International Journal of Hygiene and Environmental Health 275 (2026) 114809

Contents lists available at ScienceDirect

International Journal of Hygiene and Environmental Health

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

Do carbon dioxide, volatile organic compounds and atmospheric pressure affect the cognitive performance of occupants in indoor environments? Results of a large-scale experiment with simulated flights

Britta Herbig [a,*], Florian Mayer [b], Victor Norrefeldt [b], Pawel Wargocki [c]

[a] Institute and Clinic for Occupational, Social and Environmental Medicine, LMU University Hospital, LMU Munich, Ziemssenstr. 5, 80336, München, Germany
[b] Fraunhofer Institute for Building Physics IBP, Holzkirchen Branch, Fraunhoferstraße 10, 83626, Valley, Germany
[c] International Centre for Indoor Environment and Energy, Department of Environmental and Resource Engineering, Technical University of Denmark, Anker Engelunds Vej 1, Bygning 101A, 2800, Kgs. Lyngby, Denmark

ARTICLE INFO

Keywords:
Indoor air quality
Carbon dioxide
Total volatile organic compounds
Atmospheric pressure
Cognitive domains
Performance

ABSTRACT

Background: Indoor air quality, especially carbon dioxide (CO_2) and volatile organic compounds (VOCs), has been linked to cognitive impairment in previous studies, even below guideline levels. Most previous studies employed small cross-over designs and varied only one indoor air factor.

Objective: This study aimed to investigate the influence of various combinations of indoor air parameters at typical and elevated levels on cognitive performance.

Methods: We conducted a randomized controlled trial with 398 healthy adults aged 18-69 years. Participants were randomly assigned to 11 groups and exposed to different combinations of pure CO_2 (1200, 2750, 4200 ppm), atmospheric pressure (940, 755 hPa), and a VOC mixture of selected compounds at low, medium, and high concentrations. Eight different cognitive domains were assessed with established tests.

Results: Participants rated air quality as good to very good at the end of exposure in all groups. Analyses of variance with pairwise comparisons showed effects of exposure in 2 out of 8 cognitive domains (concentration and figural working memory). None of the observed differences were systematically aligned with the air quality conditions.

Conclusion: Our study provides no evidence that short-term exposure to various combinations of CO_2 (up to 4200 ppm), TVOCs (up to 2100 μg/m^3), and atmospheric pressure up to 8000 ft. elevation had any systematic negative effects on the cognitive abilities of healthy adults. The TVOC results should be viewed with caution, as a specific mixture was used, and other mixtures may occur in other indoor environments. Rigorously designed studies with longer exposure times are recommended.

1. Introduction

Carbon dioxide (CO_2) has long been used as a proxy for indoor air quality. Other air components that are considered potentially more relevant to health and well-being, such as volatile organic compounds (VOCs), are associated with CO_2 and increase with it. Current CO_2 levels in nonindustrial buildings are often above the most common 1000 parts per million (ppm) guideline value (Mendell et al., 2024) but rarely, if ever, exceed the maximum workplace concentration of 5000 ppm (OSHA, 2020). On the other hand, numerous VOCs are measured at low levels, often parts per trillion (ppt), and many of them exceed their reported odor and sometimes even irritation thresholds and health guideline values. High levels of VOCs are caused by emission sources from materials used in buildings, but also by people and their activities, or may be the result of chemical transformations.

Only in recent years has indoor CO_2 itself come into focus as a factor that can impair cognitive performance (even at concentrations below guideline values). As studies vary greatly in terms of CO_2 exposure levels, the creation and application of exposure, and the cognitive domains tested, reviews, some systematic, and meta-analyses have produced inconsistent results and emphasized different conclusions. Lax (2020) reported in general no and even one positive effect of CO_2 on

* Corresponding author.
E-mail addresses: britta.herbig@med.uni-muenchen.de (B. Herbig), florian.mayer@ibp.fraunhofer.de (F. Mayer), victor.norrefeldt@ibp.fraunhofer.de (V. Norrefeldt), pawar@dtu.dk (P. Wargocki).

https://doi.org/10.1016/j.ijheh.2026.114809
Received 11 February 2026; Received in revised form 24 March 2026; Accepted 23 April 2026
Available online 24 April 2026

improved attention, Azuma et al. (2018), Du et al. (2020), Fan et al. (2023), and Jacobson et al. (2019) stressed that CO_2 had some negative effects on human cognitive performance, although effect sizes, for example, those reported in the meta-analysis by Fan et al. (2023), were extremely small, and confidence intervals mostly included zero. Fisk et al. (2019), Lowe et al. (2018), Lowther et al. (2021), Mishra et al. (2020), and Wang et al. (2021) reported that the effects were inconsistent and might be dependent on a number of different test-, design-, sample-variables and other confounders, such as the employed cognitive tests and difficulty of the cognitive task, exposure duration, differences in CO_2 levels, and potential moderation, mediation and/or curvilinear effects depending on other environmental or individual health conditions. Almost all individual studies on CO_2 and cognitive performance employ crossover designs (the same subjects are tested at all exposure levels) with small sample sizes (Supplement A for details on most often referenced studies). Some of them find impaired cognitive performance at levels below 1000 CO_2 ppm (Allen et al., 2019; Satish et al., 2012), whereas others do not find problems at levels as high as 15,000 (Rodeheffer et al., 2018) or 20,000 ppm (Maniscalco et al., 2022). Recent crossover studies with more conventional levels of CO_2 also show inconsistent results: Chen et al. (2023) reported no effects of pure CO_2 up to 2100 ppm, whereas Cao et al. (2022) found negative effects of 5000 ppm pure CO_2 compared with lower levels in three out of six cognitive domains. Flagner et al. (2025) reported no effects of exposure to CO_2 at 3000 ppm compared with 900 ppm for various cognitive tasks.

Indoor air quality is also influenced by VOCs. While regulatory limit values exist for specific toxic and cancerogenic VOCs and these VOCs should be eliminated or kept as low as possible, typical indoor air usually contains a mix of different (mostly nonadverse) VOCs. In guidelines, this mixture is referred to as TVOC (total VOCs), but the definition and measurement methods vary greatly, which complicates comparability and leads to inconclusive results with respect to TVOC as a risk index (Andersson et al., 1997). Therefore, some recommendations were changed from staged concepts for TVOCs to concepts in which the exceedance of guideline values is considered on a substance-specific basis (Umweltbundesamt, 2025). As a general indication for precautionary measures, TVOC concentrations above 950 $\mu g/m^3$ indoors are considered unusually high and should be avoided. However, various indoor environments frequently reach this level. For example, average TVOC concentrations in Europe are reported to range from 19 to 1543 $\mu g/m^3$ in schools, 50 to 2610 $\mu g/m^3$ in homes, and 33 to 2449 $\mu g/m^3$ in hospitals (Martins et al., 2025). In subways TVOC ranged between 84.4 and 1414.7 $\mu g/m^3$ (Wang et al., 2024).

There is limited research on the effects of VOCs on cognitive performance. Typically, in studies on cognitive performance, VOCs are only monitored and build up depending on ventilation (Norrefeldt et al., 2021). For example, Wargocki and Wyon (2007) reported significant positive effects on pupils' work rates but not errors, with higher ventilation rates in classrooms. Similar results were obtained by Haverinen-Shaughnessy et al. (2011); however, the authors stressed that the effects on academic performance might not be linear but "taper off" with higher ventilation rates. For primary schools, a systematic review (Honan et al., 2024) reported ventilation-related CO_2 effects on cognitive performance but cautioned that CO_2 was significantly correlated with other indoor air quality parameters, including TVOCs. Wolkoff et al. (2021) reach a similar conclusion for work performance in offices. The effects of (T)VOCs and CO_2 are often confounded, as reduced ventilation in spaces occupied by people increases the concentration of both. Experimental dosing and/or characterization of VOCs has been carried out in only a few studies to date. Exposures to toluene at 380 mg/m^3 (Bælum et al., 1985) or a mixture of 22 common indoor air pollutants at concentrations up to 25 mg/m^3 (Mølhave et al., 1986) led to a decrease in cognitive performance in some, but not all, tests, with the VOC concentration much higher than usual indoors. Allen et al. (2016) and Maddalena et al. (2015) characterized the VOCs that occurred in their experimental studies, but only Allen and colleagues used the data to model the effects separately. They found independent associations of VOCs and CO_2 with performance and better results for lower levels of both. Wargocki et al. (2002) compared two studies that attempted to disentangle the effects of pollution and CO_2 (ventilation) in office spaces on cognitive performance using a present or absent source of pollution (tufted boucle carpet). In both studies, participants performed better in some tests (e.g., text typing) but not in others (e.g., creative thinking) when the pollution source was absent. This source of pollution, however, only covers material emissions, whereas in indoor spaces bioeffluents, VOCs, whose sources are the occupants of the rooms, should also be considered (Wargocki and Wyon, 2017).

With humans occupying indoor spaces at all altitudes, the understanding of the effect of altitude on responses to CO_2 or VOCs is also inadequate. While some studies have examined the effects of high altitudes on cognitive (and physical) performance, they usually focus on hypobaric hypoxic conditions and extreme contexts (e.g., Davranche et al., 2016). There are some indications of cognitive impairment under these conditions; however, these data are not comparable to usual indoor conditions. Existing threshold limit values are only valid up to an altitude of 5000 ft. (1500 m), but there are buildings that are in daily use at even higher altitudes.

While there is a growing body of literature on the effects of controlled (dosed) CO_2 on cognitive performance, studies on the effects of only VOCs on cognition are scarce, and potential modifying effects are not well known. And although people spend time at all altitudes with a corresponding influence on respiration (as the relevant factor for the intake of CO_2 and VOCs), to our knowledge, no study has systematically varied the different combinations of these parameters. Additionally, most studies were conducted with small sample sizes and used crossover designs and heterogeneous outcome measures. Studies with large-scale experimental (RCT) designs targeting different levels of CO_2, VOCs and atmospheric pressure and using valid and reliable cognitive performance measures are lacking. One often-used test in previous studies (e.g., Allen et al., 2016; Rodeheffer et al., 2018; Satish et al., 2012), the strategic management simulation (SMS, Breuer and Streufert, 1995), aims to depict complex decision-making abilities but does not specify tasks, data analyses or the algorithms used. Mishra et al. (2020) and Rodeheffer et al. (2018) criticize that the test has unknown psychometric properties and is a 'proprietary, closed-source instrument' that does not allow independent evaluation of its results. Our research aims to address these issues by using established, transparent and well-defined cognitive performance measures and a large-scale experimental (RCT) approach to test the effects of different combinations of CO_2, VOC and pressure levels beyond the usual co-development of CO_2 and VOC. As an indoor context with well-described characteristics, aircraft cabins were used as the application context of the study, which are also considered examples for other indoor environments. Specifically, our research questions were as follows:

(1) Do realistic, but higher-than-typical concentrations of CO_2 and VOCs compared to typical indoor air conditions have negative effects on the cognitive performance of occupants?
(2) Do CO_2 and VOCs have a combined effect on cognitive performance, or are the effects independent?
(3) Is the atmospheric pressure level a modifier of any response produced by exposure to VOCs and CO_2?

