The Effect of Survey Scale Sizes on How People Assess the Effect of the Built Environment on their Work Performance

By

Issei Yuki

A thesis
submitted to Victoria University of Wellington
in fulfilment of the requirements for the degree of
Master of Building Science

Victoria University of Wellington

2020

Preface

This thesis was submitted in fulfilment of the requirements of the degree of Master of Building Science to the School of Architecture, Victoria University of Wellington.

Author

Issei Yuki

School of Architecture

Victoria University of Wellington

Email: issei_1993_nz@yahoo.co.jp

Supervisor

Dr. Michael Donn

Associate Professor

School of Architecture

Victoria University of Wellington

Email: michael.donn@vuw.ac.nz

Abstract

Increase in employee work performance could lead to additional savings for businesses, and therefore investments to improve the indoor work environment could be very cost effective. However, these business decisions are based on self-reported work performance measures from surveys. Doubts have been raised as to the accuracy of these measures.

In addition to the uncertainties of the self-reported measures, two major post-occupancy surveys by Building Use Studies (BUS) and Occupant IEQ surveys by the Center for the Built Environment (CBE) use two different scales for occupants to report the change in their work performance. This research examines how the two scales influence people in how they report the environmental effect on their work performance.

An experiment has been conducted under three pink-noise conditions to obtain self-reported work performance measures using these two different scales for comparison. Other self-reported indicators of work performance such as motivation, job satisfaction, fatigue and distraction were also collected through surveys, and cognitive tests were conducted to obtain objective measures of work performance.

The results from this experiment show trends for the BUS scales that are 2–3 times larger than results from using CBE scales. These results are significant when the results are correlated with other self-reported indicators; however, the differences were not statistically significant when the scores were correlated with environmental conditions and cognitive performance. The data suggests that the two scales generate different results.

Different scales leading to different results suggests that use of these two surveys could lead business investment analysis in misleading directions. If the productivity bonus of an investment is 5% according to one and 10% according to the other, there is a problem. Business investment analysis using self-reported work performance as an estimate of actual work performance could be misleading if self-reported work performance is used to directly calculate hours of work and savings achieved. Further studies should recruit more subjects, consider other environmental stressors and examine differences obtained by using different survey questions.

Confidentiality Statement

Due to the potential human ethical considerations within this study, the author and participants of the study have agreed that all identifiable information is kept confidential. To maintain confidentiality, pseudonyms are used throughout this paper to replace all identifiable information, such as “participants”, “groups” and “students”. For the purposes of this study, however, the participants have agreed to be identified by university year – that is, Master’s students, undergraduate students.

Contents

Figures

Tables

Acknowledgements

I would first like to thank my thesis supervisor, Associate Professor Michael Donn of the School of Architecture at Victoria University of Wellington. He was always open to questions and discussions and has consistently guided me in the right direction for my research whenever it was necessary.

I would also like to acknowledge the School of Architecture IT office staff who have organised, prepared and taught me how to set up the equipment for the experiment, and the School of Architecture Resources Centre staff for teaching me how to use the tools required for measuring the experimental environment; and Terry O'Donnell of the Centre for Translational Physiology for providing information on how environmental experiments in the laboratory are conducted and on possible options for my research.

I would also like to thank Emeritus Professor George Baird of the School of Architecture at the University of Wellington, for providing me with much information and leading me to researchers and professionals in the field; Steve Lamb of the School of Architecture at Victoria University of Wellington, who was always open to discussions on the different options for my research and gave me critical feedback on my presentations and research; James Sullivan of the School of Architecture at the University of Wellington, for supporting my research ideas, directing me to relevant studies and for teaching and helping me with statistical analysis; and my fellow Master's students, for helping with the pilot tests and discussing the results and analysis.

Finally, I express my gratitude to my parents for providing me with support and encouragement throughout my years of studies in different universities and different fields. You have helped me to focus and achieve this accomplishment. Thank you.

Special thanks to the participants who have spent their weekends on my experiment and provided data that made this research possible.

1 What is Work Performance, Why is it Important and How do we Measure It?

Like many other organisations, the Green Building Council of Australia (2013) has claimed that moving into green buildings has increased self-reported work performance of workers by 10.9%. They estimate this leads to a saving of AUS$2m per year. Their argument is based on the reasonable assumption that moving into office buildings with better environments could lead to higher work performance and so achieve savings. However, this thesis examines the reliability of the measures used to estimate the improvement in work performance. There are significant differences in the two most commonly used work performance scales that estimate performance improvement.

Many researchers have demonstrated that work performance is influenced by the indoor environmental quality of a building (Boerstra et al., 2015; Clausen and Wyon, 2008; Cui et al., 2013; Kekäläinen et al., 2010; Kershaw and Lash, 2013; Kobayashi et al., 2005; Lan et al., 2009; Lan, Wargocki and Lian, 2011; Leaman and Bordass, 1999; Niemelä et al., 2002; Tanabe, Haneda and Nishihara, 2015; Tham and Willem, 2010; Valančius & Jurelionis, 2013). To increase occupants' work performance it is important to be able to quantify people's work performance reliably; however, measuring work performance is very complex. This research started with the question:

Is there an easy method to measure the effect of the built environment on people's work performance?

The indoor environment affects people and their work performance, which could significantly affect companies' profits, because costs for salaries and wages are estimated to be significantly higher than initial construction, maintenance and operational costs (Evans et al., 1998, cited in Sullivan, Baird and Donn, 2013; Wu and Clements-Croome, 2005, cited in Clements-Croome, 2011a). Due to the complexity of measuring actual work performance in offices, surveys have commonly been used to ask people to self-assess their work performance. Questions have been raised, however, as to whether these assessments show an accurate value of people's actual work performance in offices (Cui et al., 2013; Haynes, 2008; Feige et al., 2013; Leaman and Bordass, 1999; Sullivan, Baird and Donn, 2013; Tanabe et al., 2013).

Within the overall research question, the research focused on one simple research issue:

The accuracy of self-assessed work performance as an alternative measure of actual work performance in the built environment.

The focus of the research examines two major surveys where participants self-report work performance and asked to assess the effect of the work environment on their work performance, and it asks two simple questions:

1) Do the different scales in these surveys affect the size of the participants' responses?
2) How does the size of the effect as measured via self-reporting compare with the results of simple cognitive tests of objectively measured performance?

