Journal of Environmental Psychology 97 (2024) 102367

Contents lists available at ScienceDirect

Journal of Environmental Psychology

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

Elevated ceiling heights reduce the cognitive performance of higher-education students during exams

Isabella S. Bower [a,*], Jaclyn Broadbent [b], Scott Coussens [a], Peter G. Enticott [c]

[a] Cognitive Ageing and Impairment Neurosciences Laboratory, Behaviour-Brain-Body Research Centre, UniSA Justice and Society, University of South Australia, Adelaide, 5000, Australia
[b] Centre for Assessment and Digital Learning (CRADLE) & School of Psychology, Faculty of Health, Deakin University, 3220, Geelong, Australia
[c] Cognitive Neuroscience Unit, School of Psychology, Faculty of Health, Deakin University, Geelong, 3220, Australia

ARTICLE INFO

Handling Editor: L. McCunn

Keywords:
Assessment and evaluation
Built environment
Cognitive performance
Examination
Higher education
Interior design

ABSTRACT

Examinations are a widely used assessment method in higher education. They are often conducted in large indoor environments that can accommodate high numbers of students to maximize scheduling and cost efficiency. Recent evidence, however, suggests enlarged room scale impacts brain activity that is associated with concentration, which could negatively impact cognitive performance. We analysed data (N = 15,400) from undergraduate students over eight years across three campuses at an Australian tertiary institution. Using a linear mixed model, we compared examination performance across different room scales, while accounting for coursework performance, and other variables. We found student examination performance was reduced in rooms with elevated ceiling heights. These results support the notion that built environment scale influences cognitive performance, and argue against conducting examinations in large scale, high-ceiling rooms.

1. Introduction

For more than 1300 years, examinations have enabled us to investigate knowledge and capabilities and can be pivotal in shaping a student's academic progression (O'Sullivan & Cheng, 2022). When conducting examinations, scalability is paramount, and venues are primarily selected for capacity. In both secondary school and tertiary settings where a subject or course has a sizable student cohort, large scale indoor spaces (auditoriums, indoor recreational spaces etc.) are used to group students under the same conditions. This enables student identity authentication at scale, and maximizes scheduling of facilities and cost efficiency in personnel required to proctor the examination for the institution (Butler-Henderson & Crawford, 2020). In Australia, these venues can range from hiring public spaces (racecourses, showgrounds, town halls etc.) to utilizing on-campus spaces not traditionally associated with teaching such as gymnasiums, auditoriums, and large multipurpose event rooms (for some examples see University descriptions (Deakin University, 2023; The University of Adelaide, 2022; The University of Melbourne, 2023). The outcome of examinations significantly influence a student's ability to progress to higher education, pursue postgraduate studies, or, in the context of law, enter the profession (Sheppard, 1996). While not always compulsory, the often high weighting of examinations necessitates completion, without which, it may not be possible to pass the unit of study successfully (Amigud, Arnedo-Moreno, Daradoumis, & Guerrero-Roldan, 2018). From a pedagogical perspective, a recent review found little empirical evidence to support conducting high-stakes examinations, and the problematic nature of this form of assessment (French, Dickerson, & Mulder, 2023). Thus, cumulatively, examinations can significantly influence career direction and prospects for students, alongside shaping the cohort of the students who are selected for entry to tertiary and postgraduate studies, and in turn, our future leaders in industry, academia, and government. Despite the important role and implications this form of assessment can have, there is no research into how the environment where examinations are conducted might affect student performance.

Increasing research is being conducted to understand the role of the built environment on neurophysiological response (Bower, Tucker, & Enticott, 2019). Currently, research exploring the role of the built environment on assessment and performance has focused on the role of context (Metzger, Boschee, Haugen, & Schnobrich, 1979; Saufley, Otaka, & Bavaresco, 1985; Weir & May 1988), familiarity (Cassaday, Bloomfield, & Hayward, 2002), and ambient environmental quality parameters (Brink, Loomans, Mobach, & Kort, 2021; Hoque & Weil, 2016). However, there is strikingly little research investigating the

* Corresponding author. University of South Australia, GPO Box 2471, Adelaide, South Australia, 5072, Australia.
E-mail address: isabella.bower@unisa.edu.au (I.S. Bower).

https://doi.org/10.1016/j.jenvp.2024.102367
Received 24 January 2024; Received in revised form 23 June 2024; Accepted 24 June 2024
Available online 26 June 2024
0272-4944/

impact of interior design characteristics, with even less empirical work conducted to investigate the role of room scale encompassing both internal area and ceiling height (Meyers-Levy & Zhu, 2007) on cognitive performance. A recent study investigated the impact of room scale on neurophysiological correlates using electroencephalography, autonomic and self-report indices. This work showed that enlarged indoor environment scale impacts markers of brain activity that are associated with high-order cognitive processes, which could adversely affect cognitive performance (Bower et al., 2022; Bower, Hill, & Enticott, 2023). Given the association between the brain signatures identified and their role in cognitive performance, we were interested in testing the impact of enlarged room scale in a real-world setting with a cognitive task that has been shown to induce stress (i.e., undergraduate examination).