2. Materials and methods

2.1. Study design

A single-blinded, randomized experimental design with stratification for age and sex and eleven different exposure combinations of pressure, CO_2 and VOC levels was used. A previously planned fractional factorial design with two pressure levels, three CO_2 levels and three TVOC levels had to be abandoned as not all targeted VOC levels could be achieved (see exposure section; Supplementary Fig. S1). Consequently, all analyses were performed as a one-factorial design with eleven factor levels.

To have a fixed yet broad target group, participants were predefined to be representative of flight passengers traveling for business (that is, they potentially had to work in aircraft cabins). Age/sex strata from the Airport Travel Survey (Flughafenverband, 2018) were used for stratified randomization of participants (with Winpepi program, Abramson, 2011) to reach a representative sample composition in each experimental condition. The study was carried out in accordance with the Declaration of Helsinki and approved by the Ethics Committee at the Faculty of Medicine, Ludwig-Maximilians-University, Munich (ID: 19-350; 17/06/19). Written informed consent was obtained from all participants.

2.2. Exposure

All eleven experimental sessions were conducted during simulated 4-h flights with approximately two and a half to 3 h under planned exposure conditions that took place at the Fraunhofer Flight Test Facility (FTF) in Holzkirchen (Germany) between December 2019 and January 2020, before the COVID-19 pandemic and its restrictions. The FTF consists of a low-pressure vessel with the inserted front part of a former Airbus A310, which includes a cockpit, galley, and seating for up to 80 passengers (Supplementary Fig. S2). To provide enough space for performing the cognitive tests (and to avoid cheating), the cabin was planned to be occupied with a maximum of 40 passengers seated such that there was always one empty seat between them. This arrangement also avoids mental and physiological stress due to the proximity of other people (Herbig et al., 2023; Lewis et al., 2017).

The environmental conditions that were fixed (or arose because of the cabin occupants) were as follows*: relative humidity*: approximately 10–15% with 40 participants present. *Temperature:* comfortable and typical at 23 °C (Chen et al., 2021). *Level of noise:* kept constant at 74 db (A). *Air velocities* and *outdoor and total air supply rates* for the cabin: With 40 passengers the outdoor air supply rate was 1354 m^3/h, which is 33.9 m^3/h or 9.4 L per second (L/s) per person.

The experimentally varied environmental conditions included atmospheric pressure, CO_2, and VOCs. Two *atmospheric pressure levels* were examined: the ground level pressure was 940 hPa because of the elevation of the FTF location. The cabin pressure at the flight altitude was set to 755 hPa (corresponding to 8000 ft. equivalent altitude) to cover the full range of allowable pressure conditions in aircraft cabins.

Three *levels of CO_2* in the cabin were examined: 1200 ppm (slightly lower than the average reported level in aircrafts (1500 ppm) (Chen et al., 2021), 2750 ppm and 4200 ppm. They were chosen to cover the whole range of current aircraft levels up to the FAA max. or maximum workplace concentration of 5000 ppm (FAA, 2019; OSHA, 2020) with some margin to avoid exceedance of this limit. The levels were reached by injecting CO_2 (with a purity of 99.5%) at predefined rates from a pressurized gas cylinder into the cabin supply air while the outside air supply rate was held constant. With an outdoor CO_2 level at the test site of approximately 400 ppm, CO_2 was dosed at rates of 380 L/h, 2480 L/h and 4440 L/h to reach 1200, 2750 and 4200 ppm CO_2, respectively. Several technical measures (e.g., magnetic shut-off valves) and constant monitoring of CO_2 levels by FTF technicians were used as safety precautions to exclude uncontrolled leakage of CO_2 into the cabin.

To test realistic typical and higher-than-typical *TVOC* concentrations (defined as mass traces of substances quantified via on-line selected-ion flow-tube mass spectrometry (SIFT-MS), i.e., the sum of measured VOCs), three levels were experimentally varied: low, medium and high. On the basis of the extensive review of airborne pollutants in aircraft cabins (Chen et al., 2021), pretesting with different compounds was performed. These were selected a) from among the VOCs frequently measured in aircraft cabins, b) from the VOCs measured at high concentrations and c) from the VOCs with measured concentrations higher than or close to their odor detection limits. Compounds on the IARC list of carcinogens belonging to Groups 1, 2A and 2B were not included even though they are frequently found in aircrafts. The final composition of the dosing mixture is shown in Table 1.

At the low level, no VOCs were dosed, and the resulting level was the background composed of emissions from the cabin interior and the occupants and their belongings. Medium and high levels were determined based on results from a pretest. The mean concentration of the selected VOCs in the mixture was considered the medium VOC level, 266 μg/m^3. Three times this level was considered the high level, 798 μg/m^3. The high level should avoid unbearable odor conditions. These amounts must always be considered in addition to the background. A syringe pump was used to dose the liquid VOC mixture in the respective amounts to the duct supplying the outdoor air. The amount dosed was either 0.44 mL/h or 1.32 mL/h to reach medium and high TVOC levels, respectively. All environmental conditions were monitored continuously with different sensors (Supplementary Fig. S3).

Table 2 shows targeted and measured levels of atmospheric pressure, CO_2 and TVOCs for the experimental conditions (breakdown for VOCs in the conditions in Supplementary Table S1a). Test 1 and test 3 were planned as the typical, low exposure reference cases at the two atmospheric pressure levels, with 1200 ppm CO_2 (typical level in aircrafts) and no dosing of VOCs. Although the measured CO_2 values were somewhat lower than the target values, the targeted relationships of the three levels remained. This was not the case for TVOCs. The reference

Table 1
Composition of the Dosing Mixture with odor detection thresholds and typical amounts in non-smoking commercial flights.

Substance	CAS-No.	ODT[a] in μg/m^3	Typical amount[c] in μg/m^3 M±SD	Concentration in the mixture [μg/m^3] medium	high
Ethanol	64-17-5	990	386 ± 899	82.1	246.3
Acetone	67-64-1	101000	14 ± 5.6	13.9	41.7
1-Propanol	71-23-8	240	71 ± 26	71.4	214.2
2-Propanol	67-63-0	65000	10 ± 3.4	11.5	34.5
Acetic acid	64-19-7	15	11 ± 2.7	10.7	32.1
Toluene	108-88-3	1300	15 ± 12	13.8	41.4
Hexanal	66-25-1	1.1	5.2 ± 4.8	4.2	12.6
Acetonitrile	75-05-8	22000	20 ± 2.5	20.3	60.9
6-MHO	110-93-0	18.9[b]	7.0 ± 3.5	7.0	21.0
Nonanal	124-19-6	2.0	7.8 ± 5.6	9.2	27.6
Limonene	138-86-3	210	24 ± 31	17.1	51.3
Ethyl acetate	141-78-6	3100	6.5 ± 4.4	4.8	14.4
Sum				**266**	**798**

Note. ODT=Odor detection threshold, 6-MHO = 6-Methyl-5-hepten-2-one.
[a] all ODT from Nagata (2003) unless otherwise annotated.
[b] ODT from Benzo et al. (2007); both referenced according to van Gemert (2011).
[c] Amount of VOC typically found in aircraft cabins according to review (Chen et al., 2021; only averaged data).

Table 2
Targeted and measured exposure conditions.

Test no.	Target + Measured: Atm. Pressure [hPa]	Target: CO_2 [ppm]	Measured[a]: CO_2 average [ppm]	Measured[a]: CO_2 max. [ppm]	Target: TVOC	Measured[b]: TVOC average [μg/m^3]	Measured[b]: TVOC max [μg/m^3]
1	940	1200	946	974	Low	650	680
2	755	1200	1163	1181	Medium	864	934
3	755	1200	1190	1226	Low	2103	2630
4	940	2750	2488	2502	Low	442	509
5	755	2750	2656	2767	Medium	761	854
6	755	2750	2733	2753	High	1139	1148
7	940	2750	2482	2562	Medium	1181	1214
8	940	4200	3748	3761	Medium	648	761
9	755	4200	4138	4182	High	1382	1462
10	755	4200	4036	4235	Low	1556	2147
11	940	4200	3658	3862	High	1719	1786

Note.
[a] real-time measurements with Los Gatos Research portable analyser based on Off-Axis Integrated Cavity Output Spectroscopy (OA-ICOS); pressure corrected to 940 hPa.
[b] measured with selected-ion flow-tube mass spectrometry (SIFT-MS) pressure corrected to 940 hPa.

test 3 and test 10 showed unexpectedly high levels on no-dosing days and, in fact, the highest levels among all conditions. As the cabin background was almost the same for all tests (Supplementary Table S1b) and ventilation with outdoor air was constant, participants had to be the source. The order of TVOC amounts also did not match the order of target levels for tests 6, 7, and 8. Because of these unintended exposure conditions, we had to change the statistical analyses.

2.3. Sample: size, criteria, recruitment, and screening

An a priori sample size estimation with G*Power 3.1 (Faul et al., 2009) for the planned fractional factorial design showed an overall required sample size of N = 391 participants. A second estimation for the necessary switch to a one-factorial design with eleven factor levels showed that the primary estimated sample size is also sufficient to detect small-to medium-sized effects, with a conventional alpha error of 5% and a power of 0.85 (1–β), which is still above the conventional level of 0.80 usually adopted in this type of study.

Participants were recruited via an agency for background actors to reach this sample size in the respective age and sex strata for business passengers (based on Flughafenverband, 2018). The inclusion criterion was healthy adults willing and able to give informed consent. The exclusion criteria were people at health risk in low-pressure conditions (e.g., heart conditions), people who might experience problems (e.g., people who cannot sit without pain for a few hours), unrepresentative people (people without flight experience), and people who might cause problems during experiments (e.g., people with claustrophobia or who are aggressive).

A three-step screening procedure was used to ensure that these criteria were met. First, potential participants received a prescreening self-check together with the study information. In the second onsite screening part, participants completed a short set of questions providing information on severe surgery, cardiac infarction, or pneumothorax in the past six months, as well as on ongoing cold, sinusitis, etc. In the third step, participants were individually checked by onsite physicians, who excluded unfit participants as well as participants who were aggressive, intoxicated, or showed other behavioral problems. On the days of the trial, 33 participants did not show up, although participation was confirmed. An additional 17 participants called in sick, the onsite physicians screened out three, and one person aborted during the trial and left the FTF via pressure lock (Supplementary Fig. S4). Overall, 398 subjects participated, and the necessary sample size was reached.

