36.6)	24.2 (31.1)$^{n.s.}$
Writing task			
-Quantitative aspects	Characters of final text	941 (255)	920 (253)$^{n.s.}$
	Writing fluency	528 (135)	517 (135)$^{n.s.}$
- Corrections and errors	Total number of corrected errors	34.5 (17.3)	34.6 (19.2)$^{n.s.}$
	Total time of pauses exceeding 2 s	52.8 (30.7)	57.1 (33.4)$^{n.s.}$

$^{n.s.}$non-significant.

Table 3
Means and standard deviations (in brackets) for the sum variables and other items of the questionnaires.

Variable	Experimental condition	
	A	B
Cognitive fatigue[a,c]		
Tiredness	1.8 (0.5)	2.0 (0.6)**
Lack of energy	1.5 (0.5)	1.9 (0.7)***
Lack of motivation	1.5 (0.5)	1.8 (0.7)***
Emotional reactions[a,d]		
Fear arousal	1.2 (0.4)	1.3 (0.5)$^{p=.05}$
Positive affect	2.8 (0.9)	2.5 (1.1)*
Anger/aggression	1.2 (0.6)	1.4 (0.6)n.s.
Sadness	1.2 (0.5)	1.4 (0.7)*
Working efficiency in a temperature like this[b]		
I could work effectively in a temperature like this for long periods of time	3.9 (1.0)	1.6 (0.8)***
Acoustic and visual privacy		
It would be more pleasant to work at the work site if there were more screens around it[b]	2.4 (1.2)	2.2 (1.1)*
Some work sites were too close to my own[b]	1.8 (0.9)	1.8 (0.8)n.s.
The distraction of the precense of other people[a]	1.2 (0.5)	1.3 (0.6)n.s.
Symptoms[a,d]		
A headache	1.3 (0.5)	1.5 (0.7)*
A runny or blocked nose	1.4 (0.6)	1.4 (0.6)n.s.
Dryness of the throat	1.5 (0.7)	1.7 (1.0)*
Dryness or irritation of the eyes	1.9 (0.9)	2.0 (1.0)n.s.
The feeling of being unwell	1.1 (0.5)	1.6 (0.9)***
Difficulties in concentration	2.3 (1.0)	2.8 (1.0)**

*$p < .05$, **$p < .01$, ***$p < .001$, n.s. non-significant.
[a] 1 Not at all, 5 Very much.
[b] 1 Disagree completely, 5 Agree completely.
[c] Average of the whole exposure time.
[d] At the end of the session.

Condition A. The effect of condition on *Subjective workload* was not affected by exposure time, as indexed by a non-significant interaction between exposure time and condition.

Acoustic satisfaction was significantly lower in Condition B ($Z = -4.82$, $p < .001$; Fig. 9b). The subjective assessments of the distraction of performance due to different sounds are presented in Fig. 10. Speech distracted significantly more in Condition B. Speech was more distracting both from nearby desks ($Z = -2.95$, $p = .003$) and desks from further away ($Z = -5.53$, $p < .001$). Moreover, speech from nearby distances was more distracting than speech from further away both in Condition B ($Z = -5.20$, $p < .001$) and in Condition A ($Z = -6.29$, $p < .001$). Similarly, a significant effect of condition was found for the distraction due to the sounds of computer tapping ($Z = -2.22$, $p = .026$) and other sounds made by other participants ($Z = -2.68$, $p = .007$).

The distraction of the hum of ventilation, i.e., masking sound, was significantly higher in Condition A ($Z = -4.70$, $p < .001$). However, it should be noted that speech either from nearby desks ($Z = -5.42$, $p < .001$) or desks from further away ($p = .06$) distracted significantly more than the hum of ventilation in Condition A. That is, speech was the main acoustic distractor in both experimental conditions.

Overall thermal comfort differed between conditions: participants felt warmer at the end of the exposure time in Condition B ($Z = -7.00$, $p < .001$); the difference between conditions was observed during the whole exposure time ($p < .001$). Using the 7-point scale of thermal comfort, participants felt warm ($M = 6.3$, $SD = 0.7$) at 29.5 °C and neutral ($M = 3.6$, $SD = 0.9$) at 23.5 °C after the second exposure hour ($Z = -7.00$, $p < .001$). A similar result was found for both local and overall thermal *comfort* during the whole exposure time ($p < .001$). The impairment of performance due to thermal factors is presented in Fig. 11. Overall, working efficiency was assessed to be significantly higher in Condition A ($Z = -6.5$, $p < .001$; Table 3). Heat was rated to interfere with performance

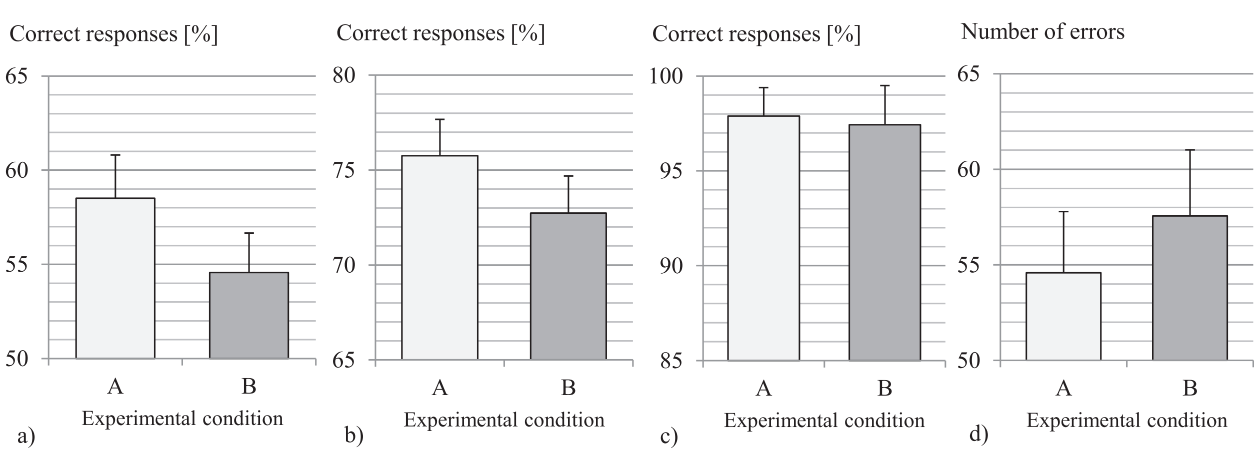

Fig. 7. Mean percentage of correct responses in the a) serial recall task, b) operation span task, and c) N-back task. The total number of errors in d) typing task. The error bars represent the standard errors of the mean.