In this study we examined whether student performance in exams is influenced by the scale of rooms. Using a naturalistic approach, we obtained historical data from three first-year undergraduate units (subject of study sometimes called courses) from 15,400 student datapoints across an eight-year retrospective period across three campuses at a University in Australia. To conduct a comprehensive analysis of the data, we employed a linear-mixed-model methodology, as detailed in the materials and methods section. This approach incorporated the coursework score as a covariate while concurrently exploring the potential impact of other variables on the observed relationship. Notably, we deliberately omitted the inclusion of internal area or volume in the model due to concerns about collinearity – that is, given the inherent characteristics of the rooms, instances of elevated ceiling height were frequently associated with larger internal areas and volumes, such as in the case of a basketball court.

2. Materials and methods

2.1. Participants

The database used for this study primarily included academic performance scores and examination locations from 15,400 students (11,532 female, 3850 male, 18 other; $M_{age} = 20.2$, $SD_{age} \pm 4.61$) who undertook at least one of three first-year units of study in introductory psychology units during their undergraduate studies at an Australian University during the period 2011–2019. Our rationale was to analyse data prior to the global pandemic as students historically sat examinations on campus. Our sample included three separate units, and exam venue data from three different geographically located campuses (1 metropolitan and 2 regional cities), eight buildings, and 11 rooms. An ethics waiver for collecting and analysing existing data in this study was granted by the University Human Research Ethics Committee. We preregistered the study prior to data analysis (https://doi.org/10.17605/OSF.IO/C9HF5). All data were supplied and analysed in an anonymous format, without access to personal identifying information. The data and code that support the findings of this study are available in Open Science Framework (https://doi.org/10.17605/OSF.IO/GY8M4).

2.2. Data inclusion

To create the database for this study, we obtained and linked two datasets. A numerical integer was applied to the original student identification (ID) number on both datasets to create a new deidentified student ID prior to the data being accessible to the research team. This enabled linking by the research team while maintaining anonymity of the students. Both datasets included the student ID number, unit code, year, and trimester to match the data (see Fig. 1).

2.2.1. Academic performance dataset

The first dataset contained the coursework score, examination score, overall grade, and grading descriptor for each student (N = 18,301). Coursework scores or 'assignments' across the three units involved between 1 and 3 tasks (dependent on unit) which could be completed in

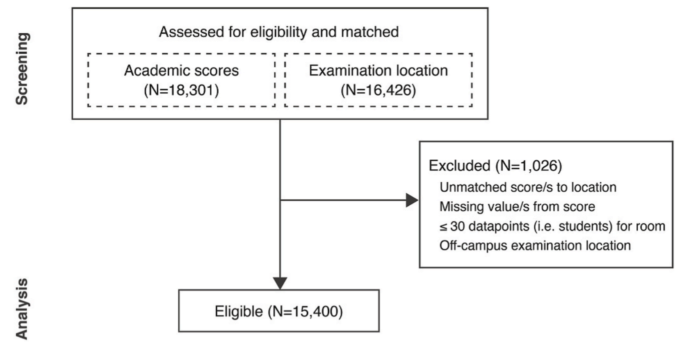

Fig. 1. Data summary. CONSORT diagram (Consolidated Standards of Reporting Trials) of the data inclusion and linking process of this study.

the student's environment of choice (often a 'home' setting). The types of assignments across these units consisted of essays (800-1200 words), laboratory reports (2000 words), short quizzes, and journal reflections (700-1200 words per entry). In contrast, examinations were highly formal supervised assessments with set location and time, ranging from 90 min to 2-h in length. We have summarised the forms of assessment and weightings in Table 1.

To include a datapoint, we required both a coursework and examination score. Datapoints which had missing values for coursework and/or examination score were removed. As variable weightings had been applied to the coursework and examination scores during the course history to determine the overall grade, we created adjusted scores for both as a percentage of 100.

2.2.2. Examination location dataset

The second dataset included the campus location, examination venue code, seat number and the age and gender of each student (N = 16,426). Of the exam venue codes which were identified to have one or more student datasets (N = 267), we required the examination code to be matched to an on-campus venue and include ≥30 datapoints (students) across the three units and over the 8-year period (N = 28). We then sourced the internal room area (m2) and ceiling height (m) for each of the included rooms from the institution's building information modelling system (Archibus Inc.). Examination locations included rooms across three-campuses and eight-buildings. To categorise the three campuses, we referred to the state foundational description for 'metropolitan' and 'regional' which are based on boundaries. However, to also capture the unique population, location, and demographics we included the local government area codes (known as LGA) to provide contextualised census data which can be viewed via the Australian Bureau of Statistics (Australian Bureau of Statistics, 2021). This resulted in ceiling heights between 2.79 and 9.50-m, internal floor area of 38 to 1562-m^2, and estimated internal volume of 115 to 14,831 m^3. We were unable to determine which room and level of the building students assigned for three of the examination location codes. Although this included 15 different rooms across levels 1–3 there was only small variation in the internal area, ceiling height, and volume. To assign values, we averaged each measure of scale. We have visually summarised this data in Fig. 2, and the descriptives for the examination venues are summarised in Table 2.