2.4. Performance measures

Cognitive performance was measured with a variety of different tests. The results were supposed to be valid for a heterogeneous group of persons from different educational and cultural backgrounds, therefore culture-/education-fair tests of basal cognitive abilities were the core of the battery. Wherever possible, psychometrically tested and established tests with parallel forms for multiple testing and standardized for the German population were used for repeated testing at the beginning and end of exposure; these tests included information processing speed, concentration performance, working memory, and logical reasoning:

Information processing speed was measured by a German variant of the Trail-Making-Test (Oswald and Roth, 2016). The respondents needed to connect numbers in sequence from 1 to 90 (30 s per matrix, 4 matrices). It takes only a few minutes and allows a highly reliable estimation of mental speed.

The d2-test (Brickenkamp et al., 2010) was used to measure *concentration/attention*. The test measures selective and sustained attention. It contains letters p and d; 14 rows with 57 letters each; 1 to 4 dashes above or below the letter. For each row, respondents have 20 s to cross out all – and only those - d with 2 dashes. The concentration performance score (correctly marked d2 minus type 2 error) is reported.

Working memory is a necessary precondition for every other cognitive ability. It was assessed with a module from the Intelligence Structure Test (Liepmann et al., 2007). We used figural memory to measure cultural fair (i.e., nonverbal) memory performance. The participants had 1 min to memorize the given figure pairs. After a period of time, they are then asked to mark the correct letter of the figure pairs previously memorized.

Logical reasoning includes the ability to discern patterns within information, to comprehend the basic structure of a set of relations and to recognize logically equivalent statements. We measured logical reasoning with the (figural) Matrices module from Liepmann et al. (2007). Respondents need to choose which of five choice figures can be placed by a question mark to logically complete the given sequence, that is, to use logical reasoning to find the underlying pattern of the presented stimuli.

The following tests were only performed once at the end of the second test battery:

Executive control was assessed with a group version of the Stroop test (Golden, 1975; Stroop, 1935). The respondents got a list of color words printed in different colors (either congruent or incongruent with the word sense). The task is to say/mark the color of a word but not the name of the word. Differences in time between reactions to congruent and incongruent stimuli reflect executive control; longer reaction times to incongruent stimuli indicate less executive control since the person needs more time to suppress the dominant reaction.

Vigilance/alertness was measured as a stimulus to be noticed while participants were occupied with other things (Simons and Chabris, 1999), which is different from previous research with vigilance tasks that are clearly identifiable as tests (e.g., Pang et al., 2021). We used an auditory stimulus presented as an emergency vehicle siren (7 s with a volume of 27 dB) in the second half of the exposure. At the end of the second test battery, participants were asked if they noticed any unusual noise(s) during the flight. To lower the guessing probability, the participants were given 10 different items, including “no” and the instruction that one or more answers might be correct, allowing a threefold scoring (“no detection/wrong”, “precise detection/only siren”, “imprecise detection/siren plus other(s)”).

Problem-solving was assessed with the NASA-Game (Pfeiffer and Jones, 1970), which was originally designed as a game in group vs. individual scenarios. The story behind the game is a failed moon landing with the goal of reaching a faraway rescue craft. The participants are given a list of 15 items to use for their survival and are asked to prioritize these items in a hierarchy of usefulness. A model solution is available so that the results can be assessed as deviation from the optimum. That is, the greater the deviation is, the worse the solution of the problem is.

As a task already used in research on the effects of indoor air quality on cognitive performance, we used *acoustic proofreading* (Wargocki and Wyon, 2006). The subjects listened to an acoustic text and simultaneously read the same text. The written text included errors that the subjects needed to mark. We used a rather complex press text on a stamp swap meet. The text was recorded and played to the participants while they had to correct the text in front of them. There were three types of errors: additional words, missing words, and exchanged words. Correctly marked errors were added, and incorrectly marked errors were subtracted from this.

2.5. Control variables: self-report measures

The control variables were measured before exposure during the setting of the respective pressure level. They consisted of single-item sociodemographic aspects. Health, last night's sleep, circadian rhythm and chronotype were included as variables that can directly influence cognitive performance independently of environmental conditions (Snow et al., 2019). With an overall experimental duration of more than 4 h, it is impossible to avoid every of the known “dips” in performance during the day that are based on the circadian rhythm (Valdez et al., 2012). We therefore used a proxy control variable ‘chronotype’ (Roenneberg, 2015) and the state of rest. For further control variables see Table 3.

At the end of the exposure, the participants were also asked if they were familiar with the cognitive performance tests, as the mere

Table 3
Sample characteristics and randomization check; targeted atmospheric pressures and CO_2 are shown and actual (not targeted) TVOC levels.

	940 hPa; 1200 CO_2; 650 VOC	755 hPa; 1200 CO2; 864 VOC	755 hPa; 1200 CO_2; 2103 VOC	940 hPa; 2750 CO_2, 442 VOC	755 hPa; 2750 CO_2; 761 VOC	755 hPa; 2750 CO_2; 1139 VOC	940 hPa; 2750 CO_2, 1181 VOC	940 hPa; 4200 CO_2; 648 VOC	755 hPa; 4200 CO_2; 1382 VOC	755 hPa; 4200 CO_2; 1556 VOC	940 hPa; 4200 CO_2; 1719 VOC	Overall difference: p-value
Male	24	17	27	26	25	27	20	20	23	19	20	0.818
Female	14	13	13	12	13	10	15	16	13	14	17	
Age	42.3 ± 12.6	41.1 ± 12.7	42.9 ± 11.8	43.0 ± 11.8	42.3 ± 13.2	41.0 ± 10.3	44 ± 13.7	42.3 ± 12.7	42.8 ± 12.8	44.3 ± 13.0	44.0 ± 13.5	0.989
Body Mass Index	25.3 ± 4.0	25.7 ± 9.1	25.7 ± 5.9	25.2 ± 4.4	24.2 ± 3.4	25.4 ± 4.3	25.1 ± 6.5	25.8 ± 6.8	24.6 ± 5.1	23.8 ± 3.8	26.5 ± 6.9	0.768
Smoking:												0.144
Never smoked	10	14	15	19	18	18	16	20	13	17	15	
Do not smoke anymore	12	5	12	12	6	5	10	7	13	7	11	
Smoke occasionally	5	4	7	4	10	8	1	4	3	1	7	
Smoke regularly	11	7	6	3	4	6	8	5	7	8	4	
Educational level:												0.919
below university-entrance diploma	15	12	14	21	17	15	16	17	14	15	15	
university-entrance diploma	23	18	26	17	21	22	19	19	22	18	22	
General health assessment[a] [8-41]	17.0 ± 4.5	15.6 ± 5.1	17.0 ± 5.4	15.5 ± 5.3	15.1 ± 4.5	15.7 ± 5.2	14.8 ± 4.6	17.0 ± 5.7	15.1 ± 3.9	15.6 ± 5.3	14.1 ± 4.5	0.173
Multiple Chemical Sensitivity[b] [0-32]	5.4 ± 5.6	5.3 ± 4.2	6.5 ± 7.2	5.5 ± 5.5	6.5 ± 5.9	4.6 ± 5.4	7.0 ± 8.2	6.7 ± 6.3	7.2 ± 7.2	5.8 ± 5.8	5.7 ± 5.3	0.780
Average number of flights p.a.	2.8 ± 3.7	3.7 ± 4.8	5.7 ± 11.2	4.8 ± 12.0	13.7 ± 30.0	2.5 ± 2.3	3.2 ± 4.0	4.1 ± 6.3	15.0 ± 43.5	4.2 ± 8.9	2.1 ± 2.0	0.009**
Sleep quality before test[c] [1-4]	3.0 ± 0.6	2.7 ± 0.6	2.8 ± 0.7	3.0 ± 0.5	3.0 ± 0.6	2.9 ± 0.6	2.9 ± 0.5	3.0 ± 0.7	2.7 ± 0.6	3.0 ± 0.6	3.0 ± 0.6	0.071
Health on day of flight[d] [1-5]	4.2 ± 0.6	3.9 ± 0.7	4.0 ± 0.6	4.1 ± 0.8	4.3 ± 0.7	4.2 ± 0.6	4.2 ± 0.6	4.3 ± 0.7	4.0 ± 0.7	4.0 ± 0.6	4.2 ± 0.6	0.263

Note: CO_2 in ppm; VOC in μg/m^3; N/M ± SD; Last column: significance level of χ^2 (categorical variables) or ANOVA (continuous variables).
[a] Short Form-8 Health Survey (SF-8, Ware et al., 2001): Lower values equal better health.
[b] Multiple Chemical Sensitivity scale (Kiesswetter et al., 1997): Higher values equal higher multiple chemical sensitivity.
[c] Single item: Higher values equal better sleep quality.
[d] Single item: Higher values equal better simple/short—rated health at day of test.

knowledge of the procedure of a test can increase performance (Reeve et al., 2009). Perceived air quality was recorded via different questions/scales: How would you rate the air quality in this flight? (very poor to very good); In comparison with other flights you experienced, how would you rate the air quality of this flight? (much worse to much better); How would you assess the odor intensity in this flight? (no smell to overwhelming smell). Moreover, a bipolar interrupted visual analog scale for air quality acceptability was used (Wargocki, 2001).

2.6. Procedure

All eleven experimental sessions took place on individual workdays. Upon arrival at the flight test facility (approximately 10:00 a.m. on each test day), the participants were welcomed and received printed informed consent forms and the second screening questionnaire. In the first hour before the actual experiment (T-60 to T0), they also received a small snack and noncarbonated nonalcoholic drinks and were familiarized with the setting, including an information- and Q&A session (for a final informed consent decision), as well as safety instructions. Moreover, the last onsite screening with a "fit-to-fly" decision by a physician was carried out before boarding.

The participants were not seated next to or behind each other so that they did not disturb each other during the tests; there was always an unoccupied seat or an aisle between them. Unoccupied seats were used for measurement equipment or not at all.

At approximately 11:00 a.m. (T0), the simulation started. The following test sequence was used:

- T0: (Boarding): subjects entering the FTF and taking their seats
- T+25 min: establishing the respective pressure level. Pretrial performance tests allowing participants to familiarize themselves with the tests so that learning curve steepness flattened. Questionnaires on control variables.
- T+25 to T+40: baseline reference (same for every condition) of approximately 1200 ppm CO_2/100 μg/m^3 TVOC to obtain physiological measures at rest (not reported here).
- **T+40:** (onset of actual exposure condition): start of dosing of CO_2 or VOC or both
- **T+55:** (start of exposure): actual planned exposure level reached
- **T+55 to T+85:** first test battery with measurement of outcomes, including cognitive performance tests, questionnaires on health symptoms, sleepiness and emotional state.
- **T+125:** questionnaires on health symptoms and comfort/satisfaction with environmental conditions (not reported here)
- **T+175 to T+205:** second test battery with second measurement for repeated cognitive performance and one-time cognitive tests
- T+205 to T+230: establishing ground level pressure. Short questionnaire to explore the perceived air quality, realism, etc., of the simulation.
- T+230 (Deboarding): subjects leaving the FTF. Debriefing.

