DOI: 10.61782/fa.2023.1134

forum acusticum 2023

Investigating the effect of realistic sound absorption variations in open-plan office simulations on short-term memory performance

Manuj Yadav[1*], Markus Georgi[2], Larissa Leist[3], Maria Klatte[3], Sabine J Schlittmeier[2], Janina Fels[1]

[1] Institute for Hearing Technology and Acoustics, RWTH Aachen University, Aachen, Germany
[2] Institute of Psychology, RWTH Aachen University, Aachen, Germany
[3] Center for Cognitive Science, RPTU Kaiserslautern-Landau, Kaiserslautern, Germany

ABSTRACT

Extant room acoustic design recommendations for open-plan offices (OPOs) include combined use of sound absorption and electroacoustic masking. However, some design trends including activity-based working sometimes eschew sound absorption. Besides, some studies suggest that increased reverberation times could sufficiently reduce changing-state characteristics of irrelevant OPO sounds, allowing short-term memory (STM) performance similar to quiet conditions. This study included participants (n=40) performing an auditory-verbal serial recall task in silence, and in conditions with irrelevant sounds (multi-talker speech and common OPO nonspeech sounds) reproduced binaurally over headphones in three reverberant conditions typical in medium-sized OPOs (mid-frequency T_{30} (s) = 0.4, 0.8, 1.1), and in an anechoic condition. Bayesian mixed-effects modelling was used with sound conditions and percentage of serial recall errors (err%) as the independent and dependent variables, respectively. The results showed err% increased in all conditions relative to *Quiet* and higher err% in T_{30} = 0.4 s vs. T_{30} = 0.8 and 1.1 s conditions. The err% differences between the latter two conditions were not statistically robust. This suggests cognitive impairment regardless of typical T_{30} in OPOs. Moreover, STM performance with inordinately high T_{30} may not reach performance in quiet, reinforcing careful considerations of sound absorption usage in OPOs.

Keywords: *open-plan offices, short-term memory, sound absorption, reverberation time, activity-based working*

*[*]**Corresponding author**: manuj.yadav@akustik.rwth-aachen.de*

1. INTRODUCTION

Task-irrelevant sounds demonstrably impair cognitive performances in certain tasks. This phenomenon has been studied extensively for verbal short-term memory (STM) performance, typically using serial recall tasks, and is nowadays referred to as the irrelevant sound effect (ISE; [1]). Herein, the presence of to-be-ignored, task-irrelevant *changing-state* background sound stream has been shown to impede verbal STM performance compared to steady-state sound stream and/or quiet conditions.

A model based on verbal STM serial recall performance is even integral to ISO 3383-3, which is the room acoustic standard for open-plan offices (OPOs) [2]. ISO 3382-3 recommends simultaneous use of sound absorption and sound masking to achieve adequate spatial decay of speech sound pressure level (SPL) and speech intelligibility [2]. However, it is not uncommon for many contemporary office designs to eschew sound absorption to varying degrees for more 'open/lively' workspaces, including those based around activity-based working (ABW) [3].

Moreover, previous studies have explored reduced changing-state characteristics – hence, potentially increased verbal STM performance – due to spectrotemporal smearing of irrelevant multi-talker speech with relatively high reverberation times (e.g., low sound absorption) in OPOs [4], [5]. Interestingly, for a condition with 15 voices and reverberation time (T in s; presumably broadband) of 1.0 s, serial recall task performance was not significantly different to the quiet condition in Vachon et al. [5]. According to the authors, this suggested sufficiently reduced changing-state characteristics, and recommended similarly long reverberation times as a possible solution to speech distraction in OPOs, which typically have several voices simultaneously active as in their study design [5].

However, these studies [4], [5] do not consider realistic spatial arrangement of sound sources in OPOs, wherein most speech-based disruption tends to be due to intelligible speech from spatially separated nearby workstations [3], [6], [7]. Besides, T around 1.0 s may represent a relatively high value for most medium-large sized OPOs [3].

Hence, both from conceptual (e.g., changing-state characteristics) and practical purposes (e.g., considering ABW offices), it is worth investigating the role of reasonable room absorption variations in realistic OPO simulations. Therefore, this study aims to study the ISE within a simulated medium sized room with variable acoustic profiles, and with spatially spread multi-talker speech and nonspeech sources resembling OPOs.