2.3. Data linking and analysis

To merge the datasets, we used a database management system (Microsoft Access, Version 2307, Build 16.0.1.16626.20170) imported both datasets as tables, created a query, and then matched three field properties (deidentified student ID, year, and unit code). This resulted in a database of N = 15,400 which we exported as a.csv spreadsheet for analysis. Of the datapoints, N = 9075 unique deidentified participant IDs were present (N = 6325 students had two or more entries reflecting

Table 1
Assessment tasks and weightings for units included in this study. Note: Only data from 2018 to 2019 was available and included for Unit 1, and 2015–2019 for unit 2. These weightings and assessment tasks are derived from the published student handbooks from the institution.

Year	Unit 1					Unit 2					Unit 3				
	Coursework (%)				Exam (%)	Coursework (%)				Exam (%)	Coursework (%)				Exam (%)
	Essay	Report	Journal	Quiz		Essay	Report	Journal	Quiz		Essay	Report	Journal	Quiz	
2011											0	0	45	10	45
2012											0	0	45	10	45
2013											0	0	45	10	45
2014											0	0	45	10	45
2015						0	40	0	0	60	0	0	45	10	45
2016						0	40	0	0	60	0	0	45	10	45
2017						0	40	0	0	60	0	0	45	10	45
2018	50	0	0	20	30	0	40	0	0	60	0	0	45	10	45
2019	50	0	0	20	30	0	40	0	0	60	0	0	45	10	45

Table 2
Descriptives of rooms examinations conducted within.

Room	Performance	Spatial properties					Students							
	Examination	Ceiling height	Int. floor area	Volume	Characteristics		Total		Female		Male		Other	
	M (SD)	m	m2	m3	Function	Location	N	*M* (SD)	N	*M* (SD)	N	*M* (SD)	N	*M* (SD)
1	68.9 (13.7)	2.79	73.00	203.31	Tutorial room	Metropolitan	251	22.9 (7.74)	212	22.8 (7.91)	37	23.4 (6.97)	2	20 (0.00)
2	66.6 (12.7)	9.50	1562.00	14,839.00	Gymnasium	Metropolitan	8802	19.9 (3.89)	6560	19.8 (3.84)	2231	20.3 (3.99)	11	20.4 (4.30)
3	73.0 (10.0)	4.30	177.00	761.10	Multipurpose	Metropolitan	146	19.6 (2.69)	103	19.6 (2.78)	43	19.6 (2.49)	0	N/A
4	70.0 (9.17)	3.00	160.00	480.00	Multipurpose	Metropolitan	140	20.0 (3.81)	103	19.9 (3.48)	37	20.4 (4.65)	0	N/A
5	62.6 (15.3)	3.52	68.86	242.10	Tutorial room	Metropolitan	491	20.1 (4.43)	313	20.4 (5.09)	178	19.7 (2.90)	0	N/A
6	64.5 (14.5)	6.45	1079.84	6964.97	Gymnasium	Regional 1	4734	20.3 (4.95)	3566	20.2 (5.02)	1163	20.4 (4.73)	5	19.6 (2.61)
7	58.5 (10.9)	2.85	246.13	700.24	Multipurpose	Regional 1	126	20.5 (4.23)	97	20.6 (4.16)	29	20.3 (4.53)	0	N/A
8	60.9 (18.7)	3.00	40.56	121.68	Tutorial room	Regional 1	97	22.4 (7.47)	76	21.8 (6.26)	21	24.5 (10.7)	0	N/A
9	71.5 (12.1)	3.00	38.28	114.84	Tutorial room	Regional 1	41	23.9 (5.94)	34	23.4 (5.80)	7	26.3 (6.47)	0	N/A
10	65.8 (13.5)	6.00	227.02	1362.12	Lecture theatre	Regional 2	484	21.8 (6.32)	395	21.6 (6.13)	89	22.9 (7.05)	0	N/A
11	65.7 (16.2)	5.50	136.97	753.34	Lecture theatre	Regional 2	88	29.2 (10.9)	73	29.6 (11.4)	15	27.2 (7.92)	0	N/A

they had undertaken more than one of the three of the units analysed within this study). Refer to Table 1 for the descriptive statistics of the rooms analysed in the study.