The experimental conditions were maintained for approximately 165 min (marked in bold above). To avoid effects of boredom between measurements, participants were provided with emotionally neutral entertainment (nature journal). This is important, as the mood-inducing effect of film, music or other material might interfere with performance-related or well-being outcomes (Gendolla and Krüsken, 2002). Uncontrolled, potentially mood-inducing activities, such as smartphone use, newspapers or own music, were prohibited.

2.7. Statistical analyses

All performance data were converted to percentage-of-maximum-possible values (Cohen et al., 1999) to provide a common metric and to allow easy evaluation of the level of performance across tests. One-factorial analyses of (co)variance (with repeated measures) were carried out via IBM SPSS (Version 29.0), with the exposure condition as a factor with eleven levels and post hoc pairwise comparisons. The statistical significance level was set to $p < .05$ (2-tailed). All pairwise comparisons were corrected for multiple testing using the false discovery rate (FDR) method to control for the proportion of type I errors while still minimizing the risk of overlooking relevant effects (Benjamini and Hochberg, 1995).

The criteria for the exclusion of participants for the respective analyses were predefined according to the observation protocols from the experimenter, when people were clearly cheating or incapable of understanding/performing the test. Overall, this was true for N = 10 participants in the working memory test, N = 9 in the attention/concentration test, N = 9 in the information processing speed tests, N = 6 in the proofreading test, and N = 2 in the vigilance test. With N = 90 participants who were found in the evaluation to have probably not understood the instruction correctly, the results on executive control (Stroop test) are probably the least reliable. A distribution analysis of all excluded and missing values showed no systematic differences between exposure groups.

As large-scale experimental designs might include high heterogeneity of subsamples, all analyses of cognitive performance were controlled for a parsimonious set of variables, namely, age, chronotype, and familiarity with the test, as these might influence the outcome. For problem-solving and proofreading, which are the only tests that depend on knowledge and education, the analyses were additionally adjusted for the level of education of the participants. Crude models are reported in Supplementary Tables S2 and S3.

3. Results

3.1. Sample characteristics and randomization check

In total, 248 men (62.3%) and 150 women (37.7%) (N = 398) participated, with an overall mean age of 42.7 years (SD = 12.4; range: 18-69 years), which is in line with the targeted population of business passengers. Overall, 57.1% achieved a general qualification for the university entrance/A level, and 0.5% left school without graduation. The subjects flew 5.7 times a year on average (SD = 17.6; range: 0-240). The high maximum number of flights was caused by the participation of a few cabin crew members and (ex-)pilots, especially in two of the exposure groups. To ensure that the samples in the different experimental conditions were comparable for all types of potentially relevant sociodemographic and control variables, statistical comparisons were conducted (Table 3). Familiarity with at least one of the cognitive tests was prevalent in more than half of the participants in each group, without significant differences between groups. Except for the average number of flights, randomization was successful; that is, the participants' characteristics were comparable across all eleven exposure groups.

3.2. Perception of air quality at the end of the exposure

The air quality control questions showed no differences between exposure groups in either the average ratings (Kruskal-Wallis test; all $p > .14$) or the distribution of judgments (Chi^2 test; all $p > .30$). Descriptively, the air quality assessments were mostly positive (with at least 57% voting that the air quality was very good or good): the overall mean was 3.83 ± 0.79 (median 4). The worst rating was in the group with the (unexpected) highest VOCs (755 hPa/1200 CO_2/2103 μg/m^3 VOC), with a mean of 3.60 ± 0.84 (median also 4) (Fig. 1).

Air quality was rated as comparable or better than in other flights (overall: 3.56 ± 0.80, median 4; worst rating in the group 755 hPa/4200 CO_2/1382 μg/m^3 VOC with the most flight experience with 3.31 ± 0.86, median 3). Smell was hardly perceived (overall: 1.61 ± 0.75, median 1; worst assessment in group 755 hPa/1200 CO_2/864 μg/m^3 VOC with 1.80 ± 0.93, median 2), and acceptability was high (overall: 82.34 ±

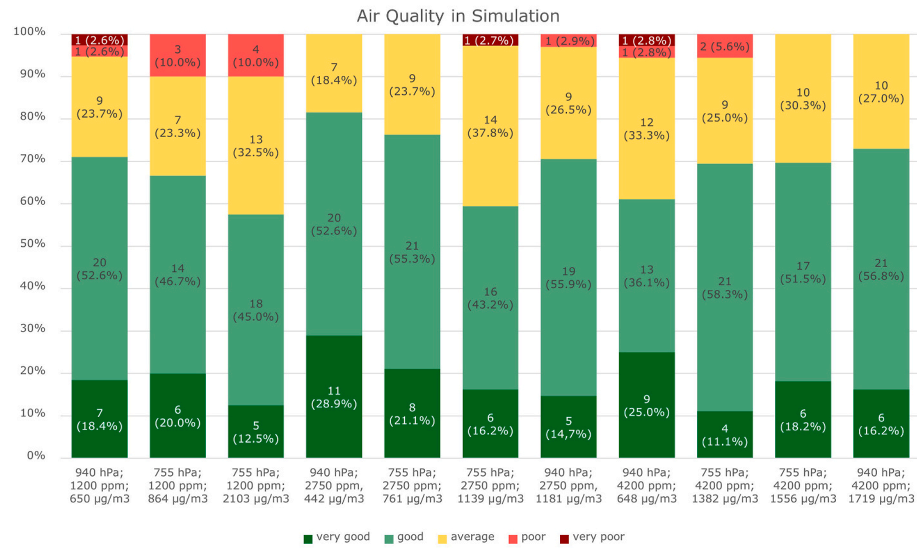

Fig. 1. Frequency of responses to air quality assessment at different exposure conditions.
Note. X-axis: pressure in hPa, CO_2 in ppm, TVOC in µg/m.[3].

Table 4
Repeated cognitive performance in exposure groups.

#	Exposure group	Concentration Performance EMM (95% CI) T1	Concentration Performance EMM (95% CI) T2	Information processing speed EMM (95% CI) T1	Information processing speed EMM (95% CI) T2
1	940 hPa/1200 CO_2/650 VOC	**58.47 (54.01;62.93)**[2,4]	**63.32 (58.67;67,97)**[2,4]	52.23 (48.68; 55.79)	56.19 (52.84; 59.55)
2	755 hPa/1200 CO_2/864 VOC	**69.63 (64.52;74.74)**[1,3,7,10]	**75.25 (69.92;80.58)**[1,3,7,10]	53.24 (49.09; 57.39)	57.64 (53.71; 61.56)
3	755 hPa/1200 CO_2/2103 VOC	**59.76 (55.37;64.15)**[2,]	**64.64 (60.06;69.22)**[2,4]	50.97 (47.46; 54.47)	54.44 (51.13; 57.75)
4	940 hPa/2750 CO_2/442 VOC	**68.16 (63.37;72.94)**[1,10]	**74.65 (69.67;79.64)**[1,3,7,10]	51.38 (47.73; 55.03)	57.31 (53.86; 60.76)
5	755 hPa/2750 CO_2/761 VOC	64.87 (60.42; 69.31)	68.34 (63.70; 72.98)	52.79 (49.24; 56.33)	54.74 (51.39; 58.09)
6	755 hPa/2750 CO_2/1139 VOC	62.47 (57.95; 66.98)	68.47 (63.76; 73.18)	53.52 (49.92; 57.12)	57.06 (53.66; 60.47)
7	940 hPa/2750 CO_2/1181 VOC	**59.32 (54.48;64.17)**[2]	**64.13 (59.07;69.19)**[2,4]	49.24 (45.43; 53.05)	54.44 (50.84; 58.05)
8	940 hPa/4200 CO_2/648 VOC	65.16 (60.58; 69.75)	70.92 (66.14; 75.70)	51.55 (47.89; 55.21)	54.57 (51.11; 58.03)
9	755 hPa/4200 CO_2/1382 VOC	61.28 (56.69; 65.87)	66.58 (61.79; 71.37)	50.60 (46.84; 54.36)	55.68 (52.13; 59.24)
10	755 hPa/4200 CO_2/1556 VOC	**57.91 (53.14;62.68)**[2,4]	**63.25 (58.27;68.22)**[2,4]	51.64 (47.77; 55.52)	56.56 (52.90; 60.22)
11	940 hPa/4200 CO_2/1719 VOC	60.09 (55.58; 64.60)	65.93 (61.23; 70.64)	47.36 (43.76; 50.96)	54.53 (51.13; 57.93)
	Repeated measures	**F(1,374)=22.67, p<.001**		**F(1,374)=9.20, p=.003**	
	Exposure	**F(10,374)=2.81, p=.002**		F(10,374) = 0.58, p = .835	
	Interaction	F(10,374) = 0.88, p = .549		**F(10,374)=4.39, p<.001**	
	Group differences (overall model)*	Group 2: vs. 1 (p = .016)/vs. 3 (p = .039)/vs. 7 (p = .039)/vs. 10 (p = .016) Group 4: vs. 1 (p = .024)/vs. 10 (p = .025)		none	