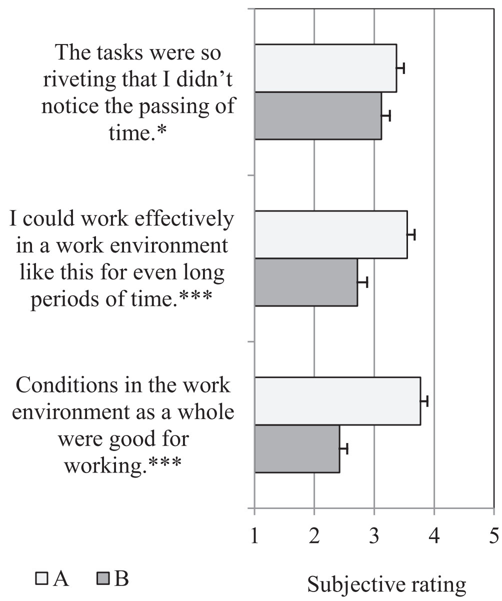

Fig. 8. Mean subjective ratings of the working conditions in experimental conditions A and B. The error bars represent the standard errors of the mean. Scale: 1 = Disagree completely, 5 = Agree completely. *p < .05, ***p < .001.

significantly more in Condition B ($Z = -4.60$, $p < .001$), whereas cold ($Z = -2.20$, $p = .028$) and draught ($Z = -2.08$, $p = .037$) interfered significantly more in Condition A. However, the mean ratings of cold and draught were very low in both conditions.

Self-rated air quality differed between conditions as expected (Fig. 11). Stuffiness of the indoor air ($Z = -3.71$, $p < .001$) interfered with self-rated performance significantly more in Condition B. No impairment caused by odours was reported in either condition.

Condition also affected the perception of acoustic and visual privacy (Table 3). Participants would have preferred more screens around the desk in Condition A in order to make working more pleasant ($Z = -2.20$, $p = .028$). The distraction caused by other people and closeness of nearby desks did not differ between experimental conditions ($p > .05$).

There was a significant main effect of condition for the perceived tiredness ($F_{1,64} = 10.40$, $p = .002$, $\eta^2 = 0.14$), the lack of energy ($F_{1,64} = 20.18$, $p < .001$, $\eta^2 = 0.24$) and the lack of motivation ($F_{1,64} = 20.55$, $p < .001$, $\eta^2 = 0.24$; Fig. 12). Participants were more tired and the lack of energy and motivation was higher in Condition B. There was also a significant interaction between the exposure time and condition in the lack of energy ($F_{2,128} = 9.62$, $p < .001$, $\eta^2 = 0.13$) and motivation ($F_{2,128} = 9.56$, $p < .001$, $\eta^2 = 0.13$). Paired comparisons revealed that the lack of energy was significantly higher in Condition B in the beginning phase ($t(64) = 7.05$, $p < .001$, two-tailed) and after the first ($t(64) = -3.68$, $p < .001$, two-tailed) and the second ($t(64) = -4.98$, $p < .001$, two-tailed) exposure hour. There was no significant difference in the lack of motivation between the conditions in the beginning phase but the lack of motivation was significantly higher in Condition B after the first ($t(64) = -3.10$, $p = .004$, two-tailed) and the second ($t(64) = -4.90$, $p < .001$, two-tailed) exposure hour. In Condition B, an increase in the lack of energy ($t(64) = 2.65$, $p = .013$, two-tailed) and motivation ($t(64) = 2.28$, $p = .033$, two-tailed) was observed already between the beginning phase and the first exposure hour. The lack of energy ($t(64) = 6.39$, $p < .001$, two-tailed) and motivation ($t(64) = 5.57$, $p < .001$, two-tailed) continued to increase, being significantly higher after two hours of exposure compared to one hour of exposure. In Condition A, no significant difference in the lack of energy and motivation was found between the beginning phase and the first exposure hour, but the lack of energy ($t(64) = 4.50$, $p < .001$, two-tailed) and motivation ($t(64) = 3.60$, $p = .001$, two-tailed) increased significantly between the first and the second exposure hour.

Regarding emotional reactions (Table 3), there was a significant main effect of condition for *Positive affect* ($F_{1,64} = 4.18$, $p = .045$, $\eta^2 = 0.06$; Fig. 13) with higher values in Condition A. There was also a significant interaction between condition and exposure time ($F_{2,116} = 4.15$, $p = .022$, $\eta^2 = 0.06$). Paired comparisons revealed that there was no significant difference between the conditions in the beginning phase or after the first exposure hour but *Positive affect* was stronger in Condition A after two hours of exposure ($t(64) = -3.80$, $p < .001$, two-tailed). In Condition B, *Positive affect* decreased significantly during the whole exposure time ($p < .001$) and difference was found between the beginning phase and the first exposure hour ($t(64) = 4.10$, $p < .001$, two-tailed) and between the first and the second exposure hour ($t(64) = 3.86$, $p < .001$, two-tailed). In Condition A, *Positive affect* decreased significantly between the beginning phase and the first exposure hour ($t(64) = 3.33$, $p = .002$, two-tailed) but no significant difference was found later between the first and the second exposure hour ($p > .05$).

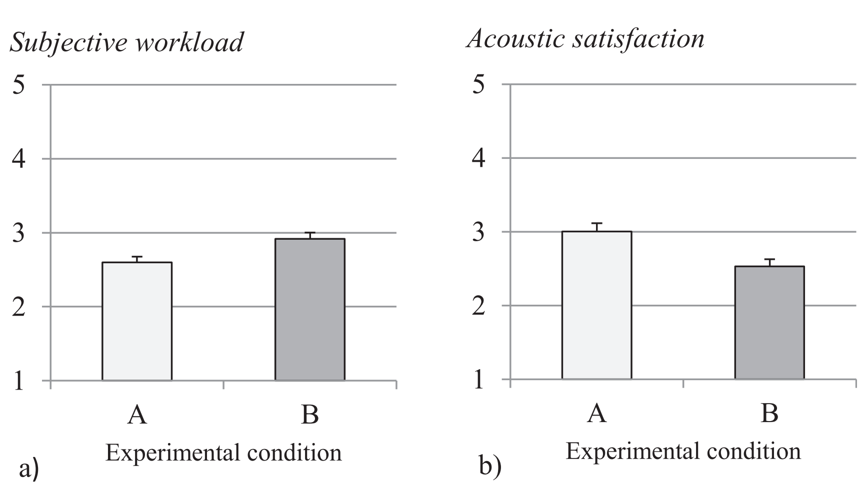

Fig. 9. Mean values of the sum scores of a) *Subjective workload* and b) *Acoustic satisfaction* after two hours of exposure (the error bars represent the standard errors of the mean). Scale: 1 = Minimum value, 5 = Maximum value.

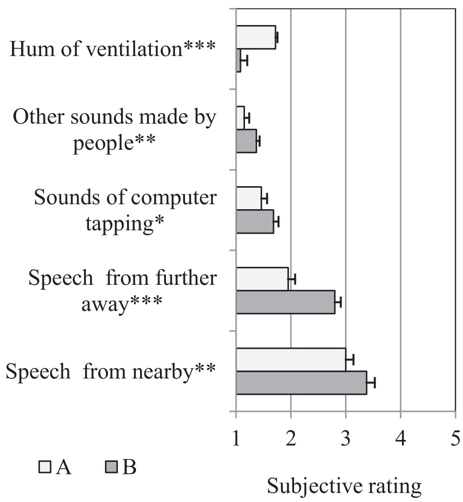

Fig. 10. Mean self-rated distraction of work performance due to different sounds in experimental conditions A and B (the error bars represent the standard errors of the mean). Scale: 1 = Not at all, 5 = Very much. *p < .05, **p < .01, ***p < .001.