To investigate the continuous relationship between examination score and ceiling height, we fitted a linear mixed model (LMM), and bootstrapped 95% confidence intervals (10,000 iterations) to increase the precision of our parameter estimates. Deidentified participant ID was entered into the model as a random intercept to model subject specific differences. To find the best fit, we entered all possible variables that were significant and didn't show collinearity, comparing the Akaike information criterion (AIC) to ensure it was lower than previous models. Significant results were based on $p < 0.05$. As its important to statistically control for student performance outside of the examination, we used the coursework score (as described in 2.2.2) as a co-variate. To understand if the relationship survived when factoring in other variables, we added simple effects for year, age, gender, unit, campus, and previous exposure. Previous exposure was calculated by checking if the student had sat one of the three units in this study previously. Here, 0 = they had not, 1 = they had sat multiple exams the same year (we cannot determine in what order), 2 = had sat an exam in a previous year, and 3 = had sat two or more exams in a previous year. We note this is limited, as we are unable to determine if the student had sat an examination outside of one of these three subjects. Lastly, we were able to obtain data indicating the seat number assigned to students who sat their examination in larger examination venues, however as we could not get this data for all spaces, and we were unable to confirm the numbering system (layout of where numbers were placed and whether this was consistent across years) we excluded this variable from our analysis. In reporting we have followed the standard practice framework (Monsalves, Bangdiwala, Thabane, & Bangdiwala, 2020). All analyses were performed using R Statistical Software v 3.6.3 (R Core Team, 2020). Linear mixed modelling was conducted using the 'lme' and 'lmer' functions in packages lme4 and nlme (Bates, Mächler, Bolker, & Walker, 2015); Pinheiro J (2023) and data visualizations from R were created using 'ggplot' within the ggeffects package (Lüdecke, 2018), illustrating the raw data.

3. Results

As predicted, we found a significant relationship between ceiling height and examination performance [SE = 0.059, $t(3567) = 3.132$, *P* 0.002, 95% CI [0.07, 0.30]]. Overall, the model explained 41.07% of variance in examination scores (La Huis' R^2 approximation = 0.4109), which provides an indication of the percentage of variance observed (LaHuis, Hartman, Hakoyama, & Clark, 2014). Importantly, the effect of

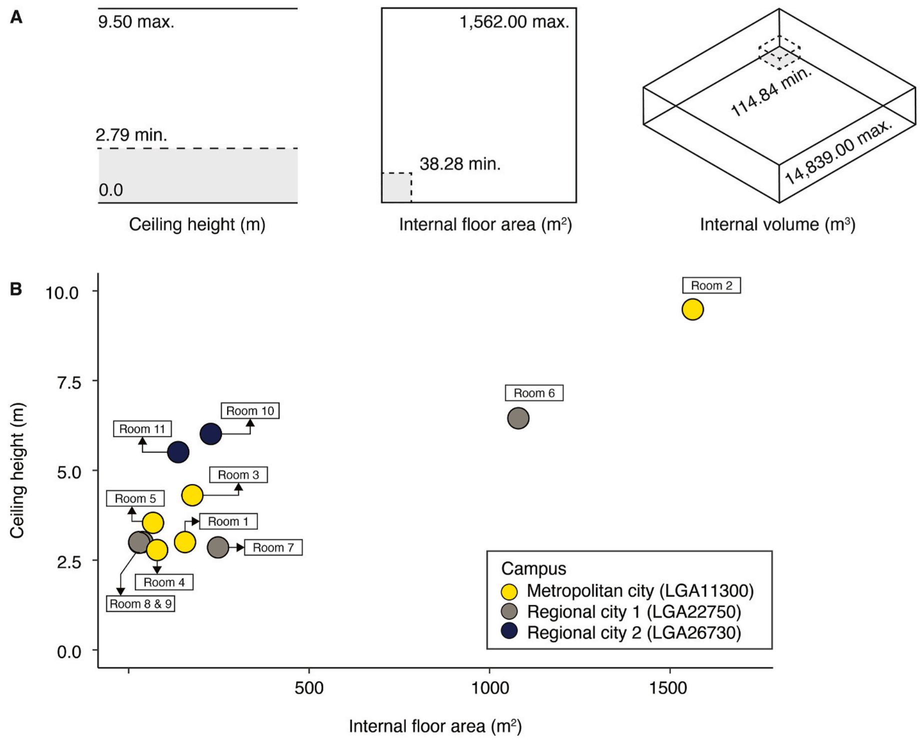

Fig. 2. (A) Two-dimensional orthographic and three-dimensional isometric representation of the relative difference in scale showing the minimum and maximum values of the room's examinations were held within. (B) Dot plot showing the different rooms included in the study based on their ceiling height and internal floor area and color coded by the campus location. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