#	Exposure group	Figural working memory EMM (95% CI) T1	Figural working memory EMM (95% CI) T2	Logical reasoning EMM (95% CI) T1	Logical reasoning EMM (95% CI) T2
1	940 hPa/1200 CO_2/650 VOC	53.64 (46.94; 60.34)	77.52 (70.40; 84.64)	50.97 (46.15; 55.78)	53.13 (47.86; 58.41)
2	755 hPa/1200 CO_2/864 VOC	59.67 (51.84; 67.51)	81.61 (73.28; 89.93)	48.14 (42.72; 53.57)	54.59 (48.65; 60.54)
3	755 hPa/1200 CO_2/2103 VOC	**46.80 (40.19;53.41)**[6]	72.35 (65.33; 79.38)	47.72 (43.04; 52.40)	55.39 (50.26; 60.52)
4	940 hPa/2750 CO_2/442 VOC	**53.22 (46.34;60.10)**[6]	76.86 (65.33; 79.38)	47.48 (42.68; 52.27)	56.02 (50.77; 61.28)
5	755 hPa/2750 CO_2/761 VOC	54.51 (47.82; 61.20)	74.60 (67.49; 81.71)	50.68 (45.88; 55.48)	55.77 (50.51; 61.03)
6	755 hPa/2750 CO_2/1139 VOC	**67.62 (60.63; 74.62)**[3,4,7,8,9,10,11]	81.57 (74.13; 89.01)	48.84 (43.98; 53.71)	48.93 (43.59; 54.26)
7	940 hPa/2750 CO_2/1181 VOC	**49.88 (42.70;57.05)**[6]	73.34 (65.71; 80.97)	44.41 (39.41; 49.42)	49.12 (43.64; 54.61)
8	940 hPa/4200 CO_2/648 VOC	**51.24 (44.22;58.26)**[6]	78.66 (71.20; 86.12)	49.27 (44.33; 54.21)	50.98 (45.57; 56.40)
9	755 hPa/4200 CO_2/1382 VOC	**48.09 (41.19;54.99)**[6]	66.54 (59.21; 73.87)	48.86 (43.92; 53.81)	53.55 (48.13; 58.97)
10	755 hPa/4200 CO_2/1556 VOC	**44.44 (37.27;51.62)**[6]	72.84 (65.21; 80.47)	48.81 (43.65; 53.96)	53.74 (48.10; 59.39)
11	940 hPa/4200 CO_2/1719 VOC	**49.41 (42.61;56.20)**[6]	73.25 (66.03; 80.47)	45.82 (40.95; 50.69)	55.90 (50.57; 61.24)
	Repeated measures	**F(1,374)=10.51, p=.001**		F(1,384) = 1.78, p = .183	
	Exposure	**F(10,374)=2.39, p=.009**		F(10,384) = 0.53, p = .871	
	Interaction	F(10,374) = 1.61, p = .101		**F(10,384)=1.93, p=.040**	
	Group differences (overall model)*	Group 6: vs. 3 (p = .016)/vs. 9 (p = .004)/vs. 10 (p = .014)/vs. 11 (p = .047)		none	

Note. CO_2 in ppm; VOC in µg/m[3]; t1/t2 = measurement times at beginning and end of exposure; EMM: estimated marginal means adjusted for age, chronotype, familiarity with test; superscripted numbers/letters in columns denote significant differences between indicated groups at respective measurement time (FDR-adjusted post hoc group differences at 5% FDR), significant differences additionally highlighted in bold, *significant overall differences with Benjamini-Hochberg p-value.

16.49, median 88; worst assessment in group 755 hPa/2750 CO_2/1139 μg/m^3 VOC with 74.98 ± 19.58, median 79).

3.3. Effects of different exposures on cognitive performance with repeated measures

All four repeated cognitive performance tests provide population-based, age-adjusted percentile rank norms. Compared with the general population, performance at the beginning of exposure across all groups was average to good (concentration: median rank norm 66, that is, better than 65% of the norm population; information processing speed: 84; figural working memory: 62; and logical reasoning: 69). At the end of the exposure, all groups presented higher percentile rank norms (concentration: 82, information processing speed: 92, figural working memory: 95, logical reasoning: 76). Supplementary Fig. S5 illustrates the changes in percentile rank norms across the different exposure groups.

The mean performance scores and analyses of variance results are presented in Table 4. Performance of all groups increased at the end of the exposure. Except for logical reasoning, this was also shown by significant repeated measures F-tests (all $p \leq .003$). In addition, for concentration performance and figural working memory, the exposure had significant effects (concentration: $F(10,374) = 2.81$, $p = .002$; working memory: $F(10,374) = 2.39$, $p = .009$), for the other two repeatedly measured cognitive abilities, significant interaction effects between time and exposure were found (information processing speed: $F(10,374) = 4.39$, $p < .001$; logical reasoning: $F(10,384) = 1.93$, $p = .040$).

FDR-adjusted group comparisons across (overall model) and at the two measurement times showed that the *concentration performance* differences were dominated by two high-performing groups (755 hPa/1200 CO_2/864 μg/m^3 VOC and 940 hPa/2750 CO_2/442 μg/m^3 VOC) that differed significantly from one group with medium-level environmental conditions (940 hPa/2750 CO_2/1181 μg/m^3 VOC), one with high levels of CO_2 and VOCs (755 hPa/4200 CO_2/1556 μg/m^3 VOC) and two groups with low resp. lower levels (940 hPa/1200 CO_2/650 μg/m^3 VOC and 755 hPa/1200 CO_2/2103 μg/m^3 VOC). All groups increased their performance to comparable degrees (all simple slopes $p < .001$).

The performance differences in the *figural working memory* test are dominated by the group 755 hPa/2750 CO_2/1139 μg/m^3 VOC, which performed better than almost every other group at the beginning of the exposure, whereas it leveled out at the second measurement. All groups also significantly increased their performance (all simple slopes $p < .001$), with the highest increase in the 755 hPa/4200 CO_2/1556 μg/m^3 VOC (28.4 performance score gain) and 940 hPa/4200 CO_2/648 μg/m^3 VOC (27.4 gain) groups and the lowest increase (13.9 gain) in the highest starting performance group.

Significant interaction effects of exposure and measurement time for information processing speed and logical reasoning hint at differences in slopes. For *information processing speed*, however, all groups increased their information processing speed significantly at $p < .001$, except for the 755 hPa/2750 CO_2/761 μg/m^3 VOC group, whose information processing speed increased the least, with approximately two performance points ($p = .049$). The steepest increase, with approximately seven points, occurred at 940 hPa/4200 CO_2/1719 μg/m^3 VOC.

For *logical reasoning*, only three exposure groups significantly increased their performance, whereas all others remained at approximately the same level: Group 755 hPa/1200 CO_2/2103 μg/m^3 VOC (7.7 points increase, $p = .005$), Group 940 hPa/2750 CO_2/442 μg/m^3 VOC (8.5 points, $p = .001$), and Group 940 hPa/4200 CO_2/1719 μg/m^3 VOC, with the highest increase of approximately 10 points from the beginning to the end of exposure ($p < .001$).

To capture potential subtle changes adding up across different cognitive domains, an additional "level of cognitive performance" variable was computed (Supplementary Fig. S6). The general linear model results are similar to the results in the individual cognitive domains: All groups increased their performance level from beginning to end of exposure ($F(1,363) = 21.51$, $p < .001$; all simple slopes $p < .001$) and did so differentially ($F(10,363) = 2.98$, $p = .001$), but there were no overall differences between exposure groups ($F(10,363) = 1.39$, $p = .182$). Descriptively, the lowest sum value increase of approximately 24 performance scores was found for the 755 hPa/2750 CO_2/1139 μg/m^3 VOC group, which started with the highest overall performance in the beginning, whereas the steepest slopes were found for groups 940 hPa/4200 CO_2/1719 μg/m^3 VOC (~47 points increase), 940 hPa/2750 CO_2/442 μg/m^3 VOC (~46 points), and 755 hPa/4200 CO_2/1556 μg/m^3 VOC (~45 points). As with individual cognitive domains, there was no indication that these changes were correlated with the exposure conditions in any systematic way.

3.4. Effects of different exposures on one-time cognitive performance measures

Executive control, problem-solving, acoustic proofreading and vigilance were tested once at the end of the exposure time. There were no

Table 5
One-time cognitive performance in exposure groups.

		Executive Control	Problem-solving	Acoustic Proofreading	Level of cognitive performance in one-time measures
	Exposure group	EMM (95% CI)	EMM (95% CI)	EMM (95% CI)	EMM (95% CI)
1	940 hPa/1200 CO_2/650 VOC	61.24 (55.25; 67.23)	53.37 (48.73; 58.00)	63.53 (58.10; 68.96)	53.85 (49.34; 58.37)
2	755 hPa/1200 CO_2/864 VOC	66.93 (60.92; 72.94)	53.33 (48.10; 58.55)	69.64 (63.58; 75.69)	51.10 (46.05; 56.14)
3	755 hPa/1200 CO_2/2103 VOC	64.23 (58.70; 69.75)	58.15 (53.64; 62.65)	69.27 (64.04; 74.50)	53.91 (49.63; 58.18)
4	940 hPa/2750 CO_2/442 VOC	63.00 (57.58; 68.43)	53.87 (49.24; 58.49)	72.41 (67.04; 77.77)	50.99 (46.61; 55.37)
5	755 hPa/2750 CO_2/761 VOC	59.98 (54.90; 65.07)	55.63 (51.01; 60.24)	69.27 (63.92; 74.62)	49.66 (45.29; 54.04)
6	755 hPa/2750 CO_2/1139 VOC	61.55 (56.37; 66.73)	55.19 (50.50; 59.87)	68.12 (62.54; 73.70)	53.43 (48.99; 57.86)
7	940 hPa/2750 CO_2/1181 VOC	57.22 (51.88; 62.56)	55.99 (51.17; 60.81)	62.35 (56.77; 67.92)	49.43 (44.87; 53.98)
8	940 hPa/4200 CO_2/648 VOC	61.07 (55.54; 66.60)	55.10 (50.34; 59.86)	68.87 (63.37; 74.37)	55.15 (50.65; 59.65)
9	755 hPa/4200 CO_2/1382 VOC	59.69 (53.43; 65.95)	47.05 (42.29; 51.81)	63.37 (57.69; 69.06)	47.40 (42.83; 51.97)
10	755 hPa/4200 CO_2/1556 VOC	61.67 (56.14; 67.19)	53.30 (48.35; 58.26)	69.92 (64.08; 75.75)	53.26 (48.58; 57.95)
11	940 hPa/4200 CO_2/1719 VOC	61.30 (55.96; 66.64)	53.91 (49.23; 58.59)	70.76 (65.33; 76.18)	53.12 (48.69; 57.55)
	Exposure	$F(10,307) = 0.76$, $p = .672$	$F(10,383) = 1.33$, $p = .214$	$F(10,378) = 1.40$, $p = .180$	$F(10,379) = 1.08$, $p = .376$
	Group differences[a]	none	none	none	none

Note. CO_2 in ppm; VOC in μg/m^3; EMM: estimated marginal means, all tests adjusted for age and chronotype, problem-solving and proofreading additionally for educational level, executive control and problem-solving additionally adjusted for familiarity with test.
[a] FDR-adjusted pairwise comparisons.

effects of exposure for either the four tests (F between 0.76 and 1.40; Kruskal-Wallis-Test for vigilance 12.33, p = .263) or the overall performance in these cognitive domains (Table 5, Supplementary Fig. S7, Supplementary Table S3). There were also no differences between single exposure groups.

Fig. 2 shows the detection rates in the vigilance task without adjustments. This task seemed to have had a high degree of difficulty, as only a few people noticed and specifically remembered the stimulus (21.4% of the complete sample). A further 10.3% stated that they remembered not only the stimulus but also other auditory stimuli (imprecise detection). Neither overall exposure nor pairwise comparisons between exposure groups were significant.