2. METHODS

2.1 Irrelevant OPO background sounds

Table 1. Summary of experimental conditions. SPL is power-averaged (over left and right ears) $L_{A,eq,1\ min}$. Mid-frequency T_{30} (reverberation time) and C_{50} (clarity index) presented are room averaged values. STI is averaged over each talker-receiver configuration.

Exp. condition	SPL (dB(A))	T_{30} (s)	C_{50} (dB)	STI
Quiet	41.5	-	-	-
Anechoic	55	-	-	~ 1.0
RA-1	55	0.4	13.1	0.73
RA-2	55	0.7	5.3	0.64
RA-3	55	1.1	2.9	0.57

For representative *HVAC noise*, pink noise shaped with a -5dB/octave decay was presented at $L_{A,eq,1min}$ (energy-equivalent A-weighted SPL over a 1-minute period) of 41.5 dB in all conditions in Table 1. The *nonspeech sounds* included common activity sounds in OPOs (e.g., workstation sounds, phone rings, footsteps, etc.). Each type of sound was organized in a quasi-random manner over time to resemble nonspeech sound streams in OPOs. The speech material was based on two-person conversations in German recorded in hemi-anechoic conditions. These were processed to create 4-channel (gender balanced) *multi-talker speech* where two voices were active at any time as halfalogues (one side of a conversation), based on [8].

2.2 Simulated binaural OPO scenarios

A room with a rectangular floor plate, and 12 workstations (over four rows) was modelled in Sketchup (267.3 m^3; 9 m×11 m×2.7 m). This model was imported within the geometrical room acoustics simulation software RAVEN [9]. The listener was modelled at the middle workstation on the last row and assigned generic head-related transfer function (HRTF) and headphone transfer functions.

There were 5 experimental conditions in total (Table 1). This included the *Quiet* baseline, an *Anechoic* condition, and three scenarios (Table 1) with reverberation times (T_{30} in seconds) spanning a mid-frequency (average of 500 Hz – 2 kHz octave-band center frequencies) T_{30} range of 0.4 s – 1.1 s typical of most OPOs [3], and comparable with similar T range of 0.4 s – 1 s in previous studies [4], [5]. Besides the reverberation time, the clarity index (C_{50} in dB) was kept within an acceptable range. Note that the speech transmission index (STI) values in Table 1 are for illustrative purposes only, as STI is not well defined for fluctuating background noise.

For each room acoustic condition (Table 1), there were 16 sound sources and corresponding binaural room impulse responses (BRIRs): 12 nonspeech sources (one per workstation) with omnidirectional directivities, and 4 speech sources (at 3 and 5 m, respectively, in front of the listener) with directivity of a singer. The BRIRs were convolved with the respective audio, which resulted in auralized files for 12 nonspeech and 4 speech sources per experimental condition in Table 1, except for *Quiet*.

Hence, each convolved file had various nonspeech sounds occurring quasi-randomly from the various workstation location, and non-repeating halfalogue speech from any two talkers active simultaneously from different workstations. Overall, this is a close representation of spatially separated multi-talker speech and nonspeech sounds in actual OPOs. All conditions had low SPL HVAC noise (section 2.1).

2.3 Serial recall digits

The serial recall task used included spoken digits as the to-be-recalled items instead of the more common visual presentation of digits. This was done as a first order approximation of processing spoken information during irrelevant background sounds, which is a common scenario in OPOs. In this regard, serial recall of aurally presented sequences has previously been shown to exhibit a similar effect as serial recall of visual information [10]. The to-be remembered digits were recorded as spoken digits (0-9) in German by trained professional speakers in an anechoic chamber using normal intonation and constant vocal effort. Each (anechoic) digit was 0.6 s in duration in a female voice and presented diotically.

2.4 Calibration

The auralized files with irrelevant sounds (section 2.2) were mixed down to a 2-channel audio file per experimental condition (Table 1) except for *Quiet*. The headphone

(Sennheiser HD650) output per channel for each experimental condition and each digit was then calibrated using an artificial ear [11]. The overall presentation level of the experimental conditions, except *Quiet*, was $L_{A,eq,1min}$ of 55 dB (Table 1). The to-be-recalled digits were always presented diotically at 61 dB $L_{A,Fmax}$ in all conditions.