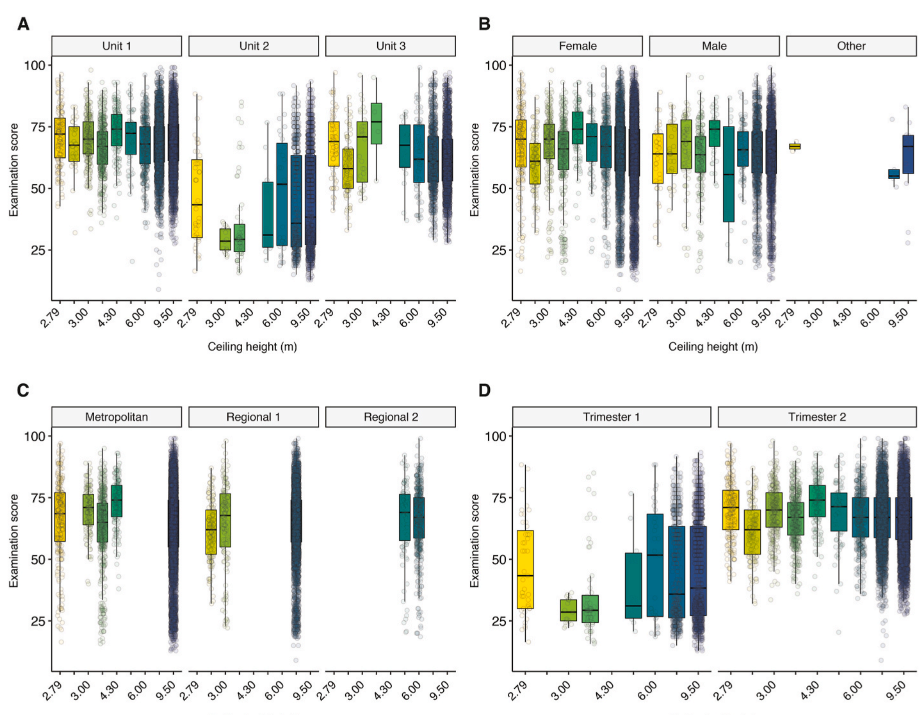

Fig. 3. Distribution of examination scores by ceiling height for different factors. Boxplots with quartile ranges and medians using raw data. Each data point represents a participant (deidentified student ID number). Note the x-axis displays the internal ceiling height as discrete for visualization, which does not reflect the continuous nature of the ceiling heights (increments between each value are treated identically rather than by their value) (A) Unit of study. (B) Gender. Note that gender 'other' refers to indeterminate/intersex/unspecified. (C) Campus location. (D) Trimester when examination occurred.

ceiling height on examination score survived after accounting for numerous variables (i.e., coursework score, year, age, gender, the subject studied, type of campus where the examination was held, and if the student had experience having taken one of more of the three subjects' examinations in a previous year; see Figs. 3 and 4 and Tables 3 and 4).

4. Discussion

We investigated whether built environment scale impacts the cognitive performance of students during examination. With previous work finding that enlarged interior environment scale invokes high frequency oscillatory activity (Bower et al., 2022, 2023) which is also involved in concentration and stress response regulation, we were interested in understanding if exposure to enlarged scale during an examination would have a detrimental effect on performance. Our results confirmed there was a relationship between students' lower examination score when sitting the assessment in a room with an elevated internal ceiling height, compared to those sitting in a regular room. Here, we are referring to what's commonly known as the 'standard' ceiling height. While the regulated height varies slightly in different countries, in Australia the National Construction Code requires habitable rooms to be of a 2.4 m minimum (Australian Building Codes Board, 2022).

A key constraint of this naturalistic retrospective study is that we are unable to probe whether the observed result is direct result of the design quality of scale, or if differences arise because of the indoor environmental parameters afforded by the scale. For instance, as enlarged gymnasium spaces are often poorly insulated and are expensive to climate control, the observed effect may be due to lowered ambient temperature on the students, which has been shown to reduce cognitive function in young adults (Mäkinen et al., 2006; Muller et al., 2012). The enlarged room scale also results in an increased number of occupants in the space, with several studies suggesting density and crowding can result in negative affect, resulting in a deterioration of performance in cognitive complex tasks (Evans, 1979; Paulus, 1976). We also cannot rule out the effect of context and familiarity with the room. Here, pre-existing associations with the space usage may have a priming effect on students, with the enlarged spaces (both gymnasiums) commonly used for team sports and activities. Lastly, the smaller room scale may allow students more opportunities to cheat. While the proctor to student ratio is far higher, the student-to-student ratio is lower reducing peer surveillance and monitoring which may influence if a student cheats by smuggling in notes. Despite not being able to invoke specific causes, the key point is that enlarged environments seem to be disadvantaging students; future studies are required to question and answer why.

Although this is an example of applied research that could have broad application across society, it is important to note it is currently unlikely this study population could be reproduced at present due to disruptions to examination procedures and practices introduced from the COVID-19 pandemic, with many students at this institution now completing their examinations online in environments of their choice. For replicability, it is important to have a measure of academic performance outside of the examination venue and preferably over the result of assessments undertaken over the course of study as a control measure of the students' overall academic performance to contrast the examination score against. Here, we note not all institutions or jurisdictions follow this practice. Lastly, we note that studies show that characteristics of our cohort (younger, female dominant, psychology field) may not generalize to other courses and cohorts. Studies have also shown links between examination stress and negative psychological and physiological health effects, such as mental distress (Fritz, Stochl, Kievit, van Harmelen, & Wilkinson, 2021) and increased blood pressure (Hughes, 2005). Given the adverse health effects, the secondary benefits of minimizing stress could result in positive improvements to mental (anxiety, depression, burnout) and physical health (blood pressure, immune system functioning) alongside societal benefits to reducing burden on systems and practitioners.