1.1 Work Performance

For business managers and owners, it is of great importance to increase the productivity of workers to increase overall profit for the business. The ratios of initial construction costs, maintenance and operations costs, and the salaries of workers in 20 years have been variously estimated to be 1:5:200 (Evans et al., 1998, cited in Sullivan, Baird and Donn, 2013) or 1:8:80 (Wu and Clements-Croome, 2005, cited in Clements-Croome, 2011a). Although not proven, these ratios have been mentioned in a number of studies (Adamson, 2004; Macmillan, 2006; Parke and Taylor, 2014; Saxon, 2005) suggesting that, although the numbers might not be accurate, business costs for salaries and wages tend to be significantly higher than initial construction and operation and maintenance costs.

This disparity between the costs means that small increases in work performance could have great value; therefore, improving the building's indoor environment through renovation or increasing running costs (such as setting HVAC systems to provide optimum indoor environmental conditions) could be a very cost-effective investment to increase occupants' work performances (Feige et al., 2013; Lan, Wargocki and Lian 2011). Cost–benefit calculations by Wargocki and Djukanovic (2003) have shown that improving the indoor air quality through the use of heating, ventilation and air-conditioning systems will generate net productivity gains that exceed the investment costs with a simple payback period of no longer than 2.1 years.

Measuring the effect of the built environment on people's work performance is very complex, however. This is because the output of work in offices is difficult to measure, and the profits or losses do not necessarily represent how efficiently people have been working. Work and task outputs in the office environment require a great variety of different skills, but only a few can be quantified (Humphreys and Nicol, 2008; Lan and Lian, 2009; Mallawaarachchi, 2017). Much office work ,such as report writing, has different levels of importance, and converting quality into quantified measures would be difficult, as quality is subjective (Sullivan, Baird and Donn, 2013). Weighting the different tasks by their importance would also be very complex, as it is difficult to state that one job is X% more important than the other, and one person's job would not be comparable between different roles and companies (Sullivan, Baird and Donn, 2013).

1.2 Perceived Work Performance

Due to the complexity of measuring actual office work, perceived work performance measures are used to obtain an indication of the occupants' work performances. Surveys and questionnaires are handed out to office occupants or experimental subjects, at low cost, and measurements are collected relatively quickly and are easy to conduct; therefore, they are an easy, cheap and fast method to measure how the environment affects people’s work performance.

Surveys and questionnaires could be applied to people with different work tasks and positions, without the limitations of their differing work natures, which allows collection and comparison of large sample sizes throughout different buildings. Surveys in offices can include scales for the occupants to assess their performance for a day, week, month or years, (Halpern, 2000, cited in Sullivan et al., 2013, p.19; Tanabe et al., 2013, p.12; Tanabe et al., 2015, p.44; Toftum et al., 2010, p.63; World Health Organization, 2001, cited in Sullivan et al., 2013, p.19), or during the test if work performance has been measured in an experiment.

1.3 Initial Research Question - Accuracy of Perceived Work Performance

There has been a number of studies on the reliability of perceived performance measures to be an alternative measure of actual office work performance. Past studies have noted that the changes in perceived work performance and perceived thermal comfort are strongly related (Leaman and Bordass, 1999; Leaman, 2010; Frontczak, 2011; Silva et al. 2000), and actual performance and perceived performance are affected in the same manner by other changes in environmental conditions (Boerstra et al., 2015; Clausen and Wyon, 2008; Kekäläinen et al., 2010).

However, studies by Cui et al. (2013), Haynes (2008), Feige et al. (2013), Leaman and Bordass (1999), Sullivan, Baird and Donn (2013) and Tanabe et al. (2013) mention that we cannot always trust people's judgements of their work performance. For example, people tend to underestimate their performance for tasks that they are not good at, and overestimate for tasks that they are good at (Sullivan, Baird and Donn, 2013). In addition, Clausen and Wyon (2008) state that people tend to overestimate their work performance when they are just asked to assess their performance.

There is also another question as to how different survey scales and questions can affect the accuracy of people’s perception of their performance. Self-reported work performance does indeed correlate with actual work performance, but the values assessed by the participants are not always an accurate reflection of their actual work performance, and in addition, the surveys could also have an effect on people's judgements.

This generates a more narrowed down research question,

"How accurate is perceived work performance as an alternative measure of actual work performance?"

1.4 Is it Correct to Make Business Decisions based on Self-Reported Work Performance Measures Only?

Businesses could make investment decisions to increase the employee work performance through improving the work environment; however, are self-reported work performance measures alone enough to make business decisions to invest in retrofits or designing a green building?

The effects of the work environment on employees' work performance and how much businesses could save from improving the indoor environment were calculated in the 1980s. The U.S. Environmental Protection Agency (1989) stated that poor indoor quality will reduce productivity by 3%, which equates to 14 minutes a day, and if this is applied to white-collar labour forces in the nation, the cost to the nation will be estimated as US$60 billion annually. Singh et al. (2010) observed an increase in perceived productivity from −0.80% to 2.18% from surveys of occupants that had moved from a conventional office to a green building, estimating an additional 39.98 hours of work per year per person from the increase in perceived productivity. Miller et al. (2009) also observed a 4.88% increase in perceived productivity from occupants; they agreed that employees are more productive in green buildings and have linked this to a net impact of US$5204 per employee, equivalent to a 4.88% increase in the average salary. Also, the Green Building Council of Australia (2013) stated that moving to a 6-star Green Star building (One Shelly Street in Sydney) increased perceived productivity by 15%, linking this to additional hours per year worked per employee. Other examples used are post-occupancy surveys after moving into Council House 2 (CH2), which found an increased productivity of 10.9%, and this is estimated to save AU$2 million annually; and staff perception of productivity increased by 6.2% by moving into Trevor Pearcey House which was estimated to produce a "small productivity improvement" that adds a benefit of $1.5 million over five years (Green Building Council of Australia, 2013).

