I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17019

Cognition, Economic Decision-Making, and Physiological Response to Indoor Carbon Dioxide: Does It Really Matter?

Stefan Flagner
Thomas Meissner
Steffen Künn
Piet Eichholtz
Nils Kok
Rick Kramer
Wouter van Marken-Lichtenbelt
Cynthia Ly
Guy Plasqui

MAY 2024

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17019

Cognition, Economic Decision-Making, and Physiological Response to Indoor Carbon Dioxide: Does It Really Matter?

Stefan Flagner, Thomas Meissner
Maastricht University

Steffen Künn
Maastricht University and IZA

Piet Eichholtz, Nils Kok
Maastricht University

Rick Kramer
Eindhoven University of Technology

Wouter van Marken-Lichtenbelt
Maastricht University

Cynthia Ly, Guy Plasqui
Maastricht University

MAY 2024

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Cognition, Economic Decision-Making, and Physiological Response to Indoor Carbon Dioxide: Does It Really Matter?

This study provides novel evidence on the isolated effect of carbon dioxide on cognition, economic decision-making, and the physiological response in healthy office workers. The experiment took place in an air-tight respiration chamber fully controlling the environmental conditions. In a single-blind, within-subject study design, 20 healthy participants were exposed to carbon dioxide concentrations of 3,000 ppm and 900 ppm in randomized order, with each exposure lasting for 8 hours. We do not find evidence on a statistically significant effect on either cognitive or physiological outcome variables. Thus, the evidence shows that the human body appears to be able to deal with exposure to indoor carbon dioxide concentration of 3,000 ppm without suffering significant cognitive decline, changes in decision-making or showing any physiological response.

JEL Classification: D87, J24, Q54

Keywords: carbon dioxide, indoor air quality, cognition, economic decision-making, physiological response

Corresponding author:
Stefan Flagner
Maastricht University
P.O. Box 616
6200 MD Maastricht
The Netherlands
E-mail: s.flagner@maastrichtuniversity.nl

1 Introduction

The real estate sector is regarded as a major factor in global energy usage and CO_2 emissions, significantly contributing to climate change. Given that half of the energy consumption in buildings is attributed to heating, cooling, and ventilation (1), there is an increased policy awareness to improve the energy efficiency of buildings (2). Modern buildings often have automated ventilation systems that adapt the ventilation rate to maintain an adequate indoor air quality. A recent article written by 43 experts in the field of indoor air quality emphasizes the importance for human health, productivity and learning to provide good air quality (3). However, increasing the ventilation rate leads to higher energy costs from fans operating at higher speed and higher energy demand for heating and air conditioning to maintain a stable indoor temperature during cold or hot weather (4; 5; 6). Building owners thus face a profound trade-off between providing healthy indoor air quality and improving the energy efficiency of buildings.

While energy consumption is tangible and its impact can be quantified, both financially and in terms of environmental impact, indoor air quality is more complex to measure, and the implications of variation in indoor air quality are harder to assess. The existing literature provides some evidence that poor indoor air quality, caused by limited supply of fresh, outside air into the room, leads to reduced cognitive performance, declining productivity, and adverse health effects for office workers (7; 8). However, air quality is a function of different components, which are correlated with each other. The negative effects of poor indoor air quality on cognition and health could be associated with volatile organic compounds and bioeffluents, which are by-products of human metabolism and are exhaled together with CO_2 (7; 9). Therefore, another stream of literature aims to estimate the isolated effect of CO_2 on health and cognitive parameters, showing mixed results. These studies are quite heterogeneous in terms of participant population, exposure time, outcome measurement, and size of groups measured at the same time. Most importantly, existing studies insufficiently control for other environmental factors which could also affect the outcomes of interest. They either do not use an air-tight climate chamber (10; 11; 12) or do not control for concentrations of volatile organic compounds during testing (9; 10; 12; 13; 14; 15; 16; 17). This makes it challenging to draw unambiguous conclusions about the effect of indoor air quality, as measured by CO_2, on human cognition and physiology.

This study examines the isolated effect of CO_2 on a broad set of outcomes in a tightly controlled environment. Using an interdisciplinary approach, we investigated the effect of an 8-hour exposure to CO_2 levels of 3,000 ppm, as compared to CO_2 levels of 900 ppm, on the

cognitive performance, economic decision-making, and physiological response of healthy office workers, keeping potentially confounding factors constant. The experiment was conducted in an air-tight respiration chamber, which is commonly used in metabolic research (18).

We hypothesize that an elevated ambient CO_2 level leads to lower cognitive performance. Cognition is measured using a neuropsychological test battery to test the domains of attention, psychomotor control, executive function, and memory. We also investigated the impact on individuals' risk and time preferences by applying multiple price lists, a tool common in the economics literature, to analyse the effect of CO_2 on economic decision-making (19; 20). It is assumed that elevated CO_2 levels lead to changes in risk-taking and time preferences. Risk and time preferences have been found to predict economic decision-making in a wide variety of settings and are often found to correlate with cognition (21). Moreover, we consider the physiological response to elevated CO_2 levels. Any effect on cognitive performance or economic decision-making is assumed to be mediated by a physiological stress reaction, such as a higher heart rate, higher blood pressure, higher physical activity level, or elevated oxygen consumption. It is further hypothesized that the respiration rate will decrease as a response to the higher indoor CO_2 concentration and this is associated with a higher blood CO_2 concentration, as has been documented in earlier work (22; 23). This would indicate that elevated levels of CO_2 lead to a respiratory acidosis caused by a change in the breathing pattern. Given the large number of hypotheses tested in the empirical analysis, we draw inferences based on p-values corrected for multiple hypothesis testing.

2 Methods

2.1 Experimental setup

This study is based on a cohort of 20 healthy individuals and was conducted in the period from November 2021 to July 2022. Participants were exposed to CO_2 levels of 3,000 ppm (High-CO_2) and 900 ppm (Low-CO_2) during an 8-hour stay in a respiration chamber of the Metabolic Research Laboratories at Maastricht University (18). Only one participant at a time was measured in the chamber, because behaviour and perception can be influenced if occupants are measured in groups (24). We applied a cross-over design where participants were exposed to both CO_2 conditions, with a break of four to six weeks between the two test days. The order of both test conditions was randomized with 10 participants starting in the High-CO_2 and 10 participants starting in the Low-CO_2 condition. Participants were blinded to the condition. The ventilation

rate was set at a high level to ensure low levels of volatile organic compounds and fine particle matter accumulation in the chamber. Additionally, a particle and volatile organic compounds filter was installed (Molekule Air Mini+, Molekule Florida, USA). The remaining fine particular matter ($PM_{2.5}$) and volatile organic compounds were measured on a 5-minute basis (Foobot SAT, Airboxlab SA, Luxembourg). In the High-CO_2 condition, after participants entered the chamber, CO_2 was induced via a gas bottle until it reached a stable level of 3,000 ppm, which took on average 11 minutes. After the 3,000 ppm concentration was achieved, the infusion of additional CO_2 was lowered to a level sufficient to maintain the steady state concentration of 3,000 ppm for the remainder of the test day. Indoor temperature, relative humidity, light conditions and noise levels were kept constant in both conditions.

2.2 Recruitment of participants

Inclusion criteria were 1) to have an office job that includes mentally demanding tasks, 2) to be between 25 and 50 years old, 3) to be generally healthy, with no intake of any medication (except for contraceptives), and 4) to not smoke. Individuals with one of the following characteristics were excluded: 1) unemployed at the moment of testing, 2) having a disorder or disease, including Parkinson, Attention Deficit Hyperactivity Disorder, Alzheimer, diabetes, cardiovascular disorder, respiratory impairments, or hypertension, 3) doing sports on a professional basis or more than five times a week for more than two hours, 4) work in shift work, 5) being colour-blind, or 6) pregnant. Participants received a lump-sum compensation of €170. The final sample included eleven female and nine male participants who were on average 31 years old, ranging from 25 years old to 46 years old. The average Body Mass Index (BMI) was 23 with the lowest BMI at 20 and the highest BMI at 27. The average body height was 174 cm (minimum of 156 cm and maximum of 191 cm) and the average body weight was 69 kg (minimum of 51 kg and maximum of 92 kg). From a socio-economic perspective, participants are quite similar. All participants have a university degree, and the majority of participants had a monthly gross income between €1,000 and €5,000 (N = 17), with two participants earning between €5,000 and €7,500 and one participant earning more than €10,000 in gross monthly salary.

2.3 Outcome variables

Cognitive tests: The Cambridge Neuropsychological Test Automated Battery (CANTAB) was used to assess the cognitive functioning of participants. Participants conducted the test on a tablet computer with a touch screen. The CANTAB tests have been validated against

other neuropsychological test batteries (25). Four domains were measured, including attention, psychomotor control, memory, and executive function. For attention and psychomotor control, the Reaction Time Task and Motor Screening Task were used. The Reaction Time Task measures movement and mental response speed in milliseconds when a stimulus is presented. The Motor Screening Task assesses movement latency in milliseconds when a stimulus is presented. The Delayed Matching to Sample and Paired Associate Learning tests were used to measure memory. Delayed Matching to Sample measures visual matching ability and short-term visual recognition memory as a percentage of correct choices. Paired Associate Learning assesses visual memory and learning as the number of correct responses. To measure executive function, the Multitasking Test, One Touch Stocking of Cambridge, Stop Signal Task, and Spatial Working Memory were used. The Multitasking Test measures the ability to manage conflicting information as the time (in milliseconds) that a participant needs to give the correct response when two contradicting stimuli are presented. One Touch Stocking of Cambridge assesses spatial planning and working memory as the number of problems solved on the first attempt. Stop Signal Task measures impulse control as the time (in milliseconds) it takes for participants to inhibit a reaction when the test initially asks for a reaction. Last, Spatial Working Memory measures strategy and working memory errors as the number of incorrect revisions from finding a specific figure among several covered fields. A detailed description of the CANTAB tests can be found on the CANTAB website (26). To ensure a balanced loading of the different cognitive domains, the same order of testing was applied during the test days: Start with the Reaction Time Task, then Paired Associate Learning, Stop Signal Task, and Spatial Working Memory , followed by a 10-minute break. After the break, the testing continued with Motor Screening Task, One Touch Stocking of Cambridge, Delayed Matching to Sample, and Multitasking Test. The eight tests, including, the 10-minute break, took approximately 60 minutes. The CANTAB test was conducted twice during a test day, first after 30 minutes of exposure and then at the end of the day, after 330 minutes of exposure.