In summary, we identified examination performance is associated with the scale of the built environment. This may be due to the scale itself or be linked to the indoor environmental parameters afforded by the scale. Regardless, it demonstrates the importance of understanding

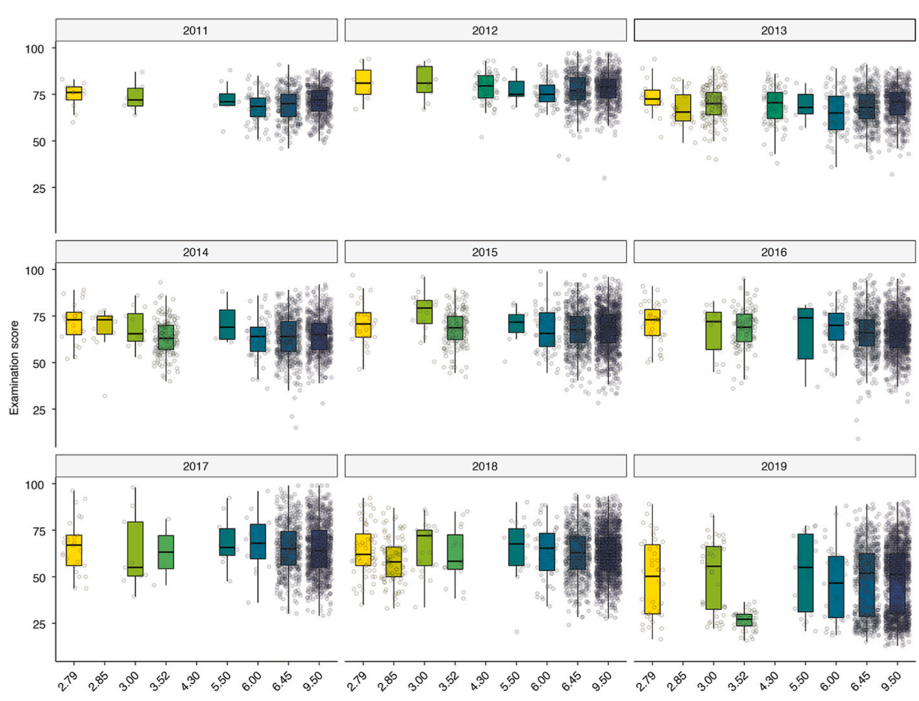

Fig. 4. Distribution of examination scores by ceiling height by year. Boxplots with quartile ranges and medians using raw data. Each data point represents a participant (deidentified student ID number). Note the x-axis displays the ceiling heights as discrete for visualization, which does not reflect the continuous nature of the ceiling heights (increments between each value are treated identically rather than by their value).

Table 3
Correlations between variables. Note: In the case of categorical values with more than two categories, dummy coding was undertaken. The first category (lowest value) was taken as a reference as these were meaningful in all instances. For example, for gender, comparisons are with reference to female, unit references the first unit of study, campus references regional and rural to metropolitan locality, and previous exposure reference taking a second or subsequent exam compared to having not sat an exam for one of these three units in a previous year.
Gender other refers to Indeterminate/Intersex/Unspecified.

	(Intercept)	Coursework score	Ceiling height	Year	Age	Gender (male)	Gender (x)	Unit 2	Unit 3	Campus (regional)	Campus (rural)	Previous exposure (1)	Previous exposure (2)
Coursework score	0.021												
Ceiling height	0.024	0.017											
Year	−1.000	−0.026	−0.029										
Age	0.036	−0.071	0.055	−0.040									
Gender (male)	0.004	0.096	0.008	−0.005	−0.033								
Gender (other)	0.012	−0.002	0.006	−0.012	−0.001	0.018							
Unit 2	0.358	0.148	−0.011	−0.359	−0.006	0.027	−0.004						
Unit 3	0.231	0.002	−0.066	−0.231	−0.019	0.010	−0.005	0.369					
Campus (regional 1)	−0.008	−0.013	0.575	0.004	0.005	0.020	0.006	0.018	0.008				
Campus (regional 2)	−0.060	0.018	0.303	0.058	−0.099	0.045	0.007	−0.001	−0.001	0.292			
Previous exposure (1)	0.418	−0.069	−0.028	−0.418	0.057	0.049	−0.006	−0.068	−0.190	−0.013	−0.001		
Previous exposure (2)	0.121	0.007	−0.006	−0.121	−0.038	0.018	−0.007	−0.028	−0.024	−0.008	−0.005	0.131	
Previous exposure (3)	0.022	−0.002	−0.004	−0.022	−0.011	0.007	0.001	−0.005	−0.008	0.003	0.002	0.023	0.035

Table 4
Fixed effects for model predicting examination performance. Note: Confidence intervals estimated using percentile bootstrapping and 10,000 iterations.