Although these studies are not business cases, managers and building owners could make business decisions based on benefits from increases in work performance. If the accuracy of perceived productivity or self-reported work performances is questionable, then is it correct to link self-reported work performance straight to actual work performance, to the hours of work saved and the annual savings achieved for the employee costs? Savings that could be achieved through an increase in work performance could change with different methods to obtain self-reported work performance, which could greatly mislead business owners to think that companies will definitely achieve certain cost savings through work environment upgrades.

1.5 How Closely is Cognitive Performance Related to Actual Work Performance in Office Environments?

Office work performance is said to have a strong relationship with cognitive performance, because the business economy has changed from manufacturing-based work to service-based and knowledge-based work that requires cognitive function work (Cascio, 1995, cited in Purdey and Leifer, 2012; Humphreys and Nicol, 2007; Mawson, 2002, cited in Haynes, 2007).

Therefore, studies by Lan et al. (2009) and Purdey and Leifer (2012) argue that office work covers a wide range of skills with increasing emphasis on cognitive demands. The neurobehavioral approach to measure work performance by Lan et al. (2009) identifies and measures the influence of environment on brain functions and its behavioural changes. Because the central nervous system shows a particular sensitivity to environmental stress, the neurological approach has been neuro-biologically justified (Hancock and Vasmatzidis, 1998, cited in Lan et al., 2009).

Lezak, Howieson and Loring (2006, cited in Lan et al., 2009) conceptualised behaviour in three functions:

1. Cognition – information handling.
2. Emotionality – feelings and motivations.
3. Executive functions – how behaviour is expressed.

Lan et al. (2009) further categorise cognitive functions into perception, memory and learning, thinking concerns, and expressive functions:

1. Perception – ability to select, acquire, classify, and integrate information (Lan et al., 2009).
2. Memory and learning – information storage, retrieval of information, working memory, short term and long term memory. Access to knowledge is central to all cognitive functions (Lezak, Howieson and Loring, 2006, cited in Lan et al., 2009).
3. Thinking – mental organisation and reorganisation of information, relating two or more items of information explicitly and implicitly (Lezak, Howieson and Loring, 2006, cited in Lan et al., 2009).
4. Expressive – how information is communicated or acted on through written or oral expression (Lan et al., 2009).

These four different functions make a distinct class of behaviour, but they work closely and are bound together by the same activity (Lan et al., 2009). To cover this wide range of cognitive functions, there are different cognitive tests to measure each one of them. Some of these tests are:

- Simple office work such as text typing and arithmetic (Boerstra, 2015; Clausen and Wyon, 2008; Hongisto et al., 2016; Lan et al., 2009; Park and Yoon, 2011; Tanabe, Nishihara and Haneda, 2007; Tanabe, Nishihara and Haneda, 2015; Tham and Willem, 2010; Toftum et al., 2010, Valančius and Jurelionis, 2013).
- Proofreading (Clausen and Wyon, 2008; Park and Yoon, 2011; Tanabe, Nishihara and Haneda, 2015; Tham and Willem, 2010; Toftum et al., 2010).
- Landolt ring task (Clausen and Wyon, 2008).
- Memory tasks (Cui et al., 2013; Hongisto et al., 2016; Lan et al., 2009, Park and Yoon, 2011).
- Letter search (Lan et al., 2009, Park and Yoon, 2011).
- Shape overlapping (Lan et al., 2009).
- Picture recognition (Lan et al., 2009).
- Verbal reasoning (Lan et al., 2009; Tham and Willem, 2010).
- Spatial image (Lan et al., 2009).
- Visual choice reaction time (Lan et al., 2009).
- Creative thinking (Tanabe, Nishihara and Haneda, 2015; Tham and Willem, 2010).
- Tsai–Partington test (Boerstra, 2015; Valančius and Jurelionis, 2013).

Since each test assesses one aspect of cognitive performance, it is common to use a series of different tests in one experiment to get a better understanding of how people perform under different environmental conditions. The results from the tests will likely have a link to how people are affected by the work environment, but, as no evidence exists for the actual relationship, they should only be considered as another self-reported indicator for work performance. This is because the relationship between cognitive performance and “actual” work performance cannot be established due to the difficulty of measuring “actual” work performance, as discussed in section 1.1. In addition, the results of cognitive performance would most likely be job dependent. Thus, a person with a higher score in cognitive performance would not necessarily be more productive at their work.

1.6 Aim of this Study

This study examines whether self-reported work performance measures could be an alternative measure of actual work performance and if it could be considered accurate enough to make investment decisions for businesses. As alternatives to measure work performance, studies can use other measures such as simulated office work performance, cognitive work performance and some studies attempt to measure actual office work performance for particular types of jobs. However, since self-reported work performance measures could be obtained easily with low cost and less time, this study will focus on how accurately the self-reported work performance measures represent actual work performance.

One of the main questions for this research is "How do the self-reported work performance measures align with the objective measures of work performance?" This is examined in section 2, through a review of literature measuring both self-reported and objective measures of work performance, as well as other questions including:

- "What is the definition of work performance/productivity?"
- "What is the definition of comfort/well-being for work performance?"
- "What environmental conditions to affect work performance have been tested, and is there a significant difference between having one or multiple environmental stressors on work performance?"
- "What are the common types of questions and scales used in surveys to measure self-reported work performance?"
- "What are some other common indicators used to estimate the change in work performance?"

Through answering these questions, I have established that there are limited studies on survey questions that ask the occupants to assess the environmental effect on their work performance. The two major surveys that use these questions, as well as using a numerical scale, use very differently sized scales. An experiment has been conducted to answer the main question of this study:

"How do scale sizes influence how people assess the indoor environmental effect on their work performance?"

2 Relationship Between People, Work Performance, and the Built Environment

This section summarises the key questions on how the built environment affects people and their work performance, and how people perceive the effect of the indoor environment on their work performance. This review of the literature about occupants' work performance in offices explores the problems that exist in self-reported work performance measures and determines what is required to sufficiently improve the quality of self-reported work performance so that it can be used as a measure for actual work performance.

The review initially focuses on studies that investigated both self-reported work performance and objective measures of work performance to observe how the self-reported measures align with changes in the objective measure of work performance. A set of key questions was formulated in order to focus the literature search on finding papers on both types of work performance measures.