The ratings of *Sadness*, *Fear/arousal* and *Anger/aggression* remained low during the sessions. However, there was a main effect of condition on sadness ($F_{1,64} = 6.13$, $p = .016$, $\eta^2 = 0.09$) with higher ratings in Condition B. There was also a marginal main effect of condition on *Fear/arousal* ($F_{1,64} = 3.98$, $p = .05$, $\eta^2 = 0.06$). Participants reported higher *Fear/arousal* ratings in Condition B. *Anger/aggression* was not affected by condition. There was no significant interaction between condition and exposure time in these ratings.

Different physiological symptoms (Table 3) were enquired on a scale from 1 (Not at all) to 5 (Very much). On the whole, the intensity of symptoms was very mild. Thus, a rating of 2 (Slightly) or higher was interpreted to indicate the existence of a symptom. The self-reported prevalence of different symptoms at the end of the experiment is shown in Fig. 14. Participants reported more feelings of being unwell in Condition B than in Condition A after the first ($Z = -2.74$, $p = .009$) and second ($Z = -3.83$, $p < .001$) hour of exposure. Symptoms increased with exposure time in Condition B (p < .05) but not in Condition A. The assessments of headache and nasal, throat and eye symptoms remained low in both conditions (Fig. 14), but the symptoms increased toward the end. Participants reported more headaches at the end of the session in Condition B than in Condition A

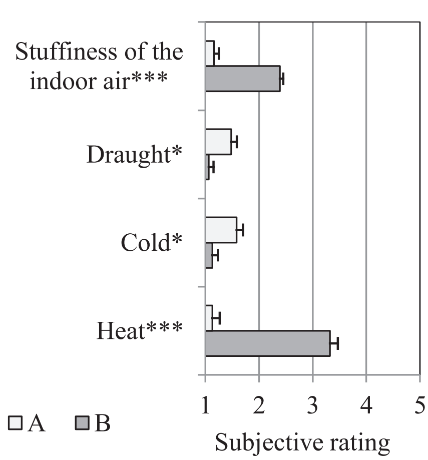

Fig. 11. Mean self-rated distraction of work performance due to thermal conditions and air quality in experimental conditions A and B (the error bars represent the standard errors of the mean; n = 31). Scale: 1 = Not at all, 5 = Very much. **p < .01, ***p < .001.

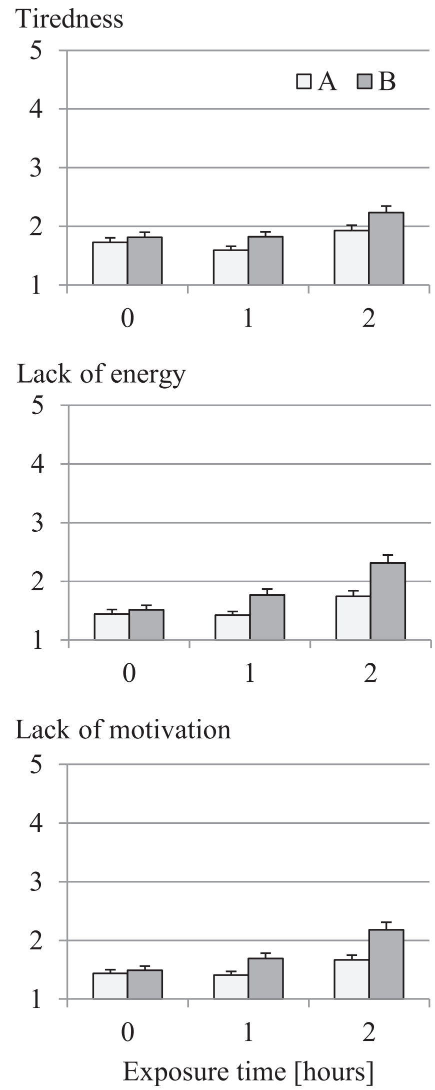

Fig. 12. Mean values of SOFI variables, as a function of experimental conditions A and B and exposure time. The error bars represent the standard errors of the mean. Scale: 1 = Smallest value, 5 = Largest value.

($Z = -2.43$, $p = .030$). No significant differences were found for other exposure times ($p > .05$). In addition, more throat symptoms were reported in Condition B after the first ($Z = -2.20$, $p = .033$) and the second ($Z = -2.29$, $p = .033$) hour of exposure compared with Condition A. Headache and throat symptoms also increased with exposure time in Condition B ($p < .05$) but not in Condition A. Ratings of nasal symptoms decreased during the first exposure hour compared to the beginning phase in Condition A ($p < .01$), but not in Condition B. However, a difference was not observed between the conditions ($p > .05$). Ratings of eye symptoms increased during the second exposure hour compared to the beginning phase in Condition B ($p < .05$). No significant changes were perceived in Condition A. However, no difference was observed between the conditions ($p > .05$).

Difficulties in concentration were higher in Condition B; the experimental conditions differed both after the first ($Z = -2.10$, $p = .046$) and second ($Z = -3.03$, $p = .004$) exposure hour.

4. Discussion

The aim of our study was to investigate the simultaneous

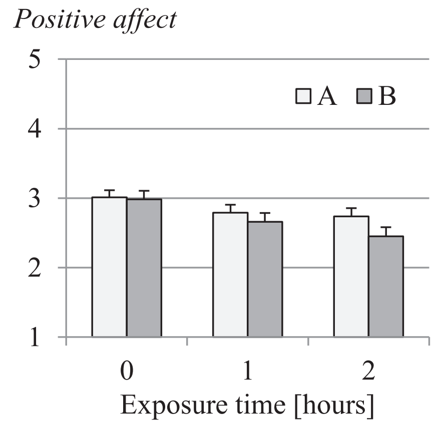

Fig. 13. Mean values of the sum scores of *Positive affect*, as a function of experimental conditions A and B and exposure time. The error bars represent the standard errors of the mean. Scale: 1 = Smallest value, 5 = Largest value.

exposure to highly intelligible irrelevant speech, high temperature and low ventilation rate and the effects of exposure on cognitive performance and environmental satisfaction in an open-plan office laboratory. As hypothesized, Condition B had detrimental effects on both cognitive performance and subjective experience. Condition A was perceived to be a better condition for office work.

4.1. Cognitive performance

The results established negative effects of Condition B on cognitive performance. An effect was observed in the percentage of correct answers and in typing errors. In the serial recall and the operation span tasks performance was worse in Condition B in comparison to Condition A. Participants also made more errors in typing when working in Condition B. Thus, Condition B proved to be poor IE for working.