Predictor	Estimate	SE	95% CI*	df	t	*p-value*
(Intercept)	4148.65	99.33	[3948.52, 4343.25]	11810	41.514	<0.001***
Coursework score	0.299	0.006	[0.29, 0.31]	3567	45.729	<0.001***
Ceiling height	0.185	0.059	[0.07, 0.30]	3567	3.132	0.002**
Year	−2.039	0.050	[-2.14, −1.94]	3567	−41.094	<0.001***
Age	0.292	0.021	[0.25, 0.33]	3567	13.679	<0.001***
Gender (male)	−0.372	0.224	[-0.83, 0.07]	11810	−1.659	0.097
Gender (other)	1.315	2.968	[-4.67, 7.02]	11810	0.443	0.658
Unit 2	−16.092	0.308	[-16.71, −15.49]	3567	−52.214	<0.001***
Unit 3	−4.309	0.222	[-4.75, −3.87]	3567	−19.383	<0.001***
Campus (regional 1)	−0.583	0.257	[-1.09, −0.08]	3567	−2.272	0.023*
Campus (regional 2)	−1.756	0.537	[-2.82, −0.72]	3567	−3.270	0.001**
Previous exposure (1)	3.526	0.252	[3.03, 4.02]	3567	13.970	<0.001***
Previous exposure (2)	2.027	0.674	[0.70, 3.34]	3567	3.007	0.003**
Previous exposure (3)	1.315	5.247	[-9.10, 11.69]	3567	0.251	0.802

how environments affect cognitive performance and suggests that if we want to achieve the best possible assessment of student capabilities, we should move away from conducting examinations in rooms with elevated ceiling-heights. This study could have significant implications for how we assess students to ensure we are not unwittingly adding disadvantage into performative evaluations. Most importantly, our data provide evidence of how the built environments we occupy impact our ability to perform tasks. This warrants further investigation and placing more effort in uncovering whether we can optimize the built environment to have a positive impact on cognitive functioning.

Competing interest statement

Authors declare that they have no competing interests.

Pre-registration and materials

This study was preregistered (https://doi.org/10.17605/OSF.IO/C9HF5). The data and code to support this study are available here (https://doi.org/10.17605/OSF.IO/GY8M4).

Funding

The lead author (ISB) was financially supported through employment by a Postgraduate Research Scholarship from Deakin University and a Postdoctoral Research Fellowship from the University of South Australia. This work received no direct funding.

CRediT authorship contribution statement

Isabella S. Bower: Writing – review & editing, Writing – original draft, Visualization, Methodology, Formal analysis, Data curation, Conceptualization. **Jaclyn Broadbent:** Writing – review & editing, Project administration, Methodology, Conceptualization. **Scott Coussens:** Writing – review & editing, Validation, Formal analysis. **Peter G. Enticott:** Writing – review & editing, Methodology, Conceptualization.

Acknowledgments

We thank M. Barham for data blinding, as well as H.A.D. Keage, T. Loetscher, M. Fuller-Tyszkiewicz, and A.C. Thompson for providing feedback on the statistical analysis design for the project.

References

Amigud, A., Arnedo-Moreno, J., Daradoumis, T., & Guerrero-Roldan, A.-E. (2018). An integrative review of security and integrity strategies in an academic environment: Current understanding and emerging perspectives. *Computers & Security, 76*, 50–70. https://doi.org/10.1016/j.cose.2018.02.021

Australian Bureau of Statistics. (2021). Census. https://www.abs.gov.au/census.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bower, I. S., Clark, G. M., Tucker, R., Hill, A. T., Lum, J. A. G., Mortimer, M. A., et al. (2022). Enlarged interior built environment scale modulates high-frequency EEG oscillations. In *eneuro* (Vol. 9, pp. 104–122). ENEURO. https://doi.org/10.1523/eneuro.0104-22.2022, 2022.

Bower, I. S., Hill, A. T., & Enticott, P. G. (2023). Functional brain connectivity during exposure to the scale and color of interior built environments. *Human Brain Mapping, 44*(2), 447–457. https://doi.org/10.1002/hbm.26061

Bower, I. S., Tucker, R., & Enticott, P. G. (2019). Impact of built environment design on emotion measured via neurophysiological correlates and subjective indicators: A systematic review. *Journal of Environmental Psychology, 66*, Article 101344. https://doi.org/10.1016/j.jenvp.2019.101344

Brink, H. W., Loomans, M. G. L. C., Mobach, M. P., & Kort, H. S. M. (2021). Classrooms' indoor environmental conditions affecting the academic achievement of students and teachers in higher education: A systematic literature review. *Indoor Air, 31*(2), 405–425. https://doi.org/10.1111/ina.12745

Butler-Henderson, K., & Crawford, J. (2020). A systematic review of online examinations: A pedagogical innovation for scalable authentication and integrity. *Computers & Education, 159*, Article 104024. https://doi.org/10.1016/j.compedu.2020.104024

Cassaday, H. J., Bloomfield, R. E., & Hayward, N. (2002). Relaxed conditions can provide memory cues in both undergraduates and primary school children. *British Journal of Educational Psychology, 72*(4), 531–547. https://doi.org/10.1348/00070990260377596

Deakin University. (2023). Campus-based exams. Retrieved from https://www.deakin.edu.au/students/study-support/assessments-and-examinations/eoua-and-exams/exams/exam-locations/campus-based-exams.