4. Discussion

In this large-scale experimental study, none of the eight cognitive domains tested showed a consistent pattern of groups affected by the respective air quality exposure. While there were a few significant differences, no consistent performance de-/increase with de-/increase in CO_2 or VOC levels or between pressure levels found. All groups showed average to good performance. Increases in performance over time in the repeated measures hint at unimpaired learning effects (e.g., Twardella et al., 2012). Overall, the results allow with fair certainty to conclude that there is no evidence that cognitive performance under typical indoor air conditions differs from performance under conditions with higher VOCs and CO_2 concentrations in the tested ranges or that it is dependent on the atmospheric pressure level.

Our study used an experimental (RCT) approach as the gold standard and the only type of design allowing to infer causation (Grimes and Schulz, 2002a, 2002b). This overcomes the shortcomings of recent research on the effects of indoor air quality on cognition, which mostly uses small crossover designs. Since these designs do not include random allocation, they introduce the risk of systematic bias, that is, effects might be related to the participants' unknown underlying features. In small samples, increased homogeneity also leads to a greater risk of false positive results than in large-scale experimental approaches. The impact of random error and greater variability in randomized approaches can be minimized with sufficiently large sample sizes, as was done in our study. In addition, the range of people examined in our study is considerably wider (Table 3) than in previous studies, which often looked at young (male) students/participants (e.g., Satish et al., 2012; Supplement A). Although our sample is far from being generally representative and does not include vulnerable groups, we add evidence across a much wider range of sociodemographic groups.

In terms of exposure, the intended pressure and CO_2 levels were reached. Both parameters seemingly have no negative effects on the performance of a range of cognitive functions. For CO_2, this result is in line with several especially rigorously designed studies, as singled out by the meta-analysis of Du et al. (2020), which tend to show no or only marginally significant small effects, except for the performance in the SMS test battery, where in some - but not all (Rodeheffer et al., 2018) - studies effects were observed at CO_2 levels as low as 900-1000 ppm (Allen et al., 2019; Satish et al., 2012).

The situation is different for VOCs as exposure parameters. The dosed levels were always in addition to those in the cabin. These could contain various degrees and types of VOCs, as is also the case in real life, especially when major sources of pollution are humans. While the measured VOCs in the different conditions were in the typical indoor (AGÖF, 2013) and aircraft cabin ranges (Chen et al., 2021; Guan et al., 2015), there were differences between the target and actual conditions. Notably, the group most affected by a deviation from the planned design at 755 hPa, low CO_2 and targeted low VOC actually had the highest TVOC level of all groups, with ethanol as the highest single VOC at 1514 μg/m^3 (Supplementary Table S1a). As VOCs were not dosed in this group, participants had to be the source and one may speculate about alcohol consumption. However, neither the study team nor the physician coming close to the subjects when performing the medical screening perceived alcohol smells, and we cannot determine the reason for this anomaly. Nevertheless, cognitive performance results showed no conspicuities either in relation to other groups or to population norms. Another issue regarding the VOC exposure parameter should be considered: Although we dosed and measured pollutants typically monitored in buildings and in aircraft, they constitute only a subset of all potential mixtures of pollutants at different concentrations that can be encountered indoors. The effects of other mixtures and levels of VOCs cannot be ruled out on the basis of the present results. For example, while we did not observe negative effects on performance for TVOC levels up to 2100 μg/m^3, Allen et al. (2016) reported reduced performance at three times lower levels of 600 μg/m^3. However, the main source of pollution in their study was not human bioeffluents but rather emissions from building materials. Similarly, in a study where the main source was a used carpet (Wargocki et al., 1999), TVOC levels of approximately 370 μg/m^3 significantly reduced typing performance.

The VOC issue also points to the major limitation of our study that we

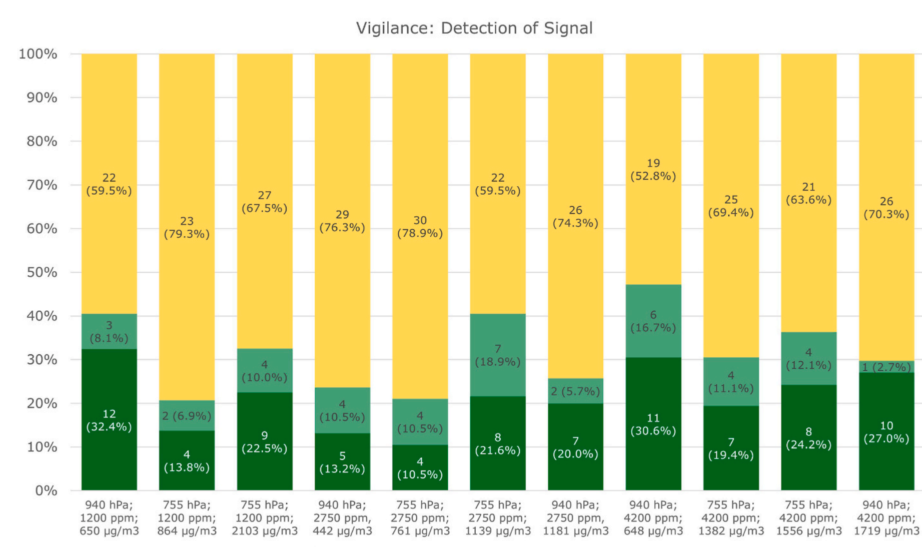

Fig. 2. Detection rates in vigilance task for different exposure groups
Note. X-axis: pressure in hPa, CO_2 in ppm, TVOC in μg/m.^{3}.

could not control exposures at the intended levels. This affects almost all exposure groups but especially two groups at low pressure levels with low and high CO_2 (Table 2). Several quality checks were performed to assess the credibility of the SIFT-MS measurements: Erroneous activation of dosing could be excluded since the syringe pump was turned off and not loaded with the VOC mixture; the cabin crew was certain not to have perceived any smell typical of the dosed pollutants (citrus smell from nonanal). Why the VOC levels were elevated on these two days remains unclear but certainly mirrors the situation that could occur indoors. Because of this problem, the strategy of statistical analysis had to be switched. While we still had enough power for the analyses, this change in the analytical approach prevented us from determining the independent main and interaction effects among the three environmental factors. However, since almost no effects were observed between the combinations, it is highly probable that no such effects exist.

Other limitations concern the generalizability of our experimental results to real indoor environments. Firstly, compounds known to be carcinogens were excluded for ethical reasons from the VOC mix, although they are often present indoors. Residual ozone (often found indoors, too, and posing a health risk) was hardly present (1-5 ppb), although it is involved in chemical reactions with bioeffluents producing VOCs other than those in the mixture (Weschler et al., 2007). Assuming that the main effects of human emissions are attributable to compounds that are products of ozone chemistry with skin oils (such as C5, C9, and C10 aldehydes and 6-MHO) (Tsushima et al., 2018), all of which were present even at this low ozone concentration (Supplementary Table S1b), adding some of them to cabin air would generally result in less than a fourfold increase in concentration, i.e., to the level of approximately 2.5 L/s per person (compared to the realized ventilation rate of 9.4 L/s). According to Zhang et al. (2017), at these levels, no effects of exposure to bioeffluents on cognitive performance are expected, as observed in our study. Furthermore, Tsushima et al. (2018) reported weak sensory effects of exposure to dermally emitted bioeffluents at a ventilation rate of approximately 2 L/s per person and at concentrations higher than those in our study, which could be one reason why we found no effects on air quality assessment (Fig. 1).

The lack of effects on self-rated air quality at the end of the exposure could also indicate some sort of habituation effect. Together with the comparably low, but typical relative humidity, that was held constant across groups and could have reduced odor perception, olfactory fatigue/adaptation is a well-known short-term phenomenon (Pellegrino et al., 2017). It may have played a role in relation to VOC effects. However, regarding the association between CO_2 and cognitive performance, it remains unclear whether people exposed to high CO_2 concentrations over a prolonged time period (e.g., in their daily work) can develop a tolerance that offsets any potential negative impact. Results from studies with miners (Maniscalco et al., 2022) and submarine-qualified sailors (Rodeheffer et al., 2018) could hint at this; however, such long-term effects were not explored in our study.

Secondly, the length of exposure was approximately two and a half to 3 h, which is significantly shorter than the normal time spent indoors (Klepeis et al., 2001) but should be valid for normal periods in public transportation, cars and airplanes. Thirdly, environmental factors that were controlled but not varied in the experiment, such as temperature, can vary greatly in real-life situations and also have an impact. Fourth, the cabin was only occupied to half of its capacity for pragmatic reasons, providing plenty of space for the seated participants, which is not typical for all indoor situations. Herbig et al. (2023) showed in the same FTF that indoor air quality manipulated by different fractions of recirculation air had an influence on physiological stress only under experimental conditions with close proximity to others but not under spacious conditions such as the one in this study. While the exact nature of this effect is not yet clear and bioeffluents in proximity might play a role, there are several associations between proximity of others, experienced stress, performance and other behavioral reactions (e.g., Aiello and Douthitt, 2001; Lewis et al., 2017) such that the effects of air quality might be different in highly occupied spaces. It would be good to additionally consider personal space when evaluating the effects of indoor air parameters. A related issue concerns the environment in general. Since contextual factors, such as appearance or odors, and individual factors, such as experiences with or within a specific context, can serve as cues for cognitive functions, the generalizability of our results to other environments is unknown, and an extrapolation to other contexts cannot be done without due consideration. However, with the exception of the varied factors, all environmental parameters were the same in all experimental conditions, so that the internal validity of the results is not compromised.

Finally, while we used typical values and levels within guideline values, one might argue that, in comparison with other studies, the CO_2 baseline was too high, as some studies seem to find effects on cognition below 1000 ppm CO_2 (Allen et al., 2019; Satish et al., 2012). However, Rodeheffer et al. (2018) and Maniscalco et al. (2022) reported no effects on cognitive performance at levels as high as 15,000 and 20,000 ppm, respectively, and one of our motivations for the study was to use realistic levels to draw practical implications.

Overall, our results confirm that, under typical and under realistic but above-average CO_2 levels and for some VOCs examined in the present study, no adverse effects on the cognitive performance in healthy adults were observed. Caution is warranted in generalizing to other VOC mixtures, longer exposures, or vulnerable populations. Future work should pursue long-term studies incorporating broader VOC mixtures (including building materials and bioeffluents), varied real-world environments, and susceptible groups, to better understand the relation between different air quality parameters and cognition, and – practically - to inform adaptive ventilation strategies that balance energy efficiency with occupant cognitive health.

CRediT authorship contribution statement

Britta Herbig: Writing – original draft, Visualization, Supervision, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Florian Mayer:** Writing – review & editing, Resources, Project administration, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization. **Victor Norrefeldt:** Writing – review & editing, Resources, Project administration, Methodology, Investigation, Conceptualization. **Pawel Wargocki:** Writing – review & editing, Resources, Methodology, Formal analysis, Conceptualization.