2.5 Experimental procedure

The experiment was conducted in a soundproofed hearing booth and lasted approximately an hour per participant. There were 40 participants (23 F, Age (years): *mean* = 25.3, *sd* = 6.4), all German native speakers. In each serial recall trial, participants heard eight digits as to-be-recalled stimulus, chosen without repetition from 0-9 (section 2.3). Each digit was 0.6 s in duration with a 1.5 s interval between digits. There were 12 trials per experimental condition, hence 60 trials overall (Table 1). Participants were instructed to remember the digits in the order of presentation, and to ignore the other sounds. There was a 6 s retention period after digit presentation, following which the participants recalled the digits by selecting (via mouse clicks) corresponding digits in a graphical user interface, where, once selected, a digit could not be selected again.

2.6 Data analysis

The statistical analysis was conducted using the software R (version 4.2.2). Per experimental condition, an incorrect response was registered for every digit recalled incorrectly in the serial order. Bayesian mixed-effects models were used to account for the repeated-measures design, with the experimental conditions as the independent variable and err% (percentage of serial recall errors per condition) as the dependent variable. The Bayesian modelling was done using the *brms* (version 2.18) package [12] with mildly informative conservative priors.

3. RESULTS AND GENERAL DISCUSSION

As seen in Table 2, serial recall performance in the *Quiet* condition was better (statistically robust differences) than other conditions and hence, exhibited the ISE. As evidenced by err% differences between *Anechoic* and RA-1 (Table 2) that were not statistically robust, a room with relatively high sound absorption but also high speech clarity may not represent the best acoustic design choice for STM performance. This is consistent with the recommendations in standards wherein speech clarity/intelligibility reduction between OPO workstations is the key consideration [2].
In an OPO simulation with spatially fused multi-talker speech Vachon et al. [5], had reported ISE in a condition with T = 0.4 s, which is consistent in principle with the current results regarding RA-1. While err% decreased with reduced sound absorption in RA-2 and RA-3 relative to RA-1, also consistent with [5] at least in terms of the direction of the effect, err% differences between RA-2 and RA-3 were not statistically robust, which is consistent in principle with [4] wherein serial recall performance in T = 0.7 s vs. 0.9 s conditions did not vary significantly. However, err% in RA-3 was still higher (and statistically robust) compared to *Quiet*. This is not consistent with the main finding in [5] wherein no ISE was exhibited in a condition with T = 1.0 s and with irrelevant speech from 15 voices at a 10 m distance.

Table 2: Median predicted differences of err% between the experimental conditions and the 95% Bayesian credible intervals (CIs) in brackets. Statistically robust differences highlighted in bold.

Exp. condition	*Anechoic*	RA-1	RA-2	RA-3
Quiet	**-12.3 [-15.5,-9.1]**	**-10.6 [-13.9,-7.5]**	**-6.5 [-9.7,-3.4]**	**-6.5 [-9.7,-3.4]**
Anechoic	-	1.6 [-1.8,5.1]	**5.8, [2.4,9.2]**	**5.8 [2.4,9.2]**
RA-1	-	-	**4.1, [0.7,7.5]**	**4.1 [0.8,7.6]**
RA-2	-	-	-	0.03 [-3.3,3.5]

Even though both Perham et al. [4] and Vachon et al. [5] used a range of T similar to the current T_{30}, they have a rather limited and ambiguous representation of room acoustics and speech sources in OPO simulations, which lead to different conclusions. Further, the speech in Vachon et al. [5] did not include nearby talkers, whose intelligible speech is generally considered the primary source of auditory distraction in OPOs [3], [6], [7]. Hence, the current results can be considered a more ecologically valid comparison between room acoustic conditions that are more representative of OPO scenarios. As such, the current findings disagree with the (rather unintuitive) suggestion in Vachon et al. [5] that increased T_{30} (around 1.0 s) with distant spatially-fused multi-talker 'babble' can exhibit serial recall performance similar to quiet conditions in OPOs. Besides, long reverberation times are generally avoided to control detrimental effects of excess reverberant energy which can contribute towards noise. Moreover, a careful *combination* of various room acoustic criteria is needed to achieve 'good' acoustics in OPOs rather than just reverberation times [2], [6], [7].