The results from the single IE factors of prior experiments (see Fig. 1) are conceivable to reflect with our study where simultaneous exposure was used. Highly intelligible irrelevant speech reduced performance and interfered with the operation of working memory (Haapakangas et al., 2014). Similarly, high temperature reduced working memory performance (Maula et al., 2015). Instead, low

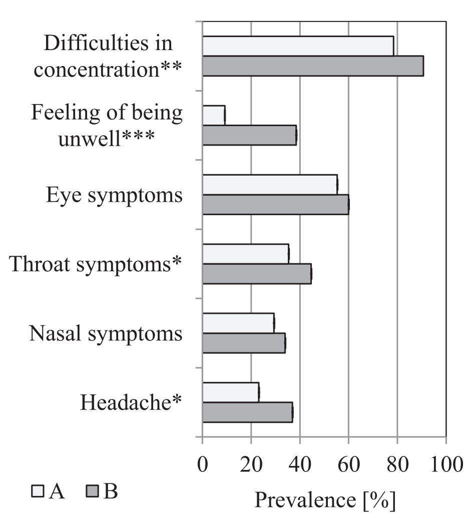

Fig. 14. Percentage of participants reporting bodily symptoms and difficulties in concentration at the end of the experiment in experimental conditions A and B. *p < .05, **p < .01, ***p < .001.

ventilation rate did not have consistently effects on cognitive performance (Koskela et al., 2014). Based on previous findings concerning single factors, high speech intelligibility and high temperature decreased performance the most. The results of our study expound that simultaneous exposure to high speech intelligibility, high temperature and low ventilation rate decrease performance. Despite of the findings of Koskela et al. (2014), the self-rated distraction of work performance caused by the stuffiness of the indoor air, heat (Fig. 11) and speech from further away and nearby desks (Fig. 10) received higher ratings in Condition B in our study. Taking our experimental design into account, where the effects of individual factors were not investigated, we can only conclude that high speech intelligibility, high temperature and low ventilation rate may have together affected the performance results through subjective responses in our study.

As pointed out above, performance differences were clearly found in the working memory tasks employed in our study. Regarding the N-back task, it is noteworthy that accuracy decreased in Condition B in comparison to Condition A, but the reaction times did not. Thus, response speed was maintained at a cost of more erroneous responses. This result is inconsistent with the speed-accuracy trade-off hypothesis (Hockey, 1984), which suggests that in noisy environments responses are given faster but with less accuracy. The reason may be that Hockey's experiment was conducted with pseudorandom noise instead of speech.

The results of the N-back task shall be compared with the prior studies conducted in the same laboratory. Similar background speech (Haapakangas et al., 2014) did not decrease the accuracy in the N-back task when three difficulty levels (0, 1 and 2) were applied as in our study. However, an effect on accuracy was observed when four difficulty levels (0, 1, 2 and 3) were used in high temperature (Maula et al., 2015). Taken together, it appears that simultaneous exposure to three IE factors in Condition B exacerbated somewhat the performance decrement in the N-back task compared with that was observed with similar single factor manipulation alone.

In the typing task, fewer typing errors were made in the better IE. In previous studies (Park & Yoon, 2011; Wargocki et al., 1999, 2000) the number of typed characters has been found to decrease with high material emissions in a similar text typing task while no significant difference has been observed in typing errors. On the other hand, Koskela et al. (2014; Fig. 1) obtained no effect of air quality on typing performance in a study conducted in the same laboratory as our study. A possible reason for this discrepancy in results is the fact that in our study not only ventilation rate but also thermal and acoustic conditions were manipulated. In other words, a set of poor IE conditions may need to be combined together to have an effect. Yet, methodological differences in manipulating air quality (human vs. material emission) and differences in task duration (10 min vs. nearly one hour) may also have contributed to the discrepancy.

In the story-writing task, participants had to produce a new text instead of copying an existing text. Story-writing was not affected by the experimental condition. Keus van de Poll et al. (2014) observed a detrimental effect of irrelevant speech on writing fluency and pauses in story-writing with STI values below 0.34. Processing of semantically meaningful speech was assumed to interfere with writing performance. It should be noted that in our study the STI values in both experimental conditions exceeded the value of 0.34 and the effect obtained by Keus van de Poll et al. could not be confirmed.

The information search task was not affected by the experimental condition. This is consistent with the study of Jahncke et al. (2013) who also observed no effect of speech intelligibility level on information search. It must be noted that neither Jahncke et al. nor our study included a silent condition (or STI 0.00) as a reference to different speech conditions so it cannot be concluded that speech does not affect performance in this task. However, the selected set of IE conditions in our study did not reveal any effect. It would be useful to investigate in the future whether this task is affected by speech or other IE conditions, because the task represents normal office work better than many other tasks normally used in experiments like this.

Working memory has a central role in the tasks employed in our study. However, because cognitive performance relies on several cognitive processes, it is difficult to identify all processes that may have been interfered by indoor environmental factors. In the area of noise-related performance effects, one suggestion is that the effects of noise on performance result from attentional capture rather than the impairment of other cognitive processes (Sörqvist, 2014). *Attentional capture* is caused by a sudden auditory change which draws attention from the task to the deviant event (Hughes, Vachon, & Jones, 2007). The effect of attentional capture is assumed to be more detrimental to performance when the task difficulty increases. In our study, the level of task difficulty varied between the tasks. Thus, one possible reason why the effects of conditions were not found in all tasks might be the variation in task difficulty. Because there were also other adverse factors than auditory ones, we suggest that the results cannot be explained purely by attentional capture hypothesis. Moreover, high temperature has been related with increased stress which activates attentional resources to cope with stress (Hancock et al., 2007). This has been shown to lead to a situation where the capacity to process task-relevant information is reduced (Hancock et al., 2007). Thus, all environmental factors should be observed together with the impairment of other cognitive processes when the results are interpreted.

Practical limitations of room acoustic design should also be considered when interpreting the overall effect of acoustic design on cognitive performance. Haapakangas et al. (2014) demonstrated that room acoustic design affects task performance and acoustic satisfaction but acoustic design has an effect only on speakers located at least 3–5 m from the listener. Nearby speech is difficult to control by room acoustic means. Our study confirms this finding. In our study and that of Haapakangas et al. (2014), speech intelligibility was temporally variable, because the location of the active speaker in the room varied every 6–33 s, and the distance between the speaker and the participant varied respectively. Thus, speaker's direction and distance was not constant. In both experimental conditions, task performance might have momentarily deteriorated when speech was heard from nearby desks. However, as task performance was overall better in Condition A, it may be assumed that performance was not adversely affected by speech when the speaker was far away from the participant. On the other hand, in Condition B speech was intelligible regardless of the location of the speech source. This probably explains the observed differences in working memory performance between the two experimental conditions.

In many previous studies silence and highly intelligible speech has been compared to each other (Hongisto, 2005). If the high speech intelligibility condition, Condition B, would have been compared to silence, differences in performance would probably have been more evident. However, our study involves a realistic open-plan office environment where silence is not expected to exist for very long periods of time. Therefore, it is argued that the methodologies and the findings of this experiment have better practical relevance than most previous studies except that of Haapakangas et al. (2014).