Economic decision-making: In addition to general cognition, we tested how varying CO_2 levels affect economic risk and time preferences, specifically risk aversion (hereafter called risk preferences) and the level of impatience when delaying a financial payment (hereafter called time preferences). To elicit risk and time preferences, we employed multiple price lists (MPLs), as introduced by Holt and Laury (20) for risk preferences, and by Coller and Williams (19) for time preferences. We used a total of six multiple price lists, each containing a total of ten choices between two options labeled neutrally as A and B. These choices were over either lotteries for

risk preferences or inter-temporal prospects for time preferences defined over monetary payoffs. Supplementary Tables S1 and S2 summarize all multiple price lists used in this experiment. To elicit risk preferences, participants repeatedly chose between lotteries with differing levels of risk in MPL1.1, MPL1.2, MPL2.1 and MPL2.2. To elicit time preferences, participants repeatedly chose between varying monetary payoffs at different points in time in MPL3.1 to MPL3.2. The order of displayed choices within each multiple price lists was randomized, while the sequence between the six multiple price lists stayed fixed. To incentivise participants to reveal their true preferences, they were informed beforehand that at the end of each test day, one of the 60 presented choices would be randomly drawn, and they would receive the corresponding payments in cash at the end of the test day. If the randomly selected decision contained a choice between lotteries (risk preferences), a coin was flipped to determine the outcome of the chosen lottery. If the randomly selected decision contained an inter-temporal choice (time preferences), participants would receive money at the end of the test day, or rather one month later, depending on their choice.

Physiological outcomes: To examine the physiological responses and thus the potential mechanism of how indoor CO_2 levels can impair cognition, several outcome parameters were measured on a minute-by-minute basis throughout the 8 hours of each test day. Blood CO_2 concentration was measured continuously with a transcutaneous monitor (SenTec, Therwil, Switzerland), which was also used in a previous study (22). For this reason, a non-invasive sensor was attached to the forehead. The software V-STATS was used to derive the data (version 5.01, SenTec AG, Switzerland). Due to measurement errors that occurred irregularly and for a short time, the blood CO_2 concentration data was cleaned in two steps. First, the monitor also measured the saturation level of oxygen, which given the respiration chamber's condition (sea level atmospheric pressure and no exercising) should stay above 95 percent (27). Thus, minute values of the partial pressure of CO_2 were removed if the saturation level of oxygen during the particular minute was below 95 percent, assuming a measurement error during this time. Second, all remaining values for blood CO_2 concentration below 30 mmHg and above 50 mmHg were removed as outliers, because partial pressure values of blood CO_2 are always between this range (28). This approach resulted in removing 10.4% of the minute values for blood CO_2. Additionally, heart rate and respiration rate were measured using the Polar H10 belt, which was attached around the thorax (H10, polar, USA, RR interval accuracy 99.6 % (29)). The mobile application Polar SDK developer kit for Android phones was used to extract the raw ECG data and the Kubios software (Biosignal Analysis and Medical Image Group, Department of Physics, Univer-

sity of Kuopio, Kuopio, Finland, (30)) was used to calculate the heart rate and respiration rate. Physical activity levels were measured using the three-axis activity monitor ActiGraph with a sampling frequency of 30 Hz (ActiGraph GT3X). The ActiGraph was placed on the right side of the hip. The Vector Magnitude counts per minute were derived from the raw data. Oxygen consumption was measured continuously in the respiration chamber using indirect calorimetry equipment (Omnical, Maastricht Instruments, Maastricht, NL), which measures changes in concentrations of oxygen over time to calculate the oxygen consumption of participants (18). Last, blood pressure was automatically measured every 15 minutes starting at each full hour using the Mobil-O-Graph device (I.E.M. GmbH, Stolberg, Germany).

2.4 Experimental protocol

The detailed test protocol is shown in Figure 1. Participants were exposed continuously to the testing conditions for 8 hours, from 09:00 h until 17:00 h. They were allowed to eat their own breakfast while in the room. Additionally, participants were either provided with lunch or brought their own lunch. The breakfast and lunch they ate during the first test day was documented to ensured that participants ate the same breakfast and lunch during the second test day. Food intake time was not standardized. Participants were provided with decaffeinated coffee if they requested coffee (participants were not informed that coffee was caffeine-free). Between the cognition tests, participants were free to work-related tasks, however, they were not allowed to watch TV or sleep. They were instructed to behave like during a normal day at work. To reduce any interference from a learning effect, participants also practiced the cognition test once during the screening session before the first test day.

2.5 Statistical analysis

Cognitive and physiological responses: We start the analysis by estimating the following basic linear fixed effect regression model described in Equation 1 to evaluate the effect of elevated CO_2 level on the cognitive and physiological responses of participants:

$$Y_{itd} = \eta + \delta HighCO2_{id} + \lambda_t + \theta_i + \gamma_d + \epsilon_{itd} \tag{1}$$

Where Y_{itd} is the outcome variable measured for participant i at the time of the day t at test day d. $HighCO2_{id}$ is a binary variable taking the value of one if participant i is exposed to the High-CO_2 condition (3,000 ppm) at test day d and zero otherwise. λ_t represents a set

Figure 1: Experimental Protocol

Time in testing condition (min)	Time of day (hh:mm)	Testing steps and cognition tests	Physiological continuous measurements	
0	09:00	Chamber door closed	Blood CO2 level, heart rate, respiration rate, physical activity level, oxygen consumption	Systolic and diastolic blood pressure (measured every 15 minutes starting with the full hour)
30	09:30	CANTAB test (60 Min)		
60	10:00			
90	10:30			
120	11:00	Multiple price lists (15 Min)		
150	11:30			
180	12:00			
210	12:30			
240	13:00			
270	13:30			
300	14:00			
330	14:30	CANTAB test (60 Min)		
360	15:00			
390	15:30			
420	16:00	Multiple price lists (15 Min)		
450	16:30			
480	17:00	Leaving the chamber		

of binary variables capturing the exact time of the day at which certain outcome variables are measured. The cognitive tests (CANTAB) were measured twice a day, i.e., in the morning and the afternoon, while the physiological parameters were continuously measured during the test day. We aggregated the physiological parameters to hourly averages. Therefore, we defined two different sets of dummy variables included in λ_t: For the cognitive outcome variables, we included one binary variable $Morning_t$ taking the value of one if the test was taken in the morning session and zero otherwise. For the physiological parameters, we included eight binary variables $Hour_t$ each representing one hour of the test day. In addition, we included individual fixed effect θ_i to restrict the analysis to within-participant comparisons, and test day fixed effect γ_d to capture potential learning effects when measuring the cognitive performance on the second test day for participant i. Last, Equation 1 includes the constant η and error term ϵ_{itd}.

With this model, the parameter of interest δ measures the impact of the participant's exposure to 3,000 ppm CO_2 concentration at test day d on the outcome variable Y_{itd}, in comparison to the same participant being exposed to 900 ppm CO_2 concentration during the other test day. Given the random assignment of participants to both conditions and test days, δ allows for a causal interpretation. Finally, given the low number of clusters in our study (N = 20) potentially violating the large-sample assumptions of analytical standard errors, we base our inference of standard errors on wild bootstrap clusters, as recommended in previous literature

(31). We applied 1,000 bootstrap replications clustered at the participant level to estimate the variance-covariance matrix. Given the relatively large number of hypotheses tested (8 tests for CANTAB and 7 tests for the physiological outcomes), we additionally provide adjusted p-values based on the method by Hommel to take the multiplicity of tests into account (32). We provide both the bootstrapped standard errors as well as the p-values based on multiple hypothesis testing with the estimated coefficients.

In a second step, we enhanced the baseline model 1 to allow the treatment effect δ_t to vary over the day. We estimated the following regression model where we added an interaction term to Equation 1 interacting the treatment parameter $HighCO2_{id}$ with the time of the day dummy variables λ_t.

$$Y_{itd} = \alpha + \delta_1 HighCO2_{id} + \delta_2(HighCO2_{id} \times \lambda_t) + \lambda_t + \gamma_d + \theta_i + \epsilon_{itd} \tag{2}$$

In Equation 2, δ_1 represents the difference in outcome variables between High- and Low-CO_2 conditions at reference time t_0 (Afternoon testing for CANTAB, and first test day hour for physiological parameters), and δ_2 represents the same difference, but measured at other times of the day in relation to t_0. Therefore, the analysis allows conclusions on whether the effect of CO_2 exposures on outcome variables is time-variant, e.g. testing a dose-response gradient and how it affects the outcome variables.