The process for reviewing the literature and answering the key questions and findings is described in the following sections:

1) Section 2.1 outlines the key questions about the effect of the built environment on occupants' work performance in offices.
2) Section 2.2 describes the process of searching the relevant studies.
3) Section 2.3 lists different definitions of keywords and compares studies researching the environmental effect on occupants' work performance.
4) Section 2.4 describes different methods of collecting self-reported work performance through surveys.
5) Section 2.5 identifies a potential problem with two of the survey methods found in the literature and revises the research question to:

"How do scale sizes influence how people assess the indoor environmental effect on their work performance?"

6) Section 2.6 summarises the main points from the review of relevant research, which leads into designing a methodology to answer the main question of this study in Section 3.

2.1 Key Questions About Self-Reported Work Performance and the Effect of the Built Environment

To obtain measures of change in work performance, self-reported measures are commonly used as discussed in section 1.2, but there are limited studies that have compared self-reported work performance against objective measures of work performance to examine the accuracy. To clearly define the key terms relating to work performance, and to obtain further information about the environmental effects on people's work performance in the built environment, the key research questions used to focus the literature review are:

1. Results from studies with both objective and subjective measures of work performance.
2. Definition of work performance/productivity.
3. Definition of comfort/well-being.
4. Environmental effect on measured performance.
5. Types of questions and scales used in surveys.
6. Common indicators used to estimate the change in work performance.

Databases and search engines used for the search, suggested by a subject librarian, include ScienceDirect, ProQuest, Taylor & Francis, Sage, Avery Index to Architectural Periodicals, JSTOR and Google Scholar.

2.2 Literature Search Process

2.2.1 Scoping Searches

Initial scoping searches were undertaken, based on keywords from the topic "work performance", productivity and building. Google Scholar showed the highest number of results, followed by Taylor & Francis, ScienceDirect and JSTOR. ProQuest and Avery Index to Architectural Periodicals (also a ProQuest database) showed the lowest number of results relating to the search terms. The numbers of search results for each database are shown inTable 1 Modifying search terms, as well as subject and publication filters, has been used to further refine and obtain information from the relevant sources.

Titles from the first pages of searches were manually screened to identify the keywords that could potentially be used for further refinement in the search terms or in subjects. Table 2shows the keywords to be included and excluded from manual scoping and search terms, discussed in the following sections. Keywords included in the search were terms relating to self-reported work performance surveys, or work performance, and terms that could relate to how the building or the environment could have an effect on people. Keywords excluded were terms relating to agricultural,

Table 1. Databases and results for scoping search.

Search term: ("work performance" OR productivity) AND building		
Database	**Field**	**Search Results**
ScienceDirect	All	171,311
ProQuest		50
Taylor & Francis		222,958
Sage		96,714
Avery Index to Architectural Periodicals		50
JSTOR		101,480
Google Scholar		3,370,000

production and mechanical science which had been included in the first search because of the search term "productivity". Building related terms such as "construction productivity" and "residential" have also been removed, because the main focus of this research is not about work performance of labourers but on work performance in commercial buildings and not residential buildings.

Table 2. Relevant and irrelevant keywords from scoping search.

Keywords relating to research	Keywords not relevant to research
Environment	Agriculture
Thermal comfort	Soil
Indoor Environmental Quality (IEQ)	Construction productivity
Perceived	Machine
Self-assessed	Manufacture
Self-estimated	Production
Worker performance	Species
Task performance	Product
Energy	Safety
	Residential
	Disease

2.2.2 Inclusion and Exclusion Criteria

The search was conducted again in ScienceDirect with a modified search term using keywords from Table 2. ScienceDirect was used as the starting database as its search options allows selection of different science categories for searches and the "Expert Search" function allows searches to be in Boolean format. The process for how the search term has developed in ScienceDirect, based on inclusion and exclusion criteria, subject and journal publication filters, is in Table 3. Searches in other databases have been repeated using the final search term, but with different subject and journal filters used in each of the databases.

Table 3. Search example in ScienceDirect.

[] = science categories, {} = journal publications, tak = search in title + abstract + keywords	
Search Term	**ScienceDirect Results**
("work performance" OR productivity) AND building [Energy, Engineering, Environmental Science]	65,090
("Work performance" OR productivity) AND building AND NOT agriculture AND NOT soil AND NOT "construction productivity" AND NOT machine AND NOT manufacture AND NOT production AND NOT species [Energy, Engineering, Environmental Science]	42,813
("Work performance" OR productivity) AND building AND (perceived OR "self-estimated" OR "self-assessed") AND NOT agriculture AND NOT soil AND NOT "construction productivity" AND NOT machine AND NOT manufacture AND NOT production AND NOT species AND NOT product AND NOT safety [Energy, Engineering, Environmental Science]	5,680
tak ("Work performance" OR productivity) AND building AND (perceived OR "self-estimated" OR "self-assessed") AND NOT agriculture AND NOT soil AND NOT "construction productivity" AND NOT machine AND NOT manufacture AND NOT production AND NOT species AND NOT product AND NOT safety AND NOT residential AND NOT disease {Building and Environment, Energy and Buildings} [Energy, Engineering, Environmental Science]	64

2.2.3 Title and Abstract Screening

After searching all the databases and removing all the duplicate references, the results were further refined through title and abstract screening. The title was screened from 1058 results from all databases to remove any articles that were still not relevant to the search. The majority of studies removed during title screening were studies that focused on other subjects that also mention work performance or productivity (e.g. main focus is on ventilation and energy performance of a ventilation system that could also enhance work performance etc.). After the title screening, 85 results remained to be screened for abstract. The abstract screening removed all search results with research subjects listed below, leaving 44 results.

- Testing a product or a particular system.
- Construction efficiency and construction labourer productivity.
- Performance of product or a building.
- Production of goods and industry workers.
- How symptoms of bad health can affect productivity.
- Productivity in education.
- Productivity in residential buildings.
- The main focus is on perceived control of the environment.
- How furnishing can enhance productivity.
- Customer satisfaction.
- The relationship between people's physical aspects and productivity.