4.2. Subjective responses

The working conditions were assessed to be significantly better in Condition A than in Condition B. On the one hand, the subjective measures support the findings on performance measures discussed above. On the other hand, they demonstrate even more pervasive effects than performance measures. This is in line with previous studies demonstrating that subjective measures are more sensitive to changes in IE than objective measures (Haapakangas et al., 2011, 2014; Schlittmeier et al., 2008). In addition, despite whether or not the open-plan office conditions are objectively measured as adequate, subjective experiences have an effect on what kind of meaning the occupant will give to the positive and negative characteristics of IE (Cox & Ferguson, 1994; Lahtinen, Huuhtanen, Kähkönen, & Reijula, 2002).

4.2.1. Room acoustics

When the acoustic environment was properly designed (Condition A), acoustic satisfaction was higher and distraction from different sounds smaller. This result was found even though the overall noise level was higher in Condition A due to the masking sound.

In both experimental conditions, the most distracting sound was speech from nearby desks. Distraction by masking sound was measured by an item called "hum of ventilation" since the masking sound resembles the hum of ventilation. This was justified because the concept of sound masking is unknown among the general population. Even though the hum of ventilation (i.e., the masking sound) was perceived as more distracting in Condition A, the masking sound was more beneficial than harmful, since the most distracting sound, i.e., speech, was assessed significantly less disturbing in Condition A. Moreover, speech coming even from a distant location was rated more distracting than the hum of ventilation. Thus, it is evident from these data that speech was the main acoustic distractor. The findings are in agreement with previous studies (Haapakangas et al., 2011, 2014; Haka et al., 2009).

4.2.2. Thermal conditions and air quality

Stuffiness of the indoor air was rated as more interfering in Condition B. In the previous study (see Fig. 1), Koskela et al. (2014) did not find any significant change in experienced stuffiness when ventilation rate was reduced from 28 l/s·person (600 ppm CO_2) to 2 l/s·person (2200 ppm CO_2) while the room temperature was kept constant (23.5 °C). Instead, a significant change in stuffiness was observed by Maula et al. (2015) when high room temperature was compared with neutral temperature. This supports the previous result that when an individual feels warm in a room, air quality is also assessed to be poor (Lan et al., 2011). Based on previous results, it can be suggested that room temperature had a significant role in increased stuffiness ratings in Condition B. However, simultaneous exposure to high room temperature and poor air quality might also have an effect on stuffiness ratings together.

4.2.3. Tiredness, energy and motivation

Perceived lack of energy and motivation increased during the two-hour period, the increase being stronger in Condition B. In addition, participants were more tired in Condition B. Previously, a quite similar background speech as in condition B did not decrease arousal (i.e., tiredness) when silence was compared with a speech condition (Haapakangas et al., 2011). Similarly, Maula et al. (2015) did not find an effect of high room temperature on tiredness, energy or motivation. However, low ventilation rate (high CO_2 concentration) increased the lack of energy and motivation although ratings remained rather low while the quality assessments of IE overall remained unaffected (Koskela et al., 2014). Based on these previous findings, it seems probable that the simultaneous exposure to highly intelligible speech, high temperature and low ventilation rate intensified the lack of energy and motivation more than the exposure to only one factor at a time could cause.

4.2.4. Subjective workload

In Condition B, subjective workload was higher than in Condition A which may reflect higher effort to compensate for the anticipated performance decrement. Experience of stress may be one mechanism leading to higher effort and subsequent disruption of performance (Hancock & Warm, 1989). As the higher subjective workload coincided with the performance decrement in Condition B, it appears that compensatory efforts were needed but were not sufficient to compensate for the worsened working conditions. Moreover, the enhanced effort hypothesis is an often-mentioned explanation for the observed differences between performance and questionnaire measures (Haapakangas et al., 2011; Schlittmeier et al., 2008). Subjective experience of disturbance, such as acoustic distraction, might lead individuals to invest more effort in their performance to compensate for the effect of distraction. In our study, the maintenance of good performance with the help of higher effort might have reduced the differences in performance between conditions; yet, significant differences were nevertheless found. Similar effects of compensatory effort have been reported by other researchers (Szalma & Hancock, 2011). Although enhanced effort could have diminished performance decrement during the two-hour work period in our study, subjective workload might accumulate across longer time periods if the IE conditions causing stress persist.

In working life, individuals are exposed to specific IE conditions several hours per day. It has been shown that coping strategies are in use to combat the disturbance of sounds in open-plan office (Kaarlela-Tuomaala, Helenius, Keskinen, & Hongisto, 2009): increased effort, longer breaks, slower working rhythm and longer working days have been reported.

4.2.5. Physical symptoms

Our results showed that perceived somatic symptoms were at a lower level in condition where the IE was well designed, which is in agreement with previous findings. The prevalence of somatic symptoms was higher in Condition B although the intensity of symptoms was relatively mild. It is notable that even with a few hours of exposure the prevalence of somatic symptoms was increased (see also e.g., Lan et al., 2011).

4.2.6. Emotions

In previous studies, effects of IE on experienced emotions have been investigated very little. In our study, positive emotional responses decreased with exposure time in both experimental conditions, but emotional responses were more positive in Condition A after the second exposure hour. This is in agreement with the study of Lan et al. (2011) who found high temperature (30 °C) to be associated with negative mood. This result further reinforces the idea that indoor environment can affect emotional comfort even in the absence of somatic symptoms.

4.3. Limitations of the study

Limitations are related to methodological issues. First, participants were exposed to the experimental conditions on two separate days on subsequent weeks since it was not possible to build two open-plan offices. This might have caused measuring error despite a repeated measures design. Retest reliability of task performance, indicating the consistency of a test across time, has been indicated to differ depending on task type (Ellermeier & Zimmer, 1997), and this might be one error source.

Second, a two-hour exposure time was used. Overall, the effects of noise on task performance have been shown to diminish with exposure time (Szalma & Hancock, 2011). This may not be the case with speech sounds (Haapakangas et al., 2014; Szalma & Hancock, 2011). Instead, as argued by Haapakangas et al. (2014), it is possible that cognitive and subjective impacts of noise might even increase over time as a result of an emerging stress response or decreasing compensatory resources. Increased exposure time associated with the intensity of thermal condition has been reported to increase the negative impact of temperature on performance (Hancock et al., 2007). This is in line with the Maximal Adaptability Model (Hancock & Warm, 1989). Because activation level in our study was low, it is probable that increasing exposure time by one or two hours would not have affected the results. Instead, it would have produced practical problems, such as a need for a lunch break. Overall, it is assumed that a significant increase in the exposure time might reveal more robust effects of temperature on task performance. In addition, the duration of thermal exposure might have more impact on performance and subjective responses when work is more mobile and the activation level is higher.