Economic decision-making: To analyse the impact of increased CO_2 on economic risk and time preferences, we used a maximum likelihood model (21; 33; 34) to estimate preference parameters of a discounted expected utility model, similar to (33). Utility over monetary gains is modeled assuming constant relative risk aversion (CRRA):

$$u(x) = \frac{x^{1-\alpha}}{1-\alpha} \tag{3}$$

where x denotes monetary gains, and α is the parameter of relative risk aversion, describing the curvature of the utility function; $\alpha = 0$ implies risk neutrality, $\alpha > 0$ risk aversion, and $\alpha < 0$ implies risk-seeking behavior. For $\alpha > 0$, the larger α, the larger risk aversion. Secondly, inter-temporal choices as a measure of time preferences are modelled using a simple expected discounted utility model:

$$U(x_t, ..., x_t) = E_t[u(x_t) + \sum_{k=1}^{T-t} \frac{1}{(1+\rho)_k} u(x_{t+k})] \tag{4}$$

Here, ρ is the discount factor. The larger ρ the more the future is discounted and thus the lower the willingness to wait for the future payment. Individuals become more impatient.

Participants repeatedly choose between two options, labeled A and B. We denote the expected discounted utility of options A and B as U^A and U^B, respectively. Our model allows for two types of decision errors to be as flexible as possible regarding the parametric assumptions. Specifically, decision noise is accounted for by a tremble error (κ) and a Fechner error (μ) (21; 35). The tremble error measures decision errors due to choice randomization, i.e. individuals may randomly choose between A and B with some probability κ. The Fechner error term accounts for errors in evaluating the expected utility of lotteries: options A and B are assessed based on their expected utility plus a random element ϵ, such that an individual chooses option B if $U^B + \epsilon^B > U^A + \epsilon^A$. Overall, the probability of choosing option B writes as follows:

$$P(B) = (1 - \kappa)F(\frac{U^B - U^A}{\mu}) + \frac{\kappa}{2} \tag{5}$$

Were the F, the cumulative distribution function of $(\epsilon^A - \epsilon^B)$ follows a standard logistic distribution. For $\kappa \to 0$, the tremble error has no effect on choice, and for $\kappa \to 1$, choice approaches uniform randomization. For $\mu \to 0$, the decision becomes deterministic (conditional on not choosing at random owing to the tremble error), and for $\mu \to \infty$, choice approaches uniform randomization. We estimate the preference parameters (α, ρ) and error parameters (κ, μ) of the model with maximum likelihood, using binary choice data from the multiple price lists. Parameters are estimated jointly for all participants as linear functions of the treatment dummy $HighCO_2$ and the interaction of $HighCO_2$ with a *Morning* dummy. Given that multiple price lists are asked two times per test day, once after 120 minutes of exposure and once after 420 minutes of exposure, we added a $Morning$ fixed effect, equals one if the multiple price lists were answered after 120 minutes of exposure time. Additional binary controls such as sex and whether it is the first test day were also added. The estimated coefficient of the treatment dummy thus indicates how much the estimated parameters differ across treatments.

3 Results

The results are presented in two parts. In the first part, we provide evidence on the validity of our experimental setting by showing that the CO_2 concentration differs across the two test days, while other parameters are kept constant. In addition to objective measures of indoor environmental quality, we also considered participants' subjective perceptions of the indoor

environment between the two test days. We then proceed to the main part of the results, where we present the findings regarding the impact of increased CO_2 concentration on participant's cognitive performance, economic decision-making, and physiological responses.

3.1 Treatment validation

Table 1 shows a comparison of environmental conditions between both testing conditions. In panel A, we show objective measures on the environmental conditions inside the respiration chamber for the two different CO_2 levels. The last column shows the resulting p-values, based on a simple t-test of equal means with the null hypothesis of no difference between 900 ppm CO_2 (Low-CO_2) and 3,000 ppm CO_2 (High-CO_2). We observe that the targeted average CO_2 concentration was achieved for both conditions, with an average CO_2 concentration of 918 ppm for the Low-CO_2 condition (122 ppm standard deviation) and 3011 ppm for the High-CO_2 condition (139 ppm standard deviation). In the Low-CO_2 condition, the ventilation rate was slightly higher (543 l/min) as compared to the High-CO_2 condition (525 l/min). However, the average concentrations of volatile organic compounds did not differ significantly between the Low- and High-CO_2 conditions. Additionally, temperature, relative humidity and fine particular matter $PM_{2.5}$ concentration were not significantly different between the two conditions.

Table 1: Comparison of environmental conditions

	Low-CO_2		High-CO_2		
	Mean	SD	Mean	SD	P-value
Panel A: Environmental quality					
CO_2 (ppm)	918	122	3011	139	0.000
Ventilation rate (l/min)	543	28	525	12	0.000
Volatile organic compounds (ppb)	519	365	574	317	0.156
Temperature (°C)	21	0.15	21	0.16	0.144
Relative humidity (%)	32	6	32	6	0.525
Fine particles $PM_{2.5}$ (in counts/L)	0.000	0.000	0.001	0.045	0.317
Panel B: Perception of environment					
Air quality	2.7	1.4	2.4	1.2	0.457
Temperature	3.0	1.4	3.6	1.5	0.213
Light	2.8	1.4	2.8	1.3	0.867
Noise	4.6	1.2	4.3	1.5	0.522

Note: The table shows the mean and standard deviation (SD) of objectively measured and perceived indoor environmental quality during the test day with low (900 ppm) and high (3,000 ppm) CO_2 levels. Panel A contains environmental conditions continuously measured inside the respiration chamber. Panel B shows measures reflecting participants' perceived indoor conditions collected via a questionnaire for which participants reported their satisfaction with the air quality, temperature, lighting and noise based on a scale ranging from "1 - Extremely Satisfied" to "7 - Extremely Dissatisfied". Column 1 and 3 show the average value, and column 2 and 4 the standard deviation. The last column shows the resulting p-value from a simple t-test of equal means (H_0 = no difference between High- and Low-CO_2 condition).

In addition to the comparison of the objective measures of indoor environmental conditions, we also collected information on participants' subjective perception of indoor environmental quality, based on a survey that participants had to complete shortly before they left the respiration chamber (at the end of each test day). We used an adapted version of the Center for the Built Environment (CBE) survey (36), asking participants how satisfied they were with the temperature, air quality, lighting conditions, and noise level. The satisfaction level with each item was reported based on a 7-point Likert scale ranging from "1 - Extremely satisfied" to "7 - Extremely dissatisfied". Panel B in Table 1 shows the mean comparisons for these variables. There were no statistically significant differences ($p > 0.05$), which indicates that participants did not perceive indoor environmental quality differently in the Low or High-CO_2 condition. This confirms the participants were successfully blinded to the testing conditions. The evidence shows no difference in objective or subjective measures of indoor environmental quality, except for the concentration of CO_2. This confirms the validity of the experimental setting, thereby allowing us to relate the outcome measures causally to CO_2 exposure.

3.2 Cognitive responses

Our main findings focus on the impact of exposure to 3,000 ppm CO_2 on cognitive performance and economic decision-making, starting with the effect on general cognitive abilities, as assessed by the Cambridge Neuropsychological Automated Test Battery (CANTAB). We provide the estimated treatment effects based on the fixed effects regression Model 1 in Table 2. The treatment coefficient $HighCO_2$ is defined as a dummy variable which is either 1 if the corresponding cognition test was done under the 3,000 ppm CO_2 exposure, or zero if it was conducted in the 900 ppm CO_2 condition. The regression results allow for direct inferences regarding the average difference in outcome variables between both testing conditions. In addition, the results of the interacted regression Model 2 in Table 2 assess the heterogeneity over the course of the day, describing a possible dose-response relationship of the effect of CO_2 on cognitive performance.

Focusing on the first model specification (without interaction terms), we do not find any statistically significant effect of elevated CO_2 concentrations on the cognitive domains of psychomotor control and attention, executive function, and the Paired Associate Learning task among the memory tasks. We document a statistically significant effect (at the 5%-level) for the Delayed Matching to Sample task, suggesting that exposure to 3,000 ppm improves participants' share of correct choices by 3.55%-points. However, based on the corrected p-values for multiple hypotheses testing, the effect becomes insignificant ($p = 0.2$).

Table 2: Elevated indoor CO_2 and CANTAB test scores

	Panel A: Attention & psychomotor control				Panel B: Memory			
	Reaction Time Task		Motor Screening Task		Delayed Matching to Sample		Paired Associate Learning	
High CO_2	3.587	2.025	0.493	4.445	3.550*	4.050	0.500	1.550**
	(4.622)	(6.064)	(12.651)	(16.93)	(1.518)	(2.432)	(0.288)	(0.552)
	[0.871]	[0.798]	[0.968]	[0.798]	[0.200]	[0.501]	[0.483]	[0.056]
High CO_2 x Morning		3.125		-7.905		-1.000		-2.100*
		(8.910)		(7.905)		(3.405)		(0.915)
		[0.960]		[0.960]		[0.960]		[0.184]
Observations	80	80	80	80	80	80	80	80
R^2	0.897	0.897	0.625	0.625	0.483	0.484	0.705	0.737
Adj. R^2	0.858	0.855	0.480	0.471	0.284	0.272	0.591	0.628
First test day FE	Y	Y	Y	Y	Y	Y	Y	Y
Morning FE	Y	Y	Y	Y	Y	Y	Y	Y
Participant FE	Y	Y	Y	Y	Y	Y	Y	Y
	Panel C: Executive function							
	Multitasking Test		One-Touch Stocking of Cambridge		Stop Signaling Task		Spatial Working Memory	
High CO_2	-11.037	-9.650	0.450	0.350	6.239	6.437	-0.400	-1,000
	(8.516)	(13.517)	(0.233)	(0.239)	(4.140)	(4.073)	(0.677)	(0.845)
	[0.715]	[0.798]	[0.311]	[0.593]	[0.570]	[0.488]	[0.968]	[0.798]
High CO_2 x Morning		-2.775		0.200		-0.396		1.200
		(14.408)		(0.358)		(8.019)		(1.035)
		[0.960]		[0.960]		[0.960]		[0.960]
Observations	80	80	80	80	80	80	80	80
R^2	0.842	0.842	0.712	0.713	0.564	0.564	0.820	0.823
Adj. R^2	0.781	0.778	0.601	0.596	0.396	0.385	0.750	0.751
First test day FE	Y	Y	Y	Y	Y	Y	Y	Y
Morning FE	Y	Y	Y	Y	Y	Y	Y	Y
Participant FE	Y	Y	Y	Y	Y	Y	Y	Y