2.2.4 Refinement

All databases were searched with the same search terms, filters applied to subjects and publications, titles and abstracts were screened, leaving 44 results as shown inTable 4. Subject and publication filters were applied to remove studies that are not relevant to the research. The list of subject exclusion and inclusion filters, and publication inclusion filters are included in Section 0, Appendix A. Duplicates were removed at title screening step, during the citation import process. However, there are studies by Clausen and Wyon (2008) and Nishihara et al. (2014) that compare self-reported work performance and measured work performance that this scoping search has not included. The search protocol is tested in the next section.

2.2.5 Protocol Test

Studies by Clausen and Wyon (2008) and Nishihara et al. (2014) both compare the self-reported performance against a quantifiable measure of performance. These two studies have been found through past literature reviews from using a less restrictive search protocol. However, the scoping search did not include these two studies, suggesting that other relevant studies for this research may also have been omitted. The search protocol was modified and tested so as to include these two studies, so that potentially more relevant studies could be included in the scoping search.

Table 4. The process of refinement to find relevant studies from all databases.

Criteria	Fields	Results
Initial Search ("Work performance" OR productivity) AND building AND (perceived OR "self-estimated" OR "self-assessed") AND NOT agriculture AND NOT soil AND NOT water AND NOT construction AND NOT machine AND NOT manufacture AND NOT production AND NOT species AND NOT product AND NOT safety AND NOT land AND NOT residential AND NOT disease	All	350'164
Subject and publication filtering	All	1058
Title screening	Title	85
Abstract screening	Abstract	44

2.2.5.1 Initial search term revision

The initial search term was revised to include the studies by Clausen and Wyon (2008) and Nishihara et al. (2014). Since studies use different terms for work performance, such as productivity, task performance, occupant performance, worker performance and human performance, all these terms were added to the search. Also, since "productivity" searches resulted in large numbers of agriculture-related results, the terms "perceived productivity", "self-assessed productivity" and "self-estimated productivity" were used for the search.

The exclusion criteria excluded articles that are also relevant, so no exclusion criteria were used for the search term. This was because, for example, "AND NOT manufacture" – will delete articles about machine productivity in manufacturing, but will also delete articles that discuss how productivity was measurable when the economy was manufacturing-based but not since the economy has moved to office-based work. The new search term used was:

"work performance" OR "perceived productivity" OR "self-assessed productivity" OR "self-estimated productivity" OR "task performance" OR "occupant performance" OR "worker performance" OR "human performance".

2.2.5.2 Search results

The same process used in the scoping search described in section 2.2.4 was followed to refine the results. The new search has added 20 results, including studies by Clausen and Wyon (2008) and

Nishihara et al. (2014), from ProQuest, ScienceDirect and Taylor and Francis. No new results were found in Sage, Avery Index to Architectural Periodicals, JSTOR and Google Scholar.

2.2.6 Quality Assessment

From the total of 64 sources identified by the search protocol, results were sorted to obtain relevant information from each of the studies. The 64 results were categorised into sources with information to answer the five key points discussed in section 2.1. Table 5shows the 32 studies which remained after sorting through the relevant information that is included in each of the results. The next section discusses the questions from section 2.1, based on the 32 results with other studies found in past literature.

Table 5. Categorised list of literature results.

Author	Self-reported is compared against Objective Measures?	Define work performance /productivity	Define comfort	Environment-al effect on measured performance	Results from laboratory or field study	Environment-al conditions affecting performance
Environmental conditions, T = temperature, RH = relative humidity, L = lighting, DL = daylight, A = acoustics, V = ventilation, CO2 = CO2 levels, P = pollutant loads, EC = environmental control, OT = office type						
Al Horr (2016)	N	N	Y	N	N/A	N/A
Balazova et al. (2008)	Y	N	N	Y	Simulated Office	T+A+OT
Boerstra et al. (2015)	Y	N	N	Y	Laboratory	EC
Byrd and Rasheed (2016)	N	Y	Y	N	N/A	N/A
Choi, Loftness and Aziz (2012)	N	N	Y	N	Field	N/A
Clausen and Wyon (2008)	Y	N	N	Y	Laboratory	T+L+A+DL+OT +P
Clements-Croome (2005)	N	Y	Y	N	N/A	N/A
Haynes (2007)	N	Y	N	N	N/A	N/A
Haynes (2008)	N	Y	N	N	N/A	N/A
Hongisto et al. (2016)	Y	N	N	Y	Laboratory	A
Humphreys and Nicol (2007)	N	Y	N	N	Field	N/A
Kekäläinen et al. (2010)	N	N	Y	N	Field	T
Kosonen and Tan (2004a)	N	N	Y	N	Laboratory	T

(continued)

Table 5. Categorised list of literature results (continued).

Author	Self-reported is compared against Objective Measures?	Define work performance /productivity	Define comfort	Environment-al effect on measured performance	Results from laboratory or field study	Environment-al conditions affecting performance
Environmental conditions, T = temperature, RH = relative humidity, L = lighting, DL = daylight, A = acoustics, V = ventilation, CO2 = CO2 levels, P = pollutant loads, EC = environmental control, OT = office type						
Lan et al. (2009)	N	N	Y	Y	Laboratory	T
Lan and Lian (2009)	Y	Y	N	Y	Laboratory	T
Lan et al. (2014)	N	N	N	Y	Office adapted for experiment	T
Leaman (1990)	N	N	Y	N	N/A	N/A
Leaman and Bordass (1999)	N	Y	Y	N	N/A	N/A
Leyten (2013)	N	N	Y	N	N/A	N/A
Mallawaarachchi (2017)	N	Y	N	N	N/A	N/A
Niemela et al. (2002)	N	N	N	Y	Field	T
Nishihara et al. (2014)	Y	N	N	Y	Laboratory	P
Park and Yoon (2011)	N	N	N	Y	Laboratory	V
Purdey and Leifer (2012)	N	Y	N	N	Field	N/A
Saari and Aalto (2006)	N	Y	N	N	N/A	N/A
Seppanen et al. (2006)	N	N	N	Y	N/A	T, V
Tanabe, Nishihara and Haneda (2007)	N	Y	N	N	Field	N/A
Tanabe et al. (2009)	N	Y	N	Y	Field	T
Tham (2004)	N	N	N	Y	Field	T+V
Tham and Willem (2010)	N	N	N	Y	Simulated Office	T
Vischer (2008)	N	N	Y	N	N/A	N/A
Zhao (2009)	N	N	Y	N	Laboratory	N/A