Third, e.g., Hancock et al. (2007) have considered that time of day might affect the relationship between temperature and performance. In our study, the sessions were conducted in the morning to create optimal condition to perform without tiredness and thus, minimize the possibility of intervening variables. Our results should be applied with reservation when considering performance at another time of day.

Fourth, the IE conditions were planned to combine the experimental conditions of prior experiments conducted in the same laboratory (Fig. 1). However, it is possible that one IE factor dominated the effect on performance or subjective response over the other factors and had more impact on the final results.

The performance results should be applied with care to real workplace environments. Our results might only be valid for individually performed tasks in visual modality and requiring intensive use of working memory. Our results may not apply to teamwork or communication tasks. In the future it is important to also include tasks in the auditory modality. For example, it would be interesting and important to study how irrelevant background speech might affect oral communication.

5. Conclusions

This study provides strong evidence that the combination of high intelligibility of irrelevant speech, high room temperature and low ventilation rate impairs the perceived working conditions and cognitive performance. It is possible to suggest that by designing room acoustic conditions, thermal conditions and ventilation rate adequately, satisfaction with work environment is increased, somatic symptoms are decreased and the possible impairments of work performance can be avoided. The experiment was the final study in a series of four experiments. The simultaneous exposure to these three adverse factors might intensify some effects of IE. This is significant because IE problems are often multifaceted in workplaces. In practice, the ventilation rate, room temperature and acoustic conditions can vary significantly between office workplaces so that our suggestions should not be generalized to all possible configurations of indoor environment. Our study supports the view that special care should be paid to the holistic design of indoor environment in open-plan offices.

Acknowledgements

This study was a part of project national research program "TOTI – User-oriented office spaces" funded by Tekes – the Finnish Funding Agency for Technology and Innovation (TOTI Grant number 855/09, Diary number 1836/31/09), Finnish Institute of Occupational Health, and 15 companies. The materials and furniture for the office laboratory were provided by the participating companies. We wish to thank Peter Backlund for his help in air quality analyses, Jukka Keränen for room acoustic measurements, David Oliva for programming and Jarkko Hakala for the construction work. Finally, we thank Helena Jahncke (University of Gävle) for providing the materials and instructions for the information search task.

Appendix 1. Room acoustic measurements

The results of this appendix present supportive data to understand how the experimental conditions related to typical conditions found in offices (see e.g., the measurements by Virjonen et al., 2009), and how the STI and speech levels behaved at different distances from a speaker. The measurements were taken according to ISO 3382–3:2012.

The measurements were taken by placing an omni-directional loudspeaker at one corner desk (source height 120 cm) and by measuring the level of speech, background noise and STI in other workstations (measurement height 120 cm). The spatial decays of A-weighted speech level and STI were determined (Fig. A1) from which the single-number quantities were derived (Table A1). The spatial decay rate of speech, $D_{2,S}$, describes how many decibels the A-weighted speech level reduces when the distance to the speaker is doubled. Large value of $D_{2,S}$ indicates the low disturbance of the speech. Field measurements have shown values between 4 and 12 dB (Virjonen et al., 2009). The distraction distance, r_D, describes how far from the speaker the speech is still perfectly intelligible. It is defined as the distance within which the value of STI is above 0.50. A small value of r_D predicts low disturbance caused by surrounding speech. Field measurements have shown values between 5 and 19 m (Virjonen et al., 2009). Fig. A1 and Table A1 indicate that Condition A and B are significantly different in respect with the objective disturbance of the speech.

It should be noted that the measurements were carried out using normal speech effort. The speech effort in the psychological experiment was 4.5 dB lower. Therefore, the levels in Fig. A1 are a little higher than in Fig. 3 and Table 1.

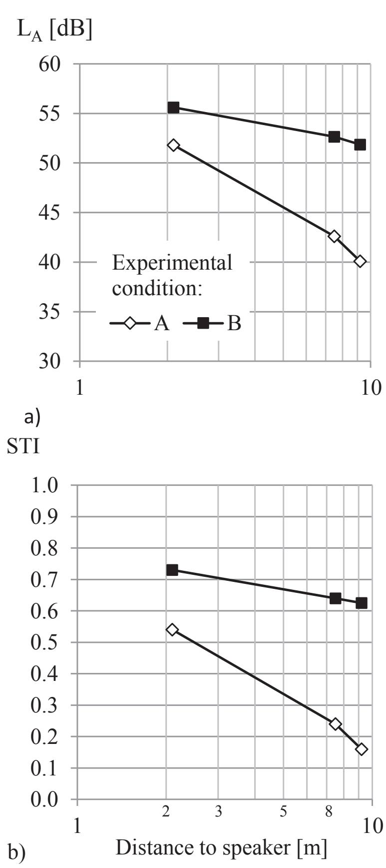

Fig. A1. Spatial attenuation of a) A-weighted level of speech, and b) Speech Transmission Index based on ISO 3382–3:2012 measurements.

Table A1
The values of the room acoustic single-number quantities in the two experimental conditions. [a] Extrapolated value.

	Experimental condition	
	A	B
$D_{2,S}$ [dB]	5.5	1.8
r_D [m]	2.9	24.1[a]

References

Åhsberg, E., Gamberale, F., & Gustafsson, K. (1998). *Upplevd trötthet efter mentalt arbete. En experimentell utvärdering av ett mätinstrument, 8*. Arbete och Hälsa (In Swedish).
Åhsberg, E., Gamberale, F., & Kjellberg, A. (1995). *Upplevd trötthetskvalitet vid olika arbetsuppgifter. Utveckling av ett mätinstrument, 20*. Arbete och Hälsa (In Swedish).
Allen, T. J., & Gerstberger, P. G. (1973). A field experiment to improve communication in a product engineering department: the non-territorial office. *Human Factors, 15*, 487–498.
Apte, M. G., Fisk, J. W., & Daisey, J. M. (2000). Indoor carbon dioxide concentrations and SBS in office workers. In *Proceedings of healthy buildings. Espoo, Finland*.
ASHRAE-62.1. (2010). *Ventilation for acceptable indoor air quality*. Atlanta, GA: American Society of Heating, Refrigerating and Air-Conditioning Engineers.
Bakó-Biró, Z., Clements-Croome, D. J., Kochhar, N., Awbi, H. B., & Williams, M. J. (2012). Ventilation rates in schools and pupils' performance. *Building and Environment, 48*, 215–223.
Balazova, I., Clausen, G., Rindel, J. H., Poulsen, T., & Wyon, D. P. (2008). Open-plan office environments: a laboratory experiment to examine the effects of office noise and temperature on human perception, comfort and office work performance. In *Proceedings of indoor air 2008. Copenhagen, Denmark*.
Balazova, I., Clausen, G., & Wyon, D. P. (2007). The influence of exposure to multiple indoor environmental parameters on human perception, performance and motivation. In *Proceedings of Clima 2007 wellbeing indoors. Helsinki, Finland*.
Banbury, S., & Berry, D. C. (1998). Disruption of office-related tasks by speech and office noise. *British Journal of Psychology, 89*, 499–517.
Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B (Methodological), 57*, 289–300.
Bodin Danielsson, C., Chungkham, H. S., Wulff, C., & Westerlund, H. (2014). Office design's impact on sick leave rates. *Ergonomics, 57*, 139–147.
Bradley, J. S. (2003). The acoustical design of conventional open plan offices. *Canadian Acoustics, 27*, 23–31.
Buchner, A., Steffens, M. C., Irmen, L., & Wender, K. F. (1998). Irrelevant auditory material affects counting. *Journal of Experimental Psychology: Learning, Memory and Cognition, 24*, 48–67.
Clausen, G., & Wyon, D. P. (2008). The combined effects of many different indoor environmental factors on acceptability and office work performance. *HVAC & R Research, 14*, 103–113.
Colle, H. A. (1980). Auditory encoding in visual short-term recall: effects of noise intensity and spatial location. *Journal of Verbal Learning and Verbal Behavior, 19*, 722–735.
Cox, T., & Ferguson, E. (1994). Measurement of the subjective work environment. *Work & Stress, 8*, 98–109.
Danielsson, C. (2005). *Office environment, health and job-satisfaction*. Licentiate Thesis. Stockholm, Sweden: KTH Technology and Health.
De Croon, E., Sluiter, J., Kuijer, P. P., & Frings-Dresen, M. (2005). The effects of office concepts on worker health and performance: a systematic review of the literature. *Ergonomics, 48*, 119–134.
Ellermeier, W., & Hellbrück, J. (1998). Is level irrelevant in irrelevant speech? Effects of loudness, signal-to-noise ratio, and binaural unmasking. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 1406–1414.
Ellermeier, W., & Zimmer, K. (1997). Individual differences in susceptibility to the "irrelevant speech effect". *The Journal of the Acoustical Society of America, 102*, 2191–2199.
EN ISO 11654. (1997). *Acoustics — Sound absorbers for use in buildings — f1Rating of*