The situation in RA-2 and especially RA-3 depicts recent trends where screen/partition and sound absorption in general are eschewed including within OPOs supporting ABW [3]. The current results do not support such practices (without considerations such as sound masking, work type, etc. [13]) since STM performance in conditions with/without high sound absorption but without screens/partitions may still exhibit substantial ISE. In RA-2 and RA-3, the STM task performance was almost the same and while higher T_{30} along with lower C_{50} than RA-3 may be achievable, it would perhaps be unreasonable for an OPO of current volume. RA-2 may represent a compromise with serial recall performance similar to RA-3 but lower (statistically robust) than RA-1. However, the 'sweet spot' for OPO room acoustics is a much broader topic (see [2, p. 338]), which needs to be tested using serial recall and other OPO tasks (e.g., writing, etc.) with variable room acoustics.

4. CONCLUSIONS

The findings here do not endorse exorbitant reverberant conditions as solutions to speech distraction, without additional room acoustic and workplace design considerations, as noted previously elsewhere [2], [3], [6], [13]. This has relevance for workplaces in general including ABW offices that typically include open-plan spaces with limited acoustic treatment. The current results however are limited to medium-sized OPOs. Future studies can explore larger offices and improve upon the ecological validity with more realistic simulations (e.g., with headtracking).

5. ACKNOWLEDGMENTS

This study was funded through the German Research Foundation (DFG) Research Grant scheme (Project number: 401278266).

6. REFERENCES

[1] W. Ellermeier and K. Zimmer, "The psychoacoustics of the irrelevant sound effect," *Acoustical Science and Technology*, vol. 35, no. 1, pp. 10–16, 2014, doi: 10.1250/ast.35.10.

[2] ISO 3382-3, "ISO 3382-3:2022. Measurement of room acoustic parameters. Part 3: Open plan offices," International Organization for Standardization, Geneva, Switzerland, 2022.

[3] M. Yadav, D. Cabrera, J. Kim, J. Fels, and R. de Dear, "Sound in occupied open-plan offices: Objective metrics with a review of historical perspectives," *Applied Acoustics*, vol. 177, p. 107943, Jun. 2021.

[4] N. Perham, S. Banbury, and D. M. Jones, "Do realistic reverberation levels reduce auditory distraction?," *Applied Cognitive Psychology*, vol. 21, no. 7, pp. 839–847, 2007.

[5] F. Vachon, E. Winder, M. Lavandier, and R. W. Hughes, "The bigger the better and the more the merrier? Realistic office reverberation levels abolish cognitive distraction by multiple-voice speech," in *12th ICBEN*, Zurich, Aug. 2017, p. 10.

[6] A. Haapakangas, V. Hongisto, M. Eerola, and T. Kuusisto, "Distraction distance and perceived disturbance by noise—An analysis of 21 open-plan offices," *JASA*, vol. 141, no. 1, pp. 127–136, Jan. 2017.

[7] M. Yadav *et al.*, "Investigating noise disturbance in open-plan offices using measurements of the room acoustics, and of the sound environment during occupancy," in *Proceedings of Euronoise 2021*, Madeira, Portugal, Oct. 2021.

[8] M. Yadav and D. Cabrera, "Two simultaneous talkers distract more than one in simulated multi-talker environments, regardless of overall sound levels typical of open-plan offices," *Applied Acoustics*, vol. 148, pp. 46–54, May 2019.

[9] D. Schröder, *Physically based real-time auralization of interactive virtual environments*, vol. 11. Logos Verlag Berlin GmbH, 2011.

[10] S. J. Schlittmeier, J. Hellbrück, and M. Klatte, "Does irrelevant music cause an irrelevant sound effect for auditory items?," *European Journal of Cognitive Psychology*, vol. 20, no. 2, pp. 252–271, Mar. 2008.

[11] IEC 60318-1, "Simulators of human head and ear — Part 1: Ear simulator for the calibration of supra-aural earphones," International Electrotechnical Commission, Geneva, Switzerland, 1998.

[12] P.-C. Bürkner, "brms: An R Package for Bayesian Multilevel Models Using Stan," *Journal of Statistical Software*, vol. 80, pp. 1–28, Aug. 2017.

[13] ISO 22955, "ISO 22955 Acoustics — Acoustic quality of open office spaces." International Organization for Standardization, Geneva, Switzerland, 2021.