Note: The table shows the results of the parametric regression analysis as presented in section 2.5 with regards to the cognitive performance of participants in the Cambridge Neuropsychological Test Automated Battery (CANTAB) tests. For each outcome variable, we show two columns, with the first column containing the estimated treatment parameter δ based on Equation 1, and column 2 showing δ_1 and δ_2 based on Equation 2. Bootstrapped standard errors based on wild bootstrap clusters with 1,000 replications are shown in (parentheses). Significance levels before multiple hypothesis testing are indicated as $***p < 0.001$; $**p < 0.01$; $*p < 0.05$. In addition, p-values resulting from multiple hypotheses testing, based on the method by Hommel (32) are in [brackets]. Fixed effects (FE) on whether participants conduct the tests on their first test day independent of the CO_2 condition, if they conduct the test in the morning after 30 minutes of exposure, and participant fixed effect has been added. Outcome variables: The CANTAB tests are measured as follow: Reaction Time Tasks as median movement time (in ms), Motor Screening Task as mean latency time (in ms), Delayed Matching to Sample as percentage of correct choices (in %), Paired Associate Learning as number of correct choices on first attempt, Multitasking test as median of multitasking cost (in ms), One-Touch Stocking of Cambridge as problems solved on first choice, Stop Signal Task as stop signal reaction time (in ms), and Spatial Working Memory as number of incorrect revisions. See section 2.3 for a detailed description of the outcome variables.

For the second model specification, including the interaction term with the time of the day (*Morning* dummy), we find a similar effect pattern. Only the Paired Associate Learning task measuring memory is affected at the 1% significance level in the High-CO_2 condition, indicating that participants made, on average, one and a half more correct choices when they were asked to correctly memorize the presented figure. There is also a significant time effect at a 5% level, which offsets the positive effect of being exposed to a 3,000 ppm CO_2 concentration. The results suggest that participants who were exposed for 30 minutes to the higher CO_2 concentration made on average 2 times fewer correct choices, compared to when they were exposed for 330 minutes to 3,000 ppm CO_2. However, these coefficients are also statistically insignificant once corrected for multiple hypothesis testing ($p = 0.056$ and $p = 0.184$, respectively). Overall, we document that statistically significant effects disappear once we correct for multiple hypotheses testing. The results in Table 2 do not provide robust evidence that exposure to a CO_2 concentration of 3,000 ppm (compared to 900 ppm) affects cognitive performance, at least as measured by the CANTAB tests.

3.3 Economic decision-making

We estimated the effects of CO_2 on risk and time preferences, using structural maximum likelihood estimations, similar to previous studies (21; 33; 34). The estimated results for economic decision-making are shown in Table 3. Both the parameter of relative risk aversion α and the monthly discount rate ρ are jointly estimated as linear functions of the treatment dummy while controlling for sex, time of day, and test day. Larger levels of α and ρ indicate higher levels of risk aversion and time discounting, respectively. The results show that neither risk nor time preferences are significantly affected by the higher CO_2 concentrations. To control for decision noise, we included a Fechner error and tremble error in the structural estimations. The tremble error captures random decision-making among individuals answering the multiple price lists, and the Fechner error captures errors in evaluating the expected utility of lotteries. Both errors are not significantly affected by the levels of CO_2 exposure. Summing up, we conclude that we cannot detect any effect of exposure to CO_2 levels of 3,000 ppm on individuals' risk and time preferences and decision errors.

Table 3: Elevated indoor CO_2 and economic decision-making

	Risk aversion α	Discounting ρ	Fechner error μ	Tremble error κ
High CO_2	0.000	0.005	0.116	-0.090
	(0.050)	(0.015)	(0.12)	(0.063)
	[0.998]	[0.998]	[0.984]	[0.616]
High CO_2 x Morning	0.022	0.009	-0.093	-0.030
	(0.061)	(0.010)	(0.144)	(0.093)
	[0.752]	[0.752]	[0.752]	[0.752]
Observations	4800	4800	4800	4800
Log likelihood	-1910	-1910	-1910	-1910
First test day FE	Y	Y	Y	Y
Morning FE	Y	Y	Y	Y
Sex FE	Y	Y	Y	Y

Note: Standard errors clustered at the subject level are in (parentheses). Significance levels before multiple hypothesis testing are indicated as $***p < 0.001$; $**p < 0.01$; $*p < 0.05$. In addition, p-values resulting from multiple hypotheses testing based on the method by Hommel (32) are in [brackets]. The maximum likelihood estimation includes controls for the time of the day when the multiple price lists were conducted including a morning fixed effect (FE), because participants answered the multiple price lists twice within a test day, once after 120 minutes and once after 420 minutes of being exposed to the corresponding CO_2 condition. Additionally, the estimation controls with fixed effects (FE) for the sex of the participant and whether it was the first test day for the participant.

3.4 Physiological responses

As a next step, we evaluated the physiological responses to exposure to CO_2 concentration of 3,000 ppm versus 900 ppm . We present boxplots in Figure 2, providing the unconditional distribution of the physiological parameters in each testing condition, and the results of the regression analysis in Table 4. Similar to the regression analysis conducted for the CANTAB

tests, we first examine the effect of elevated CO_2 concentration on the hourly average level of the corresponding physiological parameter, controlling for the hour of the test day, adding fixed effects for each participant, and a dummy measuring the first test day of the participant. Then, we conducted a second analysis based on Equation 2 in which we interacted the CO_2 dummy coefficient with the dummy variables of each hour that participants were in the testing condition, to examine a dose-response relationship.

Table 4: Elevated indoor CO_2 and physiological response

	Blood CO_2		Heart rate		Respiration rate		Systole blood pressure		Diastole blood pressure		Physical activity		Oxygen consumption	
High CO_2	-0.628	-0.379	0.637	1.151	-0.007	0.002	1.071	1.382	1.482	0.498	12.379	3.284	6.087	9.935
	(0.586)	(0.719)	(2.226)	(2.548)	(0.011)	(0.012)	(1.552)	(2.519)	(1.017)	(1.306)	(7.925)	(9.805)	(7.167)	(8.700)
	[0.647]	[0.817]	[0.677]	[0.817]	[0.677]	[0.817]	[0.677]	[0.817]	[0.530]	[0.817]	[0.424]	[0.817]	[0.677]	[0.817]
High CO_2 x Hour 2		-0.189		-0.486		-0.013		-2.653		0.299		21.363*		3.404
		(0.288)		(1.117)		(0.008)		(3.143)		(1.672)		(9.151)		(6.272)
		[0.864]		[0.864]		[0.752]		[0.864]		[0.864]		[0.287]		[0.864]
High CO_2 x Hour 3		-0.442		-1.898		-0.007		3.486		-0.187		8.884		-0.6814
		(0.374)		(1.958)		(0.010)		(2.953)		(1.621)		(11.058)		(9.395)
		[0.824]		[0.824]		[0.824]		[0.733]		[0.895]		[0.895]		[0.895]
High CO_2 x Hour 4		-0.466		-1.062		-0.003		-1.947		0.473		-4.414		-5.584
		(0.395)		(2.248)		(0.009)		(2.903)		(1.841)		(9.473)		(7.622)
		[0.839]		[0.839]		[0.839]		[0.839]		[0.839]		[0.839]		[0.839]
High CO_2 x Hour 5		-0.365		0.523		-0.008		-0.255		1.130		18.072		-5.404
		(0.436)		(2.019)		(0.006)		(3.239)		(1.671)		(15.324)		(8.468)
		[0.925]		[0.925]		[0.856]		[0.925]		[0.925]		[0.925]		[0.925]
High CO_2 x Hour 6		-0.469		0.611		-0.008		-1.511		2.317		21.034		-3.222
		(0.371)		(2.423)		(0.006)		(2.766)		(2.616)		(11.179)		(8.947)
		[0.750]		[0.750]		[0.619]		[0.750]		[0.750]		[0.426]		[0.750]
High CO_2 x Hour 7		-0.105		-1.464		-0.017		4.632		3.759		17.817		-10.479
		(0.423)		(2.263)		(0.009)		(3.294)		(2.086)		(14.248)		(11.642)
		[0.845]		[0.845]		[0.262]		[0.621]		[0.419]		[0.713]		[0.844]
High CO_2 x Hour 8		0.052		-0.331		-0.016		-4.491		-0.012		-9.999		-2.684
		(0.592)		(2.260)		(0.009)		(3.740)		(3.282)		(14.788)		(10.398)
		[0.997]		[0.997]		[0.425]		[0.997]		[0.997]		[0.997]		[0.997]
Observations	311	311	288	288	288	288	295	295	295	295	312	312	288	288
R^2	0.745	0.746	0.816	0.817	0.759	0.762	0.657	0.670	0.603	0.609	0.456	0.471	0.844	0.845
Adj. R^2	0.720	0.714	0.796	0.792	0.732	0.729	0.620	0.626	0.561	0.556	0.402	0.404	0.828	0.825
First test day FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Hour into test day FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Participant FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