sound absorption.
Evans, G. W., & Johnson, D. (2000). Stress and open-office noise. *Journal of Applied Psychology, 85*, 779–783.
Frontczak, M., Schiavon, S., Goins, J., Arens, E., Zhang, H., & Wargocki, P. (2012). Quantitative relationships between occupant satisfaction aspects of indoor environmental quality and building design. *Indoor Air, 22*, 119–131.
Gray, J. R., Chabris, C. F., & Braver, T. S. (2003). Neural mechanisms of general fluid intelligence. *Nature Neuroscience, 6*, 316–322.
Haapakangas, A., Helenius, R., Keskinen, E., & Hongisto, V. (2008). Perceived acoustic environment, work performance and well-being – survey results from Finnish offices. In *Proceedings of the 9th International Congress on Noise as a Public Health Problem (ICBEN) 2008* (pp. 434–441) (Mashantucket, Connecticut, USA).
Haapakangas, A., Hongisto, V., Hyönä, J., Kokko, J., & Keränen, J. (2014). Effects of unattended speech on performance and subjective distraction: the role of acoustic design in open-plan offices. *Applied Acoustics, 86*, 1–16.
Haapakangas, A., Kankkunen, E., Hongisto, V., Virjonen, P., Oliva, D., & Keskinen, E. (2011). Effects of five speech masking sounds on performance and acoustic satisfaction. Implication for open-plan offices. *Acta Acustica United with Acustica, 97*, 641–655.
Häggblom, H., Hongisto, V., Haapakangas, A., & Koskela, H. (2011). The effect of temperature on work performance and thermal comfort – laboratory experiment. In *Proceedings of indoor air 2011, the 12th International Conference on Indoor Air Quality and Climate. Austin, Texas*.
Haka, M., Haapakangas, A., Keränen, J., Hakala, J., Keskinen, E., & Hongisto, V. (2009). Performance effects and subjective disturbance of speech in acoustically different office types – a laboratory experiment. *Indoor Air, 19*, 454–467.
Hancock, P. A., Ross, J. M., & Szalma, J. L. (2007). A meta-analysis of performance response under thermal stressors. *Human Factors, 49*, 851–877.
Hancock, P. A., & Warm, J. S. (1989). A dynamic model of stress and sustained attention. *Human Factors, 31*, 519–537.
Hardy, H. C. (1957). A guide to office acoustics. *Architectural Record, 121*, 235–240.
Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index). Results of empirical and theoretical research. In P. A. Hancock, & N. Meshkati (Eds.), *Human mental workload* (pp. 139–183). Amsterdam: Elsevier Science Publishers.
Hockey, R. (1984). Varieties of attentional state: the effects of environment. In R. Parasuraman, & D. R. Davies (Eds.), *Varieties of attention* (pp. 449–483). New York: Academic Press.
Hongisto, V. (2005). A model predicting the effect of speech of varying intelligibility on work performance. *Indoor Air, 15*, 458–468.
Hongisto, V. (2008). *Effect of sound masking on workers in an open office*. Acoustics'08 (pp. 537–542) (Paris, France).
Hongisto, V., Haapakangas, A., Helenius, R., Keränen, J., & Oliva, D. (2012). *Acoustic satisfaction in an open-plan office before and after the renovation* (pp. 654–659). Prague, Czech Republic: Euronoise.
Hongisto, V., Oliva, D., & Rekola, L. (2015). Subjective and objective rating of spectrally different pseudorandom noises – implications for speech masking design. *The Journal of the Acoustical Society of America, 137*, 1344–1355.
Hughes, R. W., Vachon, F., & Jones, D. M. (2007). Disruption of short-term memory by changing and deviant sounds: support for a duplex-mechanism account of auditory distraction. *Journal of Experimental Psychology: Learning, Memory and Cognition, 33*, 1050–1061.
Hundert, A. T., & Greenfield, N. (1969). Physical space and organizational behavior: a study of office landscape. In *Proceedings of the 77th annual convention of the American Psychological Association* (pp. 601–602).
Hygge, S., & Knez, I. (2001). Effects of noise, heat and indoor lighting on cognitive performance and self-reported affect. *Journal of Environmental Psychology, 21*, 291–299.
ISO 3382–3. (2012). *Acoustics – Measurement of room acoustic parameters – Part 3: Open-plan offices*. Geneva, Switzerland: International Organization for Standardization.
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.
ISO 7730. (2005). *Ergonomics of the thermal environment. Analytical determination and interpretation of thermal comfort using calculation of the PMV and PPD indices and local thermal comfort criteria*. Geneva, Switzerland: International Organization for Standardization.
Jahncke, H., & Halin, N. (2012). Performance, fatigue and stress in open-plan offices: the effects of noise and restoration on hearing impaired and normal hearing individuals. *Noise and Health, 14*, 260–272.
Jahncke, H., Hongisto, V., & Virjonen, P. (2013). Cognitive performance during irrelevant speech: effects of speech intelligibility and semantic difficulty. *Applied Acoustics, 74*, 307–316.
Jones, D., Madden, C., & Miles, C. (1992). Privileged access by irrelevant speech to short-term memory: the role of changing state. *The Quarterly Journal of Experimental Psychology, 44A*, 645–669.
Juvonen, R. (2012a). *Kultainen syksypuu*, 4.5.12 (In Finnish) www.lastenmaa.net/00010592-kultainen-syksypuu.
Juvonen, R. (2012b). *Pieni merihevonen*, 4.5.12 (In Finnish) www.lastenmaa.net/00010594-pieni-merihevonen.
Kaarlela-Tuomaala, A., Helenius, R., Keskinen, E., & Hongisto, V. (2009). Effects of acoustic environment on work in private office rooms and open-plan offices – longitudinal study during relocation. *Ergonomics, 52*, 1423–1444.
Keränen, J., & Hongisto, V. (2013). Prediction of the spatial decay of speech in open-plan offices. *Applied Acoustics, 74*, 1315–1325.