Data availability statement

The deidentified data can be made available upon reasonable request to the corresponding author.

Funding

This study received cash funding from Honeywell Aerospace UK, Airbus Operations GmbH, Pall Aeropower Corporation, Embraer S.A., Liebherr Aerospace Toulouse SAS and was co-funded by ASHRAE. The funding sources had no role in the study design; the collection, analysis, and interpretation of data; the writing of the report; and the decision to submit the article for publication.

Acknowledgments

We want to thank Ines Englmann, Ivana Ivandic, Anna Reichherzer, Ria Ströhlein, and Anne-Grit Eisenschmid for their help with the data collection and pre-processing as well as the LMU physicians Drs. Matthias Demmeler, Barbara Heiden, Veronika Hofbauer, and Carolin Quartucci for their help with participants' medical screening and monitoring of tests. We thank also Maximilian Kienberger and Marie Pschirer for controlling the Fraunhofer Flight Test Facility during the exposure tests, and Maximilian Taubenberger and Michael Buschhaus for creation of the VOC atmospheres in the cabin during the exposure tests.

We would also like to thank the National Physical Laboratory NPL, Teddington, UK (namely David Butterfield and Gabriel Garcia) very much for measuring and analysing the air quality parameters during the tests (especially VOCs), and Prof. Lei Fang for his help in the initial set-up of the protocol.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijheh.2026.114809.

References

Abramson, J.H., 2011. WINPEPI updated: computer programs for epidemiologists, and their teaching potential. Epidemiol. Perspect. Innovat. 8, 1. https://doi.org/10.1186/1742-5573-8-1.

Aiello, J.R., Douthitt, E.A., 2001. Social facilitation from triplett to electronic performance monitoring. Group Dyn. 5, 163–180. https://doi.org/10.1037/1089-2699.5.3.163.

Allen, J.G., MacNaughton, P., Satish, U., Santanam, S., Vallarino, J., Spengler, J.D., 2016. Associations of cognitive function scores with carbon dioxide, ventilation, and volatile organic compound exposures in office workers: a controlled exposure study of green and conventional office environments. Environ. Health Perspect. 124, 805–812. https://doi.org/10.1289/ehp.1510037.

Allen, J.G., MacNaughton, P., Cedeno-Laurent, J.G., Cao, X., Flanigan, S., Vallarino, J., et al., 2019. Airplane pilot flight performance on 21 maneuvers in a flight simulator under varying carbon dioxide concentrations. J. Expo. Sci. Environ. Epidemiol. 29, 457–468. https://doi.org/10.1038/s41370-018-0055-8.

Andersson, K., Bakke, J.V., Bjørseth, O., Bornehag, C.G., Clausen, G., Hongslo, J.K., et al., 1997. TVOC and health in non-industrial indoor environments. Indoor Air 7, 78–91. https://doi.org/10.1111/j.1600-0668.1997.t01-2-00002.x.

Agöf, Arbeitsgemeinschaft ökologischer Forschungsinstitute e.V., 2013. Guidance values for volatile organic compounds in indoor air. https://www.agoef.de/orientierungswerte/agoef-voc-orientierungswerte. (Accessed 24 November 2025).

Azuma, K., Kagi, N., Yanagi, U., Osawa, H., 2018. Effects of low-level inhalation exposure to carbon dioxide in indoor environments: a short review on human health and psychomotor performance. Environ. Int. 121 (Pt 1), 51–56. https://doi.org/10.1016/j.envint.2018.08.059.

Bælum, J., Andersen, I.B., Lundqvist, G.R., Mølhave, L., Pedersen, O.F., Vaeth, M., et al., 1985. Response of solvent-exposed printers and unexposed controls to six-hour toluene exposure. Scand. J. Work. Environ. Health 11, 271–280. https://doi.org/10.5271/sjweh.2221.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B Methodol 57 (1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

Benzo, M., Gilardoni, G., Gandini, C., Caccialanza, G., Vita Finzi, P., Vidari, G., et al., 2007. Determination of the threshold odor concentration of main odorants in essential oils using gas chromatography-olfactometry incremental dilution technique. J. Chromatogr. A 1150 (1–2), 131–135. https://doi.org/10.1016/j.chroma.2007.02.031.

Breuer, K., Streufert, S., 1995. The strategic management simulation (SMS): a case comparison analysis of the German SMS version. In: Mulder, M., Brinkerhoff, R.O. (Eds.), Corporate Training for Effective Performance. Kluwer, Boston, pp. 1–17.

Brickenkamp, R., Schmidt-Atzert, L., Liepmann, D., 2010. d2-R: Test d2 - Revision. Hogrefe, Göttingen.

Cao, X., Li, P., Zhang, J., Pang, L., 2022. Associations of human cognitive abilities with elevated carbon dioxide concentrations in an enclosed chamber. Atmosphere 13, 891. https://doi.org/10.3390/atmos13060891.

Chen, D., Huebner, G., Bagkeris, E., Ucci, M., Mumovic, D., 2023. Effects of short-term exposure to moderate pure carbon dioxide levels on cognitive performance, health symptoms and perceived indoor environment quality. Build. Environ. 245, 110967. https://doi.org/10.1016/j.buildenv.2023.110967.

Chen, R., Fang, L., Liu, J., Herbig, B., Norrefeldt, V., Mayer, F., 2021. Cabin air quality on non-smoking commercial flights: a review of published data on airborne pollutants. Indoor Air 31, 926–957. https://doi.org/10.1111/ina.12831.

Cohen, P., Cohen, J., Aiken, L.S., West, S.G., 1999. The problem of units and the circumstances for POMP. Multivariate Behav. Res. 34, 315–346. https://doi.org/10.1207/S15327906MBR3403_2.

Davranche, K., Casini, L., Arnal, P.J., Rupp, T., Perrey, S., Verges, S., 2016. Cognitive functions and cerebral oxygenation changes during acute and prolonged hypoxic exposure. Physiol. Behav. 164 (Pt A), 189–197. https://doi.org/10.1016/j.physbeh.2016.06.001.

Du, B., Tandoc, M.C., Mack, M.L., Siegel, J.A., 2020. Indoor CO_2 concentrations and cognitive function: a critical review. Indoor Air 30, 1067–1082. https://doi.org/10.1111/ina.12706.

FAA (Federal Aviation Administration), 2019. Federal Aviation Regulations (FAR) part 25 Airworthiness standards: transport category airplanes. https://www.ecfr.gov/current/title-14/chapter-I/subchapter-C/part-25. (Accessed 24 November 2025).

Fan, Y., Cao, X., Zhang, J., Lai, D., Pang, L., 2023. Short-term exposure to indoor carbon dioxide and cognitive task performance: a systematic review and meta-analysis. Build. Environ. 237, 110331. https://doi.org/10.1016/j.buildenv.2023.110331.

Faul, F., Erdfelder, E., Buchner, A., Lang, A.G., 2009. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. Behav. Res. Methods 41, 1149–1160. https://doi.org/10.3758/BRM.41.4.1149.

Fisk, W., Wargocki, P., Zhang, X., 2019. Do indoor CO_2 levels directly affect perceived air quality, health, or work performance? ASHRAE J. 61, 70–77.

Flagner, S., Meissner, T., Künn, S., Eichholtz, P., Kok, N., Kramer, R., et al., 2025. Cognition, economic decision-making, and physiological response to carbon dioxide. Indoor Environ. 2, 100074. https://doi.org/10.1016/j.indenv.2025.100074.

Flughafenverband, A.D.V., 2018. *Airport Travel Survey 2018*. Zahlen, Fakten und Trends. https://www.adv.aero/wp-content/uploads/2016/02/Airport-Travel-Survey-2018_Brosch%C3%BCre.pdf. (Accessed 24 November 2025).

Gendolla, G.H.E., Krüsken, J., 2002. Informational mood impact on effort-related cardiovascular response: moods' diagnostic value counts. Emotion 2, 251–262. https://doi.org/10.1037/1528-3542.2.3.251.

Golden, C.J., 1975. A group version of the Stroop Color and Word Test. J. Pers. Assess. 39, 386–388. https://doi.org/10.1207/s15327752jpa3904_10.

Grimes, D.A., Schulz, K.F., 2002a. An overview of clinical research: the lay of the land. Lancet 359, 57–61. https://doi.org/10.1016/S0140-6736(02)07283-5.

Grimes, D.A., Schulz, K.F., 2002b. Bias and causal associations in observational research. Lancet 359, 248–252. https://doi.org/10.1016/S0140-6736(02)07451-2.

Guan, J., Li, Z., Yang, X., 2015. Net in-cabin emission rates of VOCs and contributions from outside and inside the aircraft cabin. Atmos. Environ. 111, 1–9. https://doi.org/10.1016/j.atmosenv.2015.04.002.

Haverinen-Shaughnessy, U., Moschandreas, D.J., Shaughnessy, R.J., 2011. Association between substandard classroom ventilation rates and students' academic achievement. Indoor Air 21, 121–131. https://doi.org/10.1111/j.1600-0668.2010.00686.x.

Herbig, B., Norrefeldt, V., Mayer, F., Reichherzer, A., Lei, F., Wargocki, P., 2023. Effects of increased recirculation air rate and aircraft cabin occupancy on passengers' health and well-being - results from a randomized controlled trial. Environ. Res. 216 (Pt 4), 114770. https://doi.org/10.1016/j.envres.2022.114770.

Honan, D., Gallagher, J., Garvey, J., Littlewood, J., 2024. Indoor air quality in naturally ventilated primary schools: a systematic review of the assessment & impacts of CO_2 levels. Build 14, 4003. https://doi.org/10.3390/buildings14124003.

Jacobson, T.A., Kler, J.S., Hernke, M.T., Braun, R.K., Meyer, K.C., Funk, W.E., 2019. Direct human health risks of increased atmospheric carbon dioxide. Nat. Sustain. 2, 691–701. https://doi.org/10.1038/s41893-019-0323-1.

Kiesswetter, E., Sietmann, B., Golka, K., Zupanic, M., Seeber, A., 1997. Discriminant validity of a new questionnaire for chemical and general environment sensitivity. Neurotoxicol 18, 902–903.

Klepeis, N.E., Nelson, W.C., Ott, W.R., Robinson, J.P., Tsang, A.M., Switzer, P., et al., 2001. The National Human Activity Pattern Survey (NHAPS): a resource for assessing exposure to environmental pollutants. J. Expo. Anal. Environ. Epidemiol. 11, 231–252. https://doi.org/10.1038/sj.jea.7500165.