2.3 Results from Review of Literature – Definitions and Comparisons

2.3.1 Definition of Productivity and Work Performance

2.3.1.1 Common definition

People have defined productivity in offices in different ways, but the definition is vague due to the complexity of the nature of work in offices. The following is a summary of the diversity of definitions identified in these papers:

- Productivity is a ratio of output to input (Byrd and Rasheed, 2016; Humphreys and Nicol, 2007; Leaman and Bordass, 1999; Sink, 1985, cited in Haynes, 2007; Tanabe et al., 2007).
- Is measurable in manufacturing, but the majority of the economy has moved from manufacturing-based work to servicing-based to knowledge-based work (Cascio, 1995, cited in Purdey and Leifer, 2012; Humphreys and Nicol, 2007; Mawson, 2002, cited in Haynes, 2007).
- The input can be defined, but the output is almost impossible to measure (Byrd and Rasheed, 2016; Humphreys and Nicol, 2007; Lan and Lian, 2009; Leaman and Bordass, 1999; Mallawaarachchi, 2017; Tanabe et al., 2009).
- There is no set standard for measuring office productivity (Haynes, 2008; Lan and Lian, 2009).

Because output is complex to measure, productivity itself is also difficult to measure. One of the reasons for this is that the work output is intangible, and the profit or losses may not reflect the amount of work intensity (Byrd and Rasheed, 2016). Also, because office work can consist of different jobs and tasks of great variety, it is difficult to compare the outputs (Mallawaarachchi, 2017). Humphreys and Nicol (2007) and Lan and Lian (2009) have also stated that office work involves many different aspects of skills and sets of skills that affect productivity, such as intellectual, social, analytical, creative and other essential skills (word processing, filing, etc.), but only a few can be quantified. For these reasons, productivity in offices is complex to measure and therefore there are no standards for measuring productivity in an office (Haynes, 2008; Lan and Lian, 2009).

2.3.1.2 Other definitions of productivity

Heerwagen (1998, cited in Clements-Croome, 2005) describes performance as being “motivation × ability × opportunity”. However, this refers to how an individual wanting to do the task should be capable of doing it, and will also require resources and amenities to be available for the task to be accomplished.

Saari and Aalto (2006) mention that “productivity” is the economic expression of human performance. Because this thesis focuses on the how people perform in the built environment and not the organisational or individual output, work performance is a better term to describe how office occupants perform under environmental conditions.

2.3.2 Definition of “Comfort”

2.3.2.1 Optimum conditions for work performance

Although not all the studies have quantifiable evidence or have referenced other studies with quantified results, different studies have varying opinions on how a comfortable work environment affects work performance. Studies by Byrd and Rasheed (2016), Choi et al. (2012), Kekäläinen et al. (2010), Leyten et al. (2013) and Zhao (2009) state that a comfortable environment is essential for optimum or maximum levels of work performance to occur. In contrast, studies by Clements-Croome (2005), Kosonen and Tan (2004a), Lan et al. (2009), Leaman (1990), Oseland (1999, cited in Al Horr et al., 2016) and Pepler and Warner (1968, cited in Leaman and Bordass, 1999) state that optimum or maximum levels of work performance do not necessarily occur in comfortable environments. The different results and opinions for the “optimum conditions” for better work performance could be caused by different definitions of comfort in relation to the different aspects of human needs, which is further discussed in the next section.

2.3.2.2 Levels of comfort for work

Vischer (2008) has separated the comfort required for occupant satisfaction and well-being into three categories: physical comfort, functional comfort and psychological comfort. Vischer (2008) also claimed that people require support from the environment that is suited for their activities and tasks, more than just health and safety from the buildings that they occupy.

- Physical comfort refers to providing for the basic human needs of the occupants, such as for safety, hygiene and accessibility so that the building is inhabitable for people.
- Functional comfort is achieved when the environment supports occupants’ performance of work-related tasks and activities. This also links to psychological aspects of occupants’ environmental preferences with results such as improved team effectiveness and improved work performance.
- Psychological comfort is achieved through having borders and territory and having control over the environment

The environmental effects on occupants and their work performance would also affect their functional comfort. A functionally comfortable environment supports the occupants to conserve their energy and attention for their activities and tasks, whereas an uncomfortable and unsupportive environment could force the occupants to use their energy to adjust to the environmental conditions, having a negative effect on productivity. Functional comfort is derived from environmental standards, with extra additions required for occupants to be comfortable to perform their task; therefore, functional comfort is what is required in office workplaces, and is used to differentiate it from physical comfort. The levels of comfort proposed by Vischer (2008) are illustrated in Figure 1.

This content is unavailable.

Please refer to image by Vischer (2008).

Figure 1 Levels of comfort with the lowest level being discomfort; a satisfactory environment will provide physical comfort, an improved environment will provide functional comfort and the thought of having territories and control over the environment will provide psychological comfort.

2.3.3 Comparing Self-Reported and Objective Measure of Work Performance

The database search found six studies that have compared self-reported and measurable work performance, including studies by Balazova et al. (2008), Boerstra et al. (2015), Clausen and Wyon (2008), Hongisto et al. (2016), Lan and Lian (2009) and Nishihara et al. (2014), shown in Table 6.

Of the six studies, the studies by Clausen and Wyon (2008) and Lan and Lian (2009) show results where changes in self-reported performance do not correspond to the objective measure of work performance.

The experiment by Clausen and Wyon (2008) showed trends between self-reported and simulated office work performance in the same direction, that is, an improved environment increases work performance for some conditions. However, it challenged expectations when the occupants reported their performance to be increasing by 20% when the simulated office work performance decreased by 1%.

Similarly, in the study by Lan and Lian (2009), the cognitive tests that had the highest results for correct percentage answer were for those working at 17°C, when the occupants reported that 17°C required the most effort to complete the tasks. Also, the 17°C working temperature produced a greater number of cognitive tests with highest results than did a temperature of 27°C for both the percentage of correct answers and response times, when occupants had reported that it took less effort to work in 27°C than 17°C.

All six studies except Lan and Lian (2009) show results where self-reported and measured performance are positively correlated. However, effect size differs between self-reported and measured performance in some studies.