Keus van de Poll, M., Ljung, R., Odelius, J., & Sörqvist, P. (2014). Disruption of writing by background speech: the role of speech transmission index. *Applied Acoustics, 81*, 15–18.
Koskela, H., Maula, H., Haapakangas, A., Moberg, V., & Hongisto, V. (2014). Effect of low ventilation rate on office work performance and perception of air quality – a laboratory study. In *Proceedings of indoor air 2014, Hong Kong*.
Lahtinen, M., Huuhtanen, P., Kähkönen, E., & Reijula, K. (2002). Psychosocial dimensions of solving and indoor air problems. *Indoor Air, 12*, 33–46.
Lan, L., Wargocki, P., Wyon, D. P., & Lian, Z. (2011). Effects of thermal discomfort in an office on perceived air quality, SBS symptoms, physiological responses, and human performance. *Indoor Air, 21*, 376–390.
LVI 05–10440 en. (2008). *Classification of indoor environment*. Helsinki, Finland: Rakennustieto Publishing Ltd (In English).
Maula, H., Hongisto, V., Östman, L., Haapakangas, A., Koskela, H., & Hyönä, J. (2015). *The effect of slightly warm temperature on work performance and comfort in open-plan offices – A laboratory study*. http://dx.doi.org/10.1111/ina.12209. *Indoor Air*, published online.
Moroney, W. F., Biers, D. W., & Eggemeier, F. T. (1995). Some measurement and methodological considerations in the application of subjective workload measurement techniques. *The International Journal of Aviation Psychology, 5*, 87–106.
Park, J. S., & Yoon, C. H. (2011). The effects of outdoor air supply rate on work performance during 8-h work period. *Indoor Air, 21*, 284–290.
Pejtersen, J., Allermann, L., Kristensen, T. S., & Poulsen, O. M. (2006). Indoor climate, psychosocial work environment and symptoms in open-plan offices. *Indoor Air, 16*, 392–401.
Pejtersen, J. H., Feveile, H., Christensen, K. B., & Burr, H. (2011). Sickness absence associated with shared and open-plan offices – a national cross sectional questionnaire survey. *Scandinavian Journal of Work, Environment and Health, 37*, 376–382.
Pilcher, J. J., Nadler, E., & Busch, C. (2002). Effects of hot and cold temperature exposure on performance: a meta-analytic review. *Ergonomics, 45*, 682–698.
Rumelhart, D. E., & Norman, D. A. (1982). Simulating a skilled typist: a study of skilled cognitive-motor performance. *Cognitive Science, 6*, 1–36.
Satish, U., Mendell, M. J., Sheklar, K., Hotchi, T., Sullivan, D., Streufert, S., et al. (2012). Is CO_2 an indoor pollutant? direct effects of low-to-moderate CO_2 concentrations on human decision-making performance. *Environmental Health Perspectives, 120*, 1671–1677.
Schlittmeier, S. J., Hellbrück, J., Thaden, R., & Vorländer, M. (2008). The impact of background speech varying in intelligibility: effects on cognitive performance and perceived disturbance. *Ergonomics, 51*, 719–736.
Schutte, M., Marks, A., Wenning, E., & Griefahn, B. (2007). The development of the noise sensitivity questionnaire. *Noise and Health, 9*, 15–24.
Seppänen, O. A., Fisk, W. J., & Mendell, M. J. (1999). Association of ventilation rates and CO_2 concentrations with health and other responses in commercial and institutional buildings. *Indoor Air, 9*, 226–252.
Sörqvist, P. (2014). On interpretation and task selection in studies on the effects of noise on cognitive performance. *Frontiers in Psychology, 5*, 1–4.
Sörqvist, P., Nöstl, A., & Halin, N. (2012). Disruption of writing processes by the semanticity of background speech. *Scandinavian Journal of Psychology, 53*, 97–102.
Strömqvist, S., & Karlsson, H. (2002). *ScriptLog for windows – User's manual*. Technical Report. University of Lund: Department of Linguistics and University College of Stavanger: Centre for reading Research.
Szalma, J. L., & Hancock, P. A. (2011). Noise effects on human performance: a meta-analytic synthesis. *Psychological Bulletin, 137*, 682–707.
Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language, 28*, 127–154.
Veitch, J. A., Charles, K. E., Farley, K. M. J., & Newsham, G. R. (2007). A model of satisfaction with open-plan office conditions: COPE field findings. *Journal of Environmental Psychology, 27*, 177–189.
Venetjoki, N., Kaarlela-Tuomaala, A., Keskinen, E., & Hongisto, V. (2006). The effects of speech and speech intelligibility on task performance. *Ergonomics, 49*, 1068–1091.
Virjonen, P., Keränen, J., & Hongisto, V. (2009). Determination of acoustical conditions in open-plan offices: proposal for new measurement method and target values. *Acta Acustica United with Acustica, 95*, 279–290.
Wargocki, P., Bakó-Biró, Z., Clausen, G., & Fanger, P. O. (2002). Air quality in a simulated office environment as a result of reducing pollution source and increasing ventilation. *Energy and Buildings, 34*, 775–783.
Wargocki, P., Wyon, D. P., Baik, Y. K. B., Clausen, G., & Fanger, P. O. (1999). Perceived air quality, sick building syndrome (SBS) symptoms and productivity in an office with two different pollution loads. *Indoor Air, 9*, 165–179.
Wargocki, P., Wyon, D. P., Sundell, J., Clausen, G., & Fanger, P. O. (2000). The effects of outdoor air supply rate in an office on perceived air quality, sick building syndrome (SBS) symptoms and productivity. *Indoor Air, 10*, 222–236.
Weinstein, N. D. (1978). Individual differences in reactions to noise: a longitudinal study in a college dormitory. *Journal of Applied Psychology, 63*, 458–466.
Witterseh, T., Wyon, D. P., & Clausen, G. (2004). The effects of moderate heat stress and open-plan office noise distraction on SBS symptoms and on the performance of office work. *Indoor Air, 14*, 30–40.
Zuckerman, M. (1977). Development of a situation-specific trait-state test for the prediction and measurement of affective responses. *Journal of Consulting and Clinical Psychology, 45*, 513–523.