Lax, T.L.G., 2020. The effects of elevated indoor carbon dioxide levels on cognitive performance and physiology: a meta-analysis. TU Eindhoven: MSc thesis. https://research.tue.nl/en/studentTheses/the-effects-of-elevated-indoor-carbon-dioxide-levels-on-cognitive. (Accessed 24 November 2025).

Lewis, L., Patel, H., D'Cruz, M., Cobb, S., 2017. What makes a space invader? Passenger perceptions of personal space invasion in aircraft travel. Ergon 60, 1461–1470. https://doi.org/10.1080/00140139.2017.1313456.

Liepmann, D., Beauducel, A., Brocke, B., Amthauer, R., 2007. I-S-T 2000 R (Intelligenz-Struktur-Test). Hogrefe, Göttingen.

Lowe, R.J., Huebner, G.M., Oreszczyn, T., 2018. Possible future impacts of elevated levels of atmospheric CO_2 on human cognitive performance and on the design and operation of ventilation systems in buildings. Build Serv Eng Res Technol 39, 698–711. https://doi.org/10.1177/0143624418790.

Lowther, S.D., Dimitroulopoulou, S., Foxall, K., Shrubsole, C., Cheek, E., Gadeberg, B., et al., 2021. Low level carbon dioxide indoors - a pollution indicator or a pollutant? A health-based perspective. Environ. Times 8, 125. https://doi.org/10.3390/environments8110125.

Maddalena, R., Mendell, M.J., Eliseeva, K., Chan, W.R., Sullivan, D.P., Russell, M., et al., 2015. Effects of ventilation rate per person and per floor area on perceived air quality, sick building syndrome symptoms, and decision-making. Indoor Air 25, 362–370. https://doi.org/10.1111/ina.12149.

Maniscalco, J., Hoffmeyer, F., Monsé, C., Jettkant, B., Marek, E., Brüning, T., et al., 2022. Physiological responses, self-reported health effects, and cognitive performance during exposure to carbon dioxide at 20.000 ppm. Indoor Air 32, e12939. https://doi.org/10.1111/ina.12939.

Martins, C., Teófilo, V., Clemente, M., Corda, M., Fermoso, J., Aguado, A., et al., 2025. Sources, levels, and determinants of indoor air pollutants in Europe: a systematic review. Sci. Total Environ. 964, 178574. https://doi.org/10.1016/j.scitotenv.2025.178574.

Mendell, M.J., Chen, W., Ranasinghe, D.R., Castorina, R., Kumagai, K., 2024. Carbon dioxide guidelines for indoor air quality: a review. J. Expo. Sci. Environ. Epidemiol. 34, 555–569. https://doi.org/10.1038/s41370-024-00694-7.

Mishra, A.K., Schiavon, S., Wargocki, P., Tham, K.W., 2020. Carbon dioxide and its effect on occupant cognitive performance: a literature review. In: Roaf, S., Nicol, F., Finlayson, W. (Eds.), Proceedings of the 11th Windsor Conference on Thermal Comfort: Resilient Comfort, pp. 440–452. In: https://windsorconference.com/wp-content/uploads/2020/05/WC2020_Proceedings_final_compressed.pdf. (Accessed 24 November 2025).

Mølhave, L., Bach, B., Pedersen, O.F., 1986. Human reactions to low concentrations of volatile organic compounds. Environ. Int. 12, 167–175. https://doi.org/10.1016/0160-4120(86)90027-9.

Nagata, Y., 2003. Measurement of odor threshold by triangle odor bag method. Odor Meas Rev 118, 118–127.

Norrefeldt, V., Mayer, F., Herbig, B., Ströhlein, R., Wargocki, P., Lei, F., 2021. Effect of increased cabin recirculation airflow fraction on relative humidity, CO_2 and TVOC. Aerospace 8, 15. https://doi.org/10.3390/aerospace8010015.

OSHA (Occupational Safety and Health Administration), 2020. Permissible exposure limits – Annotated tables, table Z-1. https://www.osha.gov/annotated-pels/table-z-1. (Accessed 24 November 2025).

Oswald, W.D., Roth, E., 2016. Der Zahlen-Verbindungs-Test (ZVT), third ed. Hogrefe, Göttingen.

Pang, L., Zhang, J., Cao, X., Wang, X., Liang, J., Zhang, L., et al., 2021. The effects of carbon dioxide exposure concentrations on human vigilance and sentiment in an enclosed workplace environment. Indoor Air 31, 467–479. https://doi.org/10.1111/ina.12746.

Pellegrino, R., Sinding, C., de Wijk, R.A., Hummel, T., 2017. Habituation and adaptation to odors in humans. Physiol. Behav. 177, 13–19. https://doi.org/10.1016/j.physbeh.2017.04.006.

Pfeiffer, J.W., Jones, J.E., 1970. A Handbook of Structured Experiences for Human Relations Training, I + II. University Associates Press, Iowa City.

Reeve, C.L., Heggestad, E.D., Lievens, F., 2009. Modeling the impact of test anxiety and test familiarity on the criterion-related validity of cognitive ability tests. Intell 37, 34–41. https://doi.org/10.1016/j.intell.2008.05.003.

Rodeheffer, C.D., Chabal, S., Clarke, J.M., Fothergill, D.M., 2018. Acute exposure to low-to-moderate carbon dioxide levels and submariner decision making. Aerosp. Med. Hum. Perform. 89, 520–525. https://doi.org/10.3357/AMHP.5010.2018.

Roenneberg, T., 2015. Having trouble typing? What on Earth is chronotype? J Biol Rhythms 30, 487–491. https://doi.org/10.1177/0748730415603835.

Satish, U., Mendell, M.J., Shekhar, 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. Environ. Health Perspect. 120, 1671–1677. https://doi.org/10.1289/ehp.1104789.

Simons, D.J., Chabris, C.F., 1999. Gorillas in our midst: sustained inattentional blindness for dynamic events. Perception 28, 1059–1074. https://doi.org/10.1068/p281059.

Snow, S., Boyson, A.S., Paas, K.H.W., Gough, H., King, M.F., Barlow, J., et al., 2019. Exploring the physiological, neurophysiological and cognitive performance effects of elevated carbon dioxide concentrations indoors. Build. Environ. 156, 243–252. https://doi.org/10.1016/j.buildenv.2019.04.010.

Stroop, J.R., 1935. Studies of interference in serial verbal reactions. J. Exp. Psychol. 18, 643–662. https://doi.org/10.1037/h0054651.

Tsushima, S., Wargocki, P., Tanabe, S., 2018. Sensory evaluation and chemical analysis of exhaled and dermally emitted bioeffluents. Indoor Air 28, 146–163. https://doi.org/10.1111/ina.12424.

Twardella, D., Matzen, W., Lahrz, T., Burghardt, R., Spegel, H., Hendrowarsito, L., et al., 2012. Effect of classroom air quality on students' concentration: results of a cluster-randomized cross-over experimental study. Indoor Air 22, 378–387. https://doi.org/10.1111/j.1600-0668.2012.00774.x.

Umweltbundesamt, 2025. Assessment of chemical indoor air pollution on the basis of indoor air measurement results. Guideline of the German Committee on Indoor Air Guide Values (AIR). Bundesgesundheitsblatt - Gesundheitsforsch. - Gesundheitsschutz 68, 190–200. https://doi.org/10.1007/s00103-024-03999-y.

Valdez, P., Ramírez, C., García, A., 2012. Circadian rhythms in cognitive performance: implications for neuropsychological assessment. ChronoPhysiol. Ther. 2, 81–92. https://doi.org/10.2147/CPT.S32586.

van Gemert, L.J., 2011. Odour Thresholds – Compilations of Odour Threshold Values in Air, Water and Other Media, second ed. https://www.academia.edu/117947016/ODOUR_THRESHOLDS_EDITION. (Accessed 10 February 2026).

Wang, C., Zhang, F., Wang, J., Doyle, J.K., Hancock, P.A., Mak, C.M., et al., 2021. How indoor environmental quality affects occupants' cognitive functions: a systematic review. Build. Environ. 193, 107647. https://doi.org/10.1016/j.buildenv.2021.107647.

Wang, S., Qin, T., Tu, R., Li, T., Chen, G.I., Green, D.C., et al., 2024. Indoor air quality in subway microenvironments: pollutant characteristics, adverse health impacts, and population inequity. Environ. Int. 190, 108873. https://doi.org/10.1016/j.envint.2024.108873.

Ware, J.E., Kosinski, M., Dewey, J.E., Gandek, B., 2001. How to Score and Interpret single-item Health Status Measures: a Manual for Users of the SF-8 Health Survey. QualityMetric Inc., Lincoln, RI.

Wargocki, P., 2001. Measurements of the effects of air quality on sensory perception. Chem. Senses 26, 345–348. https://doi.org/10.1093/chemse/26.3.345.

Wargocki, P., Lagercrantz, L., Witterseh, T., Sundell, J., Wyon, D.P., Fanger, P.O., 2002. Subjective perceptions, symptom intensity and performance: a comparison of two independent studies, both changing similarly the pollution load in an office. Indoor Air 12, 74–80. https://doi.org/10.1034/j.1600-0668.2002.01101.x.

Wargocki, P., Wyon, D.P., 2006. Effects of HVAC on student performance. ASHRAE J. 48, 23–28.

Wargocki, P., Wyon, D.P., 2007. The effects of outdoor air supply rate and supply air filter condition in classrooms on the performance of schoolwork by children. HVAC R Res. 13, 165–191. https://doi.org/10.1080/10789669.2007.10390950.

Wargocki, P., Wyon, D.P., 2017. Ten questions concerning thermal and indoor air quality effects on the performance of office work and schoolwork. Build. Environ. 112, 359–366. https://doi.org/10.1016/j.buildenv.2016.11.020.

Wargocki, P., Wyon, D.P., Baik, Y.K., 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. https://doi.org/10.1111/j.1600-0668.1999.t01-1-00003.x.

Weschler, C.J., Wisthaler, A., Cowlin, S., Tamás, G., Strøm-Tejsen, P., Hodgson, A.T., et al., 2007. Ozone-initiated chemistry in an occupied simulated aircraft cabin. Environ. Sci. Technol. 41, 6177–6184. https://doi.org/10.1021/es0708520.

Wolkoff, P., Azuma, K., Carrer, P., 2021. Health, work performance, and risk of infection in office-like environments: the role of indoor temperature, air humidity, and ventilation. Int. J. Hyg Environ. Health 233, 113709. https://doi.org/10.1016/j.ijheh.2021.113709.

Zhang, X., Wargocki, P., Lian, Z., Xie, J., Liu, J., 2017. Responses to human bioeffluents at levels recommended by ventilation standards. Procedia Eng. 205, 609–614. https://doi.org/10.1016/j.proeng.2017.10.415.