The study by Balazova et al. (2008), measuring the simulated office work in relation to changes in combination of the work environment (temperature, office type and sound) summarised in Table 6, showed that both self-reported and measured performance follow a similar trend overall, with potential outliers in condition D (open plan with replicated office noise and long reverberation time at 23°C) for text-typing and condition C (open plan with replicated office noise but with added sound baffles at 23°C) for proofreading, not reflected in the self-reported performance. The change in self-reported work performance of 12.7% between conditions A (quiet cellular office at 23°C) and B (open plan office with replicated office noise with usual average reverberation at 23°C) shows a difference to text typing (2.6%) and proofreading (18.7%). Studies by Boerstra et al. (2015) and Clausen and Wyon (2008) also showed that self-reported and measured performance values differed, especially the results from Clausen and Wyon (2008) where a 7% increase in measured performance correlated with a 25% increase in self-reported performance.

Table 6 List of studies comparing self-reported and measured performance, with results.

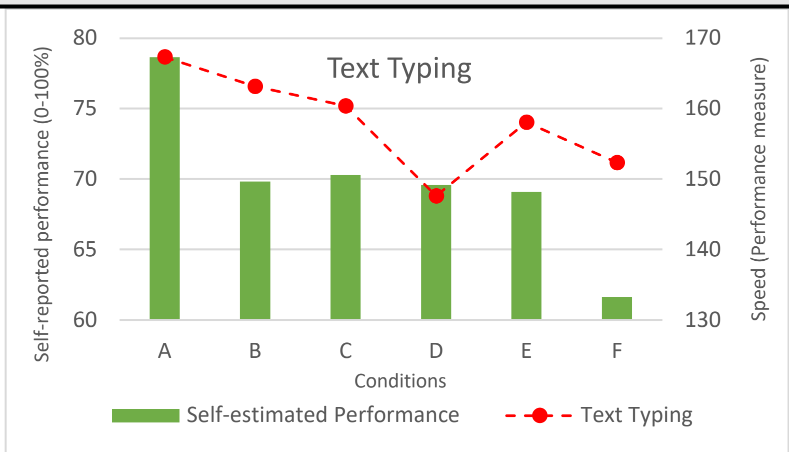
<table>
<tr><th>Author</th><th>Measured Work Performance/ Environmental Conditions Tested</th><th>Results</th></tr>
<tr><td>Balazova et al. (2008)</td><td>Simulated office work
Combination of temperature, office type, sound</td><td>


Figure 2 Self-reported VS text-typing (Balazova et al., 2008).

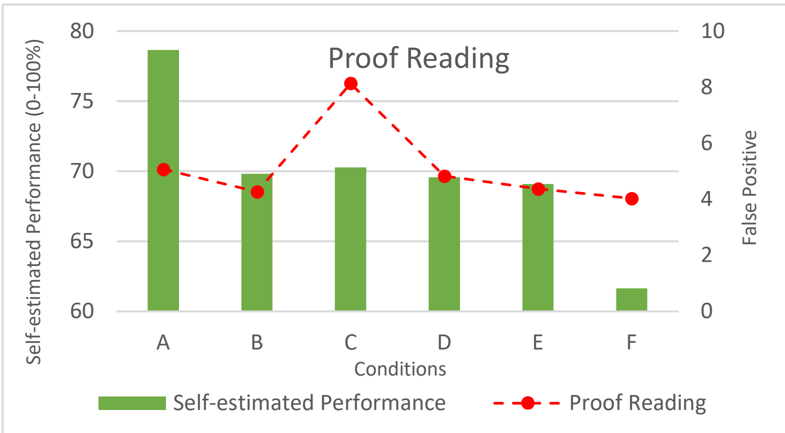
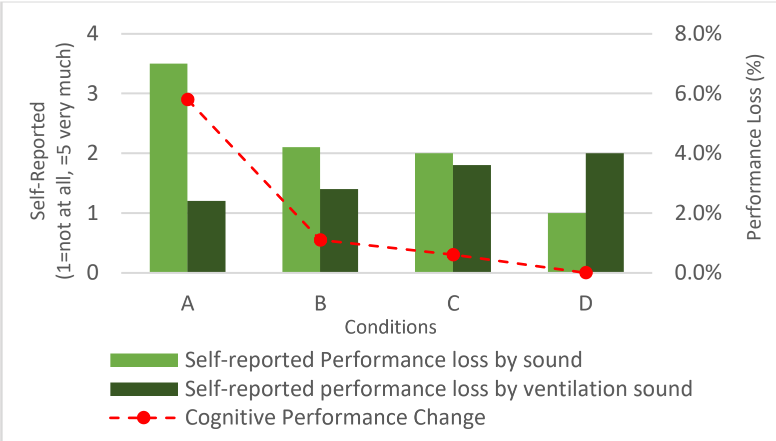
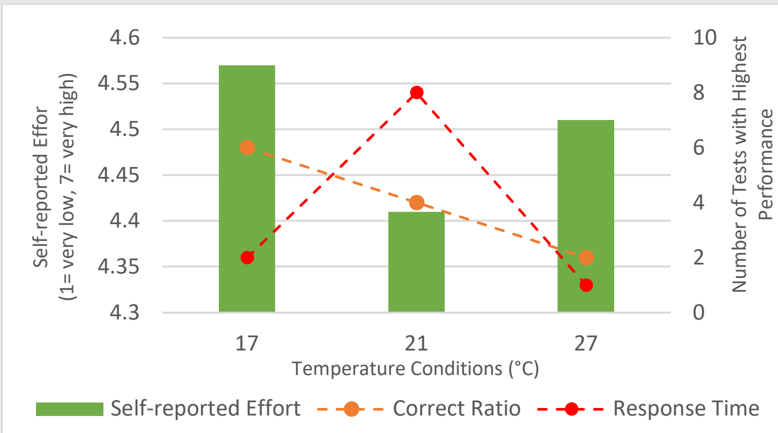


Figure 3 Self-reported VS proofreading (Balazova et al., 2008).
</td></tr>
<tr><td>Boerstra et al. (2015)</td><td>Cognitive
Environment with control/no control</td><td>▲ 4.2% self-reported
▲ 10.4% addition
▲ 8.2% multiplication
▲ 5.4% character typing</td></tr>
</table>

(continued)

Table 6 List of studies comparing self-reported and measured performance, with results (continued).