Evans, G. W. (1979). Behavioral and physiological consequences of crowding in humans. *Journal of Applied Social Psychology, 9*(1), 27–46. https://doi.org/10.1111/j.1559-1816.1979.tb00793.x

French, S., Dickerson, A., & Mulder, R. A. (2023). A review of the benefits and drawbacks of high-stakes final examinations in higher education. *Higher Education*. https://doi.org/10.1007/s10734-023-01148-z

Fritz, J., Stochl, J., Kievit, R. A., van Harmelen, A. L., & Wilkinson, P. O. (2021). Tracking stress, mental health, and resilience factors in medical students before, during, and after a stress-inducing exam period: Protocol and proof-of-principle analyses for the RESIST cohort study. *JMIR Form Res, 5*(6), Article e20128. https://doi.org/10.2196/20128

Hoque, S., & Weil, B. (2016). The relationship between comfort perceptions and academic performance in university classroom buildings. *Journal of Green Building, 11*(1), 108–117. https://doi.org/10.3992/jgb.11.1.108.1

Hughes, B. M. (2005). Study, examinations, and stress: Blood pressure assessments in college students 1. *Educational Review, 57*(1), 21–36. https://doi.org/10.1080/0013191042000274169

LaHuis, D. M., Hartman, M. J., Hakoyama, S., & Clark, P. C. (2014). Explained variance measures for multilevel models. *Organizational Research Methods, 17*(4), 433–451. https://doi.org/10.1177/1094428114541701

Lüdecke, D. (2018). ggeffects: Tidy data frames of marginal effects from regression models. *Journal of Open Source Software, 3*(26), 772. https://doi.org/10.21105/joss.00772

Mäkinen, T. M., Palinkas, L. A., Reeves, D. L., Pääkkönen, T., Rintamäki, H., Leppäluoto, J., et al. (2006). Effect of repeated exposures to cold on cognitive performance in humans. *Physiology & Behavior, 87*(1), 166–176.

Metzger, R. L., Boschee, P. F., Haugen, T., & Schnobrich, B. L. (1979). The classroom as learning context: Changing rooms affects performance. *Journal of Educational Psychology, 71*(4), 440–442. https://doi.org/10.1037/0022-0663.71.4.440

Meyers-Levy, J., & Zhu, R. (2007). The influence of ceiling height: The effect of priming on the type of processing that people use. *Journal of Consumer Research, 34*(2), 174–186. https://doi.org/10.1086/519146

Monsalves, M. J., Bangdiwala, A. S., Thabane, A., & Bangdiwala, S. I. (2020). LEVEL (logical explanations & visualizations of estimates in linear mixed models): Recommendations for reporting multilevel data and analyses. *BMC Medical Research Methodology, 20*(1), 3. https://doi.org/10.1186/s12874-019-0876-8

Muller, M. D., Gunstad, J., Alosco, M. L., Miller, L. A., Updegraff, J., Spitznagel, M. B., et al. (2012). Acute cold exposure and cognitive function: Evidence for sustained impairment. *Ergonomics, 55*(7), 792–798.

National construction code. *10.3.1*, (2022).

O'Sullivan, B., & Cheng, L. (2022). Lessons from the Chinese imperial examination system. *Language Testing in Asia, 12*(1), 52. https://doi.org/10.1186/s40468-022-00201-5

Paulus, P. B. (1976). Density does affect task performance. *Journal of Personality and Social Psychology, 34*(2), 248–253. https://doi.org/10.1037/0022-3514.34.2.248

Pinheiro, J. B. D., & R Core Team. (2023). nlme: Linear and nonlinear mixed effects models. https://CRAN.R-project.org/package=nlme.

R Core Team. (2020). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/.

Saufley, W. H., Otaka, S. R., & Bavaresco, J. L. (1985). Context effects: Classroom tests and context independence. *Memory & Cognition, 13*(6), 522–528. https://doi.org/10.3758/BF03198323

Sheppard, S. (1996). An informal history of how law schools evaluate students, with a predictable emphasis on law school final exams. *UMKC L. Rev., 65*, 657.

The University of Adelaide. (2022). *Examination venues*. https://www.adelaide.edu.au/student/exams/examinations/examination-venues.

The University of Melbourne. (2023). *On-campus written exams*. https://students.unimelb.edu.au/your-course/manage-your-course/exams-assessments-and-results/exams/on-campus-written.

Weir, W., & May, R. B. (1988). Environmental context and student performance. *Canadian Journal of Education/Revue canadienne de l'éducation, 13*(4), 505–510. https://doi.org/10.2307/1495290