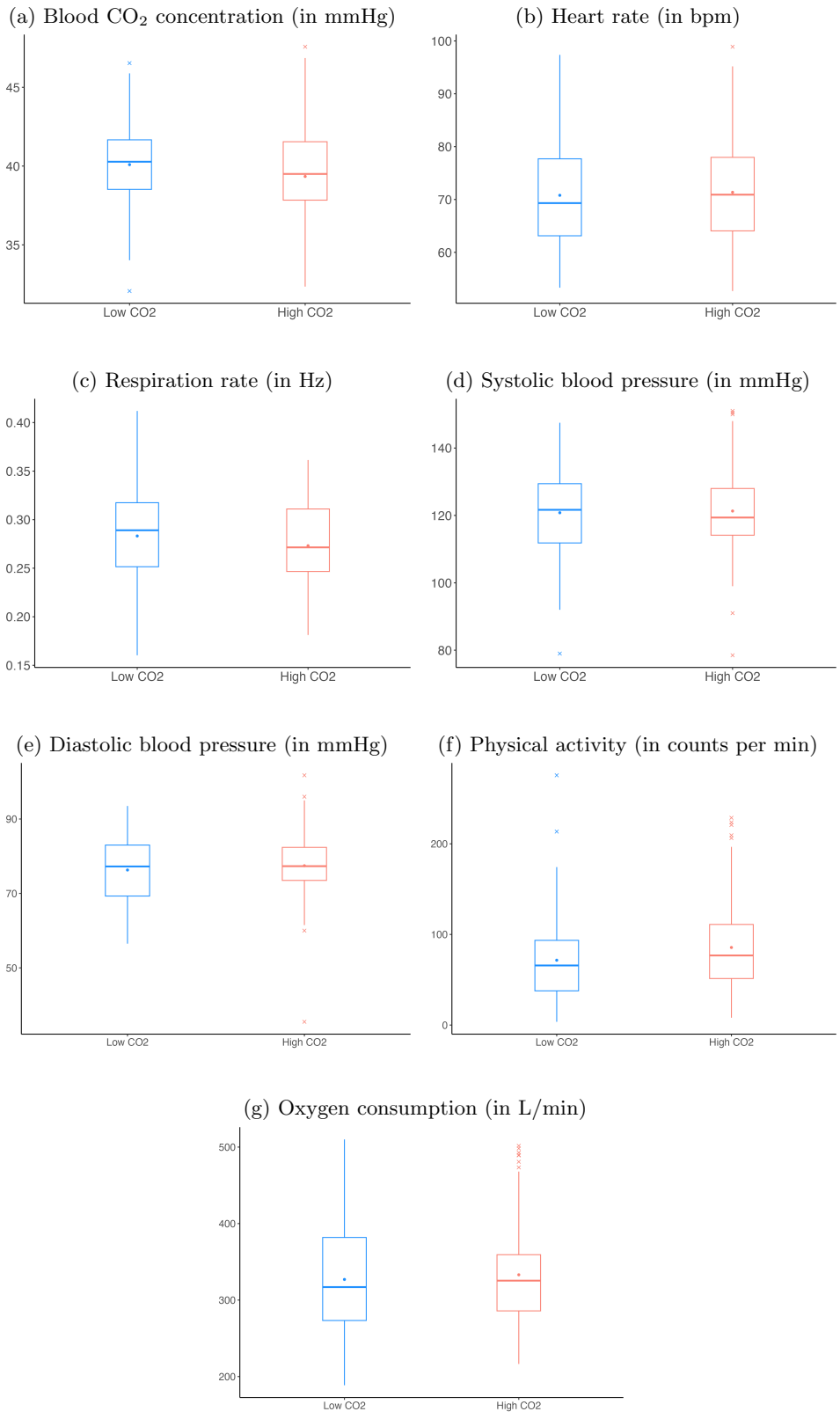
Note: The table shows the results of the parametric regression analysis as presented in section 2.5 with regards to the physiological response. For each outcome variable, we show two columns, with the first column containing the estimated treatment parameter δ based on Equation 1, and column 2 showing δ^1 and δ^2 based on Equation 2. Bootstrapped standard errors based on wild bootstrap clusters with 1,000 replications are shown in (parentheses). Significance levels before multiple hypohtesis testing are indicated as $***p < 0.001$; $**p < 0.01$; $*p < 0.05$. In addition, p-values resulting from multiple hypotheses testing based on the method by Hommel (32) are in [brackets]. The dependent variable for each physiological parameter is aggregated on an hourly average. Fixed effects (FE) on whether participants were in their first test day independent of the CO_2 condition, the hour into exposure, and participant fixed effect has been added. See section 2.3 for a detailed description of the outcome variables.

The boxplots in Figure 2 indicate quite similar distributions between the two testing conditions. This is confirmed by the regression analysis in Table 4. We do not find any significant difference in participants' physiological responses to Low- versus High-CO_2 conditions for most of the outcomes, in both model specifications. Solely for physical activity level, the regression reveals a statistically significant increase in physical activity after two hours of exposure to 3,000 ppm CO_2 concentration ($p < 0.05$). However, conducting multiple hypothesis testing, this effect becomes insignificant ($p = 0.287$).

3.5 Physiological response during cognition tests

Additionally, we examined the physiological response during the time the individual CANTAB Cognition tests were conducted. We used a similar regression model as described in Equation 1 and Equation 2, including an interaction effect of CO_2 and the *Morning* dummy, equaling 1 if the CANTAB test was answered in the morning, 30 minutes into exposure, and zero if it was

Figure 2: Physiological response between CO_2 conditions



Note: In the boxplot diagrams, the thick line in the middle is the median and the point is the average value of each corresponding outcome. The upper and lower edges of the box are the upper and lower quartiles. Values that are more than 1.5 times the interquartile range away from the box are considered to be outliers and shown as crosses (x). The whiskers that extend from the box show the minimum and maximum of the remaining, non-outlier values.

answered after 330 minutes of exposure. The dependent variables in these regressions were the average blood CO_2 concentration, average heart rate and average respiration rate during the time of the individual CANTAB tests. These physiological parameters were chosen because the human body is able to rapidly change its heart rate and respiration rate, which also impacts the blood CO_2 concentration. The individual CANTAB tests took between 1 minute for the Motor Screening Task and 12 minutes for the One-Touch Stockings of Cambridge Task to be conducted. Thus, changes in these outcomes within the short time the individual CANTAB tests taken can be expected. In addition to a fixed effect for the first test day of the participant and a fixed effect for the participant, a fixed effect for the specific CANTAB test was also included. This approach controls for variation between testing time, participant, and individual CANTAB tests.

Table 5 shows the results of this analysis. We observe that, similar to the previous analysis, elevated CO_2 levels do not trigger any physiological response, even if cognitive load is imposed through the cognition tests. However, for the model including the interaction effect, an elevated CO_2 concentration of 3,000 ppm during the morning session of the CANTAB test (after 30 minutes of exposure) is significantly associated with a higher respiration rate ($p < 0.05$), as compared to the afternoon session (after 330 minutes of exposure). This significance remains after conducting multiple hypothesis testing ($p = 0.038$ for the adjusted p-value). However, the magnitude of the effect is quite small: The coefficient indicates an increase in the respiration rate of 0.017 Hz, which is approximately one additional breath per minute, which can be considered a very small effect.

Table 5: Physiological response during cognition tests

	Blood CO_2		Heart rate		Respiration rate	
	(1)	(2)	(3)	(4)	(5)	(6)
High CO_2	−0.468	−0.596	0.056	−0.398	−0.010	−0.019
	(0.533)	(0.515)	(2.225)	(2.662)	(0.012)	(0.011)
	[0.980]	[0.860]	[0.980]	[0.980]	[0.980]	[0.652]
High CO_2 x Morning		0.256		0.909		0.017∗
		(0.414)		(2.337)		(0.007)
		[0.697]		[0.697]		[0.038]
Observations	619	619	571	571	571	571
R^2	0.744	0.766	0.758	0.809	0.579	0.632
Adj. R^2	0.731	0.752	0.745	0.797	0.556	0.609
First test day FE	Y	Y	Y	Y	Y	Y
Morning FE	Y	Y	Y	Y	Y	Y
Participant FE	Y	Y	Y	Y	Y	Y
CANTAB test FE	Y	Y	Y	Y	Y	Y

Note: The table shows the results of the regression analysis with the dependent variable as the average level of blood CO_2 concentration, heart rate, and respiration rate during the time each individual CANTAB cognition test has been conducted. Similar to the regression for the CANTAB test as shown in Table 2 and described in Model 2, next to the treatment dummy High-CO_2 for being 1 if participants were exposed to 3,000 ppm CO_2 concentration, an interaction effect as been added. This effect interact the treatment dummy High-CO_2 with the *Morning* dummy which is equals 1 if the CANTAB test has been conducted after 30 minutes of exposure, compared to zero if the CANTAB test has been conducted 330 minutes into exposure condition. Fixed effects (FE) on whether participants conduct the tests on their first test day independent of the CO_2 condition, if they conduct the test in the morning after 30 minutes of exposure, participant fixed effect, and individual CANTAB test fixed effect has been added. Bootstrapped standard errors based on wild bootstrap clusters with 1,000 replications are shown in (parentheses). Significance levels before multiple hypothesis testing are indicated as $***p < 0.001$; $**p < 0.01$; $*p < 0.05$. In addition, p-values resulting from multiple hypotheses testing based on the method by Hommel (32) are in [brackets].