Author	Measured Work Performance/ Environmental Conditions Tested	Results
Clausen and Wyon (2008)	Simulated office work Combination of noise, lighting, daylight, office layout, air temperature, and air quality	▲ 20–25% self-reported VS ▲ up to 7% simulated office work ▲ 20% self-reported VS ▼ 1% simulated office work
Hongisto et al. (2016)	Cognitive Combination of speech and ventilation sound	 *Figure 4 Self-reported loss by sound vs cognitive performance (Hongisto et al., 2016).*
Lan and Lian (2009)	Cognitive Temperature	 *Figure 5 Self-reported effort vs number of tests with highest performance (Lan and Lian, 2009)*
Nishihara et al. (2014)	Cognitive Pollution loads	No significant changes in both self-reported and cognitive performance

2.3.3.1 What is considered a significant difference in percentage of work performance?

Both self-reported and measured performance show the same trend: each increases as environmental conditions improve, and vice versa. The studies show a clear correlation between self-reported work performance and objective measures of work performance where an improved environment equates to improved work performance. However, an issue still remains in the difference between self-reported work performance and actual work performance. What should be considered a significant difference between the two measures to examine how closely self-reported work performance reflects an objective measure of work performance?

A number of different justifications could be used to determine how much of a difference is considered to be significant. Using the ratio of 1:5:200 (discussed in section 1.1, cost ratios for business in 20 years for initial construction costs:running and maintenance costs:cost of salaries and wages for workers) as an example, 2.5% could be considered accurate, because 2.5% of 200 (cost of salaries) is equal to 5 (maintenance and operational costs), or it could be argued that 0.5% is considered accurate, because 0.5% of 200 is equal to 1 (initial construction cost). For the ratio of 1:8:80, the accuracy could be considered at 10% and 1.25%.

Another justification is to compare the possible savings through an increase in work performance, with the extra costs to upgrade a conventional office building to an environmentally sustainable design (ESD) office building and through estimations of increase in profit margins for different industries. Justifying a significant difference using cost benefits from ESD buildings and the profit margin is discussed in the next section.

2.3.3.2 Using annual salaries and extra costs for upgrading a building to have ESD features, to justify how much percentage change in work performance is significant

To justify how much of people's work performance is significant for businesses, one way is to compare the average salaries of workers against the extra costs added to office building capital costs for the (ESD) features added into green buildings. Table 7 shows costs per workstation of 17.2 m^2 (Colliers International, 2016) based on building capital costs for conventional office buildings, ESD buildings, potential savings through energy and water, and the simple payback period for the energy and water savings from Fullbrook and Jackson (2005).

Table 7 Costs per workstation to upgrade an office to ESD buildings with potential savings (Colliers International, 2016; Fullbrook and Jackson, 2005).

Building Costs and Savings $ / Workstation ($ / 17.2m²)	Benchmark building capital cost	ESD building capital cost	$ difference	$ difference in %	Annual Savings (Energy)	Annual Savings (Water)	Total Annual Savings (Energy and Water)	Simple Payback Years
Office: Low-Medium ESD	34400	36636	2236	6.5%	189.2	5.16	194.36	11.50
Office: Medium-High ESD	34400	38356	3956	11.5%	292.4	10.32	302.72	13.07

Table 8 shows the average salary for New Zealand office workers, calculated using reported averages from recruitment and job search websites (Hudson, 2015; Payscale, 2017; Trademe, 2017), and the different percentages of the average annual salary with accumulated totals for a period of 50 years. This could be considered as a way of calculating savings for businesses as a 10% increase in work performance could mean that workers finish 10% more work in the same amount of time, or finish the same amount work but 10% faster. From the salaries and extra costs for ESD offices, the payback period is calculated in Table 9. The payback period shows that even with just 1% increase in work performance, the payback period is less than 10 years, and decreases greatly for a 5% increase in work performance.

Table 8 Average New Zealand office workers' salaries with percentages of the annual salaries and accumulated totals based on salary information from Hudson (2015), Payscale (2017) and Trademe (2017).

		Salary ($)	1%	2%	5%	10%	15%	20%	40%
Years	**1**	45500	455	910	2,280	4,550	6,830	9,100	18,200
	5	227,500	2,280	4,550	11,400	22,800	34,100	45,500	91,000
	10	455,000	4,550	9,100	22,800	45,500	68,300	91,000	182,000
	20	910,000	9,100	18,200	45,500	91,000	137,000	182,000	364,000
	25	1,140,000	11,400	22,800	56,900	114,000	171,000	228,000	455,000
	50	2,280,000	22,800	45,500	114,000	228,000	341,000	455,000	910,000

Table 9 Payback period for low ESD and high ESD buildings considering an increase in work performance savings only and considering all other potential savings, based on information from Tables 7 and 8.

Payback Period (years)		1%	2%	5%	10%	15%	20%	40%
Increase in work performance savings	**Low ESD**	4.9	2.5	1.0	0.5	0.3	0.3	0.1
	High ESD	8.7	4.4	1.7	0.9	0.6	0.4	0.2
Increase in work performance + energy + water savings	**Low ESD**	3.4	2.0	0.9	0.5	0.3	0.2	0.1
	High ESD	5.2	3.3	1.5	0.8	0.6	0.4	0.2

From the savings for the average salary for different percentage increases in work performance in Table 9 Payback period for low ESD and high ESD buildings considering an increase in work performance savings only and considering all other potential savings, based on information from Tables 7 and 8. Table 8 and the extra costs for ESD offices inTable 7, the total potential savings per worker (or per workstation) for the different percentage increases in work performance for up to 50 years are shown in Table 10. Calculations of the potential savings per worker were based on average New Zealand salary levels (Hudson, 2015; Payscale, 2017; Trademe, 2017) and the costs to upgrade an office building to ESD offices per workstation (per workstation is calculated as 17.2 m^2).

A 5% increase in work performance could save the business around NZ$43,300 per worker in 20 years, which is almost equivalent to the average annual salary's worth of savings for the business in