4 Discussion

Magnitude and significance: Overall, the results show that there is no effect of CO_2 concentrations of 3,000 ppm (compared to 900 ppm) on cognitive performance and physiological outcomes. Although this study uses a similar sample size as compared to previous studies (see Supplementary Table S4 for an overview of previous studies), the question remains whether the estimated effects are indeed zero or whether the sample size of 20 participants is simply too small to estimate the effects precisely enough to reject the null hypothesis. We approach this question from two angles: First, we calculated relative effects where we express the effects in terms of changes in standard deviations of the underlying distribution of the outcome variable. This approach provides insights on the magnitude of the effects, i.e., whether the estimated effects are meaningful, independently of statistical significance. For instance, if the relative effects are very small, i.e., the High-CO_2 condition hardly changes the outcome variable, statistical significance is less relevant because the magnitude of the effects would be negligible. Second, we ran a power analysis to calculate the required sample size to examine how many subjects we

actually would need in order to be able to estimate statistically significant effects in case the CO_2 concentration truly affects the outcome variable. The results of both analyses are provided in the Supplementary Table S3.

For the clear majority of outcome variables, the relative effect of elevated CO_2 levels is weak to very weak, with a change of 0.2 or less standard deviation in the outcome variable. This suggests that the magnitude of most of the parameters is negligible, even if they would have statistical significance. Only two parameters seem to be meaningful with a relative effect size of 0.41 and 0.79 of a standard deviation for the CANTAB's Delayed Matching to Sample task and tremble error for time preferences, respectively. However, conducting a power calculation based on the calculated effect sizes shows that a suggested sample size of 22 participants for the CANTAB test and 13 participants for the economic decision-making test is very close to and below our actual sample of 20 participants, respectively. This indicates that our inference tests are reliable and do allow conclusions on the statistical significance of this meaningful effect. Given that we find only a few significant effects based on the bootstrapped (analytical) standard errors and no statistical significance based on the multiple hypotheses tests, we conclude that the effects are indeed statistically insignificant and hence zero.

Link to previous literature: This study is the first, to our best knowledge, that exposed office workers uninterrupted for a full day (8 hours) to elevated CO_2 concentrations in a validated respiration chamber. The absence of significant results in our analysis shows that CO_2 does not affect basic cognitive domains in terms of attention, psychomotor control, executive function, and memory, measured with the Cambridge Neuropsychological Test Automated Battery, a commonly used and validated cognition test (25). Supplementary Table S4 gives an overview of the related studies with their setup and findings.

Our results contrast with three studies measuring complex decision-making using the Strategic Management Simulator, documenting a negative effect of CO_2 concentrations on cognitive decision making (11; 12; 16). However, one such study, which included "astronaut-like" subjects, found that the negative effect was either mitigated or even reversed at higher CO_2 concentrations (16). A fourth study could not confirm any effect of CO_2 on complex decision-making using the Strategic Management Simulator test (17).

Among this previous work, the study by Allen and co-authors (11) comes closest to our setup, focusing on office workers who are exposed to elevated CO_2 concentrations for 8 hours, while measuring the concentration of volatile organic compounds to ensure a low level of these air pollutants. However, the study was conducted in a common office room and no physiological

parameters were measured; factors that could explain the decline in complex decision-making abilities. In addition, the study used the Strategic Management Simulator for cognitive ability. Thus, our analysis extends the findings showing that for common cognitive domains, CO_2 has no effect (11).

Our findings that CO_2 concentration has no significantly negative effect on human cognition is in agreement with a series of other previous studies, which also did not find a negative effect of elevated CO_2 concentrations on cognitive performance (13; 14). These two studies used similar, although non-validated climate chambers and exposed their population for only a short period of time (255 and 150 minutes, respectively). Also, volatile organic compound concentration was not measured during the testing to validate whether the concentrations of air pollutants were been successfully reduced.

Unique to this study, we widened the cognition analysis, including multiple price lists from the economic literature to examine the potential effect of CO_2 on economic decision-making in terms of risk-taking and time discounting for monetary payouts. We found no effect on risk and time preferences. Previous literature could suggest a potentially negative effect because lower cognitive abilities lead to more random decision-making when answering multiple price lists (35). However, our results are consistent in the sense that we neither find an effect on cognitive performance measured with the CANTAB test, nor for economic decision-making.

Regarding the physiological response to CO_2 exposure, we did not find any stress response in terms of a higher heart rate or higher blood pressure. Previous studies that measured a variety of physiological parameters found a significantly higher heart rate at 2,700 ppm and 3,000 ppm CO_2, but no difference in heart rate at 5,000 ppm CO_2 compared to a concentration of 500 ppm CO_2 (9; 10; 14). However, while various parameters were measured to record the physiological response to elevated CO_2 concentrations, no multiple hypotheses testing was conducted in these studies to adjust the p-values.

Additionally, this study is the first that specifically examines the physiological stress response during the time the cognition test was administered, to derive a possible performance-induced effect of CO_2. Past studies suggest that the cognitive load, from time pressure or complexity of the task, might play a mediating role in the effect of CO_2, and more generally, indoor air quality on cognition (7; 8; 37). However, we did not find any significant change in heart rate, respiration rate, or blood CO_2 levels during the time of testing, indicating that even during times of higher cognitive load, CO_2 does not seem to trigger any physiological response.

Moreover, we could not confirm the hypothesis that an elevated indoor CO_2 concentration

of 3,000 ppm leads to a respiratory acidosis. Neither the respiration rate nor the blood CO_2 partial pressure were significantly affected, except for a marginally higher respiration rate of one breath per minute during the CANTAB test. Only one prior study found higher blood CO_2 concentrations after four hours of exposure to up to 5,000 ppm CO_2 (22). However, the CO_2 concentration in that study was achieved through a reduced indoor ventilation rate, leading to a 2.2-fold increase in volatile organic compounds concentrations in the room. The authors attribute the elevation in blood CO_2 to the increased CO_2 levels in the room, but did not elaborate on whether such a relationship could be mediated by the other air pollutants in the room which might have an impairing effect on the lungs (23; 38).

The previous literature assumes that air pollutants could cause changes in the breathing pattern, which in turn leads to a build-up of CO_2 in the blood due to an insufficient removal through exhalation (23; 39). Moreover, Snow and co-authors (10) argue that the higher heart rate for an exposure level of 2,700 ppm could be indicative of an increase in circulation to maintain CO_2 levels in the blood. However, they documented no effect on respiration rate and emphasized that no blood-gas analysis was conducted to examine this hypothesis. Because we isolated the effect of CO_2 in our study setup, we can rule out that respiratory acidosis is related to CO_2 concentrations of 3,000 ppm and no change in respiration rate was needed to mediate this. However, we did not measure tidal volume which could be affected independently of the respiration rate. Importantly, our study does not aim to examine the claim stated in previous literature that air pollutants beyond CO_2 affect the breathing pattern of individuals (23; 39).

Finally, two studies found an increase in end-tidal CO_2 in exhaled air but were not able to examine the physiological reasoning for this observation (9; 14). Increased exhalation of CO_2 can be a sign of increased cellular CO_2 production or CO_2 build-up in the blood due to increased metabolic rate. To maintain a stable pH-level, the body increases the respiration rate to remove excess CO_2 from the lungs (40). However, we could not find any significant effect on oxygen consumption as a measure of metabolic rate and also no significantly higher physical activity level in the High-CO_2 condition, which could cause a higher energy expenditure and thus higher metabolic rate. Thus, while we did not measure end-tidal CO_2 directly, our analysis cannot confirm that elevated CO_2 levels affect human metabolic rate. Therefore the higher end-tidal CO_2 concentration found in earlier studies might not be caused by an increased cellular CO_2 production from a higher metabolic rate.

Limitations and Future Recommendations: We used a validated respiration chamber and measured volatile organic compounds continuously to ensure that our method of isolating

the effect of CO_2 was successful. However, we simulated only two conditions: High- and Low-CO_2. In reality, CO_2 levels might be much higher. CO_2 levels in primary schools easily exceed 3,000 ppm if no sufficient ventilation is provided and effects of exposure to CO_2 might be non-linear (41; 42).

In addition, the study population was of average age (25 to 46 years old) and had no health-related complaints. Thus, our study does not answer how CO_2 concentrations might affect individuals with respiratory restrictions, like chronic obstructive pulmonary disease and asthma, metabolic syndrome, or mental disorders such as depression and anxiety disorder. Also, future studies could investigate the effect of CO_2 on the health and cognition of children and elderly people.

5 Conclusion

The increased focus on the energy efficiency of buildings as a route to reduce greenhouse gas emissions may lead to a situation where a trade-off needs to be made between high ventilation rates (leading to enhanced indoor air quality) and lowering energy consumption. An alternative way of creating a healthy indoor environment is the implementation of air filtering systems that clean organic compounds and fine dust from the air (43; 44). However, these air filters generally cannot remove CO_2, given limited external air intake. Therefore, it is important to understand whether CO_2, which is often used as a proxy for indoor air quality, is an air pollutant itself or whether elevated levels of CO_2 can be accepted, such that air filtering systems could substitute an increased ventilation rate to maintain an appropriate indoor air quality for occupants – be it employees or residents.

This study exploits a validated ventilation chamber in a cross-over experimental design to assess the impact of elevated CO_2 levels on cognitive performance and physiological response. The analysis reveals that a CO_2 concentration of 3,000 ppm compared to 900 ppm does not trigger any significant cognitive decline or physiological response. As such, for healthy individuals, no negative effect of a CO_2 concentration of 3,000 ppm compared to 900 ppm on cognition and economic decision-making can be expected. These findings contrast with some existing studies, which claim that elevated CO_2 levels have short-term implications for health and cognition, and suggest that somewhat lower ventilation do not necessarily harm human performance – at least in the short term (11). Of course, there might be other considerations for elevated ventilation rates, such as the prevention of disease, or the reduction of volatile organic compounds. How-

ever, such pollution could also be addressed in alternative, potentially cheaper and more efficient ways.

6 References

[1] MacNaughton P, Pegues J, Satish U, Santanam S, Spengler J, Allen J. Economic, environmental and health implications of enhanced ventilation in office buildings. International journal of environmental research and public health. 2015;12(11):14709-22.

[2] Eichholtz P, Kok N, Quigley JM. Doing well by doing good? Green office buildings. American Economic Review. 2010;100:2492-509.

[3] Morawska L, Allen J, Bahnfleth W, Bennett B, Bluyssen PM, Boerstra A, et al. Mandating indoor air quality for public buildings. Science. 2024;383(6690):1418-20.

[4] Santos HR, Leal VM. Energy vs. ventilation rate in buildings: A comprehensive scenario-based assessment in the European context. Energy and Buildings. 2012;54:111-21.

[5] Das P, Chalabi Z, Jones B, Milner J, Shrubsole C, Davies M, et al. Multi-objective methods for determining optimal ventilation rates in dwellings. Building and Environment. 2013;66:72-81.

[6] Rim D, Schiavon S, Nazaroff WW. Energy and cost associated with ventilating office buildings in a tropical climate. PloS one. 2015;10(3):e0122310.

[7] Du B, Tandoc MC, Mack ML, Siegel JA. Indoor CO2 concentrations and cognitive function: A critical review. Indoor air. 2020;30:1067-82.

[8] Wargocki P, Wyon DP. Ten questions concerning thermal and indoor air quality effects on the performance of office work and schoolwork. Building and Environment. 2017;112:359-66. Available from: `https://www.sciencedirect.com/science/article/pii/S0360132316304449`.

[9] Zhang X, Wargocki P, Lian Z. Physiological responses during exposure to carbon dioxide and bioeffluents at levels typically occurring indoors. Indoor Air. 2017;27:65-77.

[10] Snow S, Boyson AS, Paas KHW, Gough H, King MF, Barlow J, et al. Exploring the physiological, neurophysiological and cognitive performance effects of elevated carbon dioxide concentrations indoors. Building and Environment. 2019;156:243-52.

[11] Allen JG, MacNaughton P, Satish U, Santanam S, Vallarino J, Spengler JD. 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. Environmental health perspectives. 2016;124:805-12.

[12] Satish U, Mendell MJ, Shekhar K, Hotchi T, Sullivan D, Streufert S, et al. Is CO2 an indoor pollutant? Direct effects of low-to-moderate CO2 concentrations on human decision-making performance. Environmental health perspectives. 2012;120:1671-7.

[13] Zhang X, Wargocki P, Lian Z, Thyregod C. Effects of exposure to carbon dioxide and bioeffluents on perceived air quality, self-assessed acute health symptoms, and cognitive performance. Indoor air. 2017;27:47-64.

[14] Zhang X, Wargocki P, Lian Z. Human responses to carbon dioxide, a follow-up study at recommended exposure limits in non-industrial environments. Building and Environment. 2016;100:162-71. Available from: https://www.sciencedirect.com/science/article/pii/S0360132316300580.

[15] Liu W, Zhong W, Wargocki P. Performance, acute health symptoms and physiological responses during exposure to high air temperature and carbon dioxide concentration. Building and Environment. 2017;114:96-105. Available from: https://www.sciencedirect.com/science/article/pii/S0360132316305121.

[16] Scully RR, Basner M, Nasrini J, wing Lam C, Hermosillo E, Gur RC, et al. Effects of acute exposures to carbon dioxide on decision making and cognition in astronaut-like subjects. npj Microgravity. 2019;5:1-15.

[17] Rodeheffer CD, Chabal S, Clarke JM, Fothergill DM. Acute exposure to low-to-moderate carbon dioxide levels and submariner decision making. Aerospace medicine and human performance. 2018;89:520-5.

[18] Schoffelen PF, Westerterp KR, Saris WH, Ten Hoor F. A dual-respiration chamber system with automated calibration. Journal of applied physiology. 1997;83(6):2064-72.

[19] Coller M, Williams MB. Eliciting Individual Discount Rates. Experimental Economics. 1999;2:107-27.

[20] Holt CA, Laury SK. Risk Aversion and Incentive Effects. American Economic Review. 2002 12;92:1644-55. Available from: https://www.aeaweb.org/articles?id=10.1257/000282802762024700.

[21] Meissner T, Gassmann X, Faure C, Schleich J. Individual characteristics associated with risk and time preferences: A multi country representative survey. Journal of Risk and Uncertainty. 2023;66(1):77-107.

[22] Vehviläinen T, Lindholm H, Rintamäki H, Pääkkönen R, Hirvonen A, Niemi O, et al. High indoor CO2 concentrations in an office environment increases the transcutaneous CO2 level and sleepiness during cognitive work. Journal of occupational and environmental hygiene. 2016;13:19-29.

[23] Shriram S, Ramamurthy K, Ramakrishnan S. Effect of occupant-induced indoor CO2 concentration and bioeffluents on human physiology using a spirometric test. Building and Environment. 2019;149:58-67.

[24] Schweiker M, Wagner A. The effect of occupancy on perceived control, neutral temperature, and behavioral patterns. Energy and Buildings. 2016;117:246-59.

[25] Smith PJ, Need AC, Cirulli ET, Chiba-Falek O, Attix DK. A comparison of the Cambridge Automated Neuropsychological Test Battery (CANTAB) with "traditional" neuropsychological testing instruments. Journal of Clinical and Experimental Neuropsychology. 2013 3;35:319-28. Available from: https://doi.org/10.1080/13803395.2013.771618.

[26] Ltd CC. CANTAB Cognitive assessment software;. Available from: https://www.cambridgecognition.com/cantab/cognitive-tests/.

[27] Ceylan B, Khorshid L, Ülkü Yapucu Güneş, Zaybak A. Evaluation of oxygen saturation values in different body positions in healthy individuals. Journal of Clinical Nursing. 2016 4;25:1095-100. Available from: https://doi.org/10.1111/jocn.13189.

[28] Messina Z, Patrick H. Partial pressure of carbon dioxide. StatPearls Publishing; 2022. Available from: https://pubmed.ncbi.nlm.nih.gov/31869112/.

[29] Gilgen-Ammann R, Schweizer T, Wyss T. RR interval signal quality of a heart rate monitor and an ECG Holter at rest and during exercise. European Journal of Applied Physiology. 2019;119:1525-32. Available from: https://doi.org/10.1007/s00421-019-04142-5.

[30] Tarvainen MP, Niskanen JP, Lipponen JA, Ranta-aho PO, Karjalainen PA. Kubios HRV – Heart rate variability analysis software. Computer Methods and Programs in Biomedicine. 2014;113:210-20. Available from: https://www.sciencedirect.com/science/article/pii/S0169260713002599.

[31] Cameron AC, Gelbach JB, Miller DL. Bootstrap-Based Improvements for Inference with Clustered Errors. The Review of Economics and Statistics. 2008 8;90:414-27. Available from: https://doi.org/10.1162/rest.90.3.414.

[32] Hommel G. A stagewise rejective multiple test procedure based on a modified Bonferroni test. Biometrika. 1988 6;75:383-6. Available from: https://doi.org/10.1093/biomet/75.2.383.

[33] Andersen S, Harrison GW, Lau MI, Rutström EE. Eliciting Risk and Time Preferences. Econometrica. 2008 5;76:583-618. Available from: https://doi.org/10.1111/j.1468-0262.2008.00848.x.

[34] Harrison GW. Risk aversion in experiments. Emerald Group Publishing; 2008.

[35] Andersson O, Holm HJ, Tyran JR, Wengström E. Robust inference in risk elicitation tasks. Journal of Risk and Uncertainty. 2020;61:195-209. Available from: https://doi.org/10.1007/s11166-020-09341-6.

[36] Huizenga C, Abbaszadeh S, Zagreus L, Arens EA. Air quality and thermal comfort in office buildings: Results of a large indoor environmental quality survey. UC Berkeley: Center for the Built Environment. 2006.

[37] Künn S, Palacios J, Pestel N. Indoor Air Quality and Strategic Decision Making. Management Science. 2023 1. Available from: https://doi.org/10.1287/mnsc.2022.4643.

[38] Danuser B. Candidate Physiological Measures of Annoyance from Airborne Chemicals. Chemical Senses. 2001 4;26:333-7. Available from: https://doi.org/10.1093/chemse/26.3.333.

[39] Bakó-Biró ZS, Wargocki P, Wyon DP, Fanger PO. Poor indoor air quality slows down metabolic rate of office workers. Proceedings of indoor air. 2005;1:76-80.

[40] Negro CAD, Funk GD, Feldman JL. Breathing matters. Nature Reviews Neuroscience. 2018;19:351-67. Available from: https://doi.org/10.1038/s41583-018-0003-6.

[41] Eichholtz P, Kok N, Sun X. The effect of post-COVID-19 ventilation measures on indoor air quality in primary schools. PNAS nexus. 2024;3(1):pgad429.

[42] Palacios J, Eichholtz P, Kok N, Duran N. Indoor Air Quality and Learning: Evidence from a Large Field Study in Primary Schools; 2022.

[43] Liu G, Xiao M, Zhang X, Gal C, Chen X, Liu L, et al. A review of air filtration technologies for sustainable and healthy building ventilation. Sustainable cities and society. 2017;32:375-96.

[44] Cooper E, Wang Y, Stamp S, Burman E, Mumovic D. Use of portable air purifiers in homes: Operating behaviour, effect on indoor PM2.5 and perceived indoor air quality. Building and Environment. 2021;191:107621.

7 Appendix

Table S1: Multiple price lists for risk preferences

		Option A		Option B		
Multiple price list	Choice	Coin shows heads	Coin shows tails	Coin shows heads	Coin shows tails	Implied CRRA
1.1	1	6.00 €	10.00 €	1.00 €	12.00 €	1.79
	2	6.00 €	10.00 €	1.00 €	14.00 €	1.17
	3	6.00 €	10.00 €	1.00 €	16.00 €	0.87
	4	6.00 €	10.00 €	1.00 €	18.00 €	0.69
	5	6.00 €	10.00 €	1.00 €	20.00 €	0.56
	6	6.00 €	10.00 €	1.00 €	22.00 €	0.47
	7	6.00 €	10.00 €	1.00 €	24.00 €	0.40
	8	6.00 €	10.00 €	1.00 €	28.00 €	0.29
	9	6.00 €	10.00 €	1.00 €	34.00 €	0.19
	10	6.00 €	10.00 €	1.00 €	44.00 €	0.09
1.2	11	0.40 €	8.00 €	5.00 €	9.00 €	NA
	12	0.40 €	10.00 €	5.00 €	9.00 €	2.243
	13	0.40 €	11.00 €	5.00 €	9.00 €	1.61
	14	0.40 €	12.00 €	5.00 €	9.00 €	1.292
	15	0.40 €	13.00 €	5.00 €	9.00 €	1.09
	16	0.40 €	14.00 €	5.00 €	9.00 €	0.948
	17	0.40 €	15.00 €	5.00 €	9.00 €	0.841
	18	0.40 €	19.00 €	5.00 €	9.00 €	0.584
	19	0.40 €	27.00 €	5.00 €	9.00 €	0.358
	20	0.40 €	43.00 €	5.00 €	9.00 €	0.184
2.1	21	30.00 €	30.00 €	30.00 €	1.00 €	NA
	22	25.00 €	25.00 €	30.00 €	1.00 €	3.8
	23	20.00 €	20.00 €	30.00 €	1.00 €	1.7
	24	17.00 €	17.00 €	30.00 €	1.00 €	1.1
	25	16.00 €	16.00 €	30.00 €	1.00 €	1.06
	26	15.00 €	15.00 €	30.00 €	1.00 €	0.94
	27	12.00 €	12.00 €	30.00 €	1.00 €	0.63
	28	10.00 €	10.00 €	30.00 €	1.00 €	0.45
	29	5.00 €	5.00 €	30.00 €	1.00 €	-0.06
	30	1.00 €	1.00 €	30.00 €	1.00 €	NA
2.2	31	14.00 €	17.00 €	17.00 €	1.00 €	NA
	32	14.00 €	17.00 €	20.00 €	1.00 €	2.8
	33	14.00 €	17.00 €	25.00 €	1.00 €	1.4
	34	14.00 €	17.00 €	28.00 €	1.00 €	1.1
	35	14.00 €	17.00 €	29.00 €	1.00 €	1.06
	36	14.00 €	17.00 €	30.00 €	2.00 €	0.93
	37	14.00 €	17.00 €	30.00 €	3.00 €	0.87
	38	14.00 €	17.00 €	32.00 €	8.00 €	0.21
	39	14.00 €	17.00 €	32.00 €	10.00 €	-1.04
	40	14.00 €	17.00 €	32.00 €	14.00 €	NA

Note: The table shows a total of four multiple price lists, MPL1.1, MPL1.2, MPL2.1 and MPL2.2, each containing a total of ten choices between two options labeled neutrally as A and B to elicit risk preferences. Participants repeatedly chose between choices with differing levels of risk. Utility over monetary gains is modeled assuming constant relative risk aversion (CRRA), as expressed in Equation 3, described in Section 2.5.

Table S2: Multiple price lists for time preferences

Multiple price list	Choice	Option A: Today	Option B: In one month	Yearly discount factor
3.1	1	18.20 €	18.00 €	1.14
	2	18.00 €	18.00 €	1.00
	3	17.80 €	18.00 €	0.87
	4	17.30 €	18.00 €	0.62
	5	16.80 €	18.00 €	0.44
	6	16.00 €	18.00 €	0.24
	7	14.00 €	18.00 €	0.05
	8	12.00 €	18.00 €	0.01
	9	10.00 €	18.00 €	0.00
	10	8.00 €	18.00 €	0.00
3.2	11	12.00 €	11.80 €	1.22
	12	12.00 €	12.00 €	1.00
	13	12.00 €	12.20 €	0.82
	14	12.00 €	12.50 €	0.61
	15	12.00 €	13.00 €	0.38
	16	12.00 €	14.00 €	0.16
	17	12.00 €	15.00 €	0.07
	18	12.00 €	16.00 €	0.03
	19	12.00 €	18.00 €	0.01
	20	12.00 €	22.00 €	0.00

Note: The table shows a total of two multiple price lists, MPL3.1 and MPL3.2, each containing a total of ten choices between two options labeled neutrally as A and B to elicit time preferences. Participants repeatedly chose between varying monetary payoffs at different points in time. Inter-temporal choices as measure of time preferences are modelled using a simple expected discounted utility model, as expressed in Equation 4, described in Section 2.5.

Table S3: Relative effect and prospective sample size

	Coefficient	SD (Outcome)	Relative effect	Sample size
Panel A: CANTAB cognition tests				
Reaction Time Task	3.588	49.696	0.072	111
Motor Screening Task	0.493	86.484	0.006	1311
Delayed Matching to Sample	3.550	8.712	0.407	22
Paired Associate Learning	0.500	2.974	0.168	49
Multitasking Test	-11.038	73.866	-0.149	55
One-Touch Stocking of Cambridge	0.450	1.598	0.282	30
Stop Signaling Task	6.239	25.333	0.246	34
Spatial Working Memory	-0.400	4.883	-0.082	98
Panel B: Economic decision-making				
Risk aversion	0.010	0.298	0.033	242
Discounting	0.009	0.120	0.068	118
Fechner error	0.053	0.394	0.134	61
Tremble error	-0.092	0.115	-0.793	13
Panel C: Physiological parameters				
Blood CO_2	-0.628	3.016	-0.208	40
Heart rate	0.637	9.971	0.064	125
Respiration rate	-0.007	0.046	-0.154	53
Systolic blood pressure	1.071	12.122	0.088	92
Diastolic blood pressure	1.482	8.334	0.178	48
Physical activity	12.379	46.851	0.264	32
Oxygen consumption	6.087	67.367	0.090	90

Note: Column 1 shows the estimated coefficient δ based on equation 1. Column 2 contains the standard deviation (SD) of the underlying distribution of the outcome variable in the full sample. Column 3 shows the relative effect as calculated by dividing the estimated effect δ by the standard deviation. Last, column 4 presents the required sample size resulting from a power analysis. For economic decision-making in Panel B, the standard deviation was calculated as the standard error of the estimated coefficient on the treatment dummy, multiplied by the square root of the number of participants. The power analysis is conducted based on a linear multiple fixed effect regression model, two-tailed, with an alpha error rate of 0.05, a power of 0.8 and 1 predictor.

Table S4: Comparison with related papers

Panel A: Studies with a combination of cognitive and physiological parameters					
	This paper	(9; 13)	(10)	(14)	(15)
Testing room	Validated respiration chamber	Climate chamber	Office room	Climate chamber	Climate chamber
No. of subjects in room	1	In groups of 5	1	In groups of 4	In pairs of 2
Method to remove pollutants	High ventilation, pollution filter	Cleaned, baked at 40°C & high ventilation	High ventilation	High ventilation	High ventilation
Volatile organic compounds measured	Yes	No	No	No	No
Exposure levels (in ppm)	900 vs. 3,000	500 vs. 1,000 vs. 3,000	830 vs. 2,700	500 vs. 5,000	380 vs. 3,000 (at 35°C)
Exposure time (in min)	480	255	50	153	180
Time between test days	4 to 6 weeks	No days in between	Unspecified	1 day	No days in between
Study design	Within subject	Within subject	Within subject	Within subject	Within subject
Sample size	20	25	31	10	12
Blinding condition	Single blinded	Single blinded	Single blinded	Single blinded	Unspecified
Exposure order randomized	Yes	Yes	Yes	Yes	Yes
Population	Office workers	Students	Employees or students	Students	Unspecified
Significant physiological response	No	Yes	No	Yes	No
Significant cognitive response	No	No	Yes	No	No
Significant performance-induced response	Yes	Not measured	Not measured	Not measured	Not measured

Panel B: Studies with a focus on decision-making					
	This paper	(16)	(11)	(12)	(17)
Testing room	Validated respiration chamber	Climate chamber	Office room	Office-like chamber	Hypo hyperbaric chamber
No. of subjects in room	1	In groups of 4 to 6	In groups	In groups of 4	In groups of 2 to 4
Method to remove pollutants	High ventilation, pollution filter	Unspecified	High ventilation	High ventilation	Unspecified
Volatile organic compounds measured	Yes	No	Yes	No	No
Exposure levels (in ppm)	900 vs. 3,000	600 vs. 1,200 vs. 2,500 vs. 5,000	550 vs. 945 vs. 1,400	600 vs. 1,000 vs.2,500	600 vs. 2,500 vs. 15,000
Exposure time (in min)	480	175	480	150	125
Time between test days	4 to 6 weeks	1 week	No days in between	1 hour	Not applicable
Study design	Within subject	Within subject	Within subject	Within subject	Between subjects
Sample size	20	22	24	22	36 (12 per group)
Exposure order randomized	Yes	Yes	No	Yes	Yes
Population	Office workers	Astronout-like subjects	Office workers	Unspecified	Male submariners
Significant cognitive response	No	Yes, but inconsistent	Yes	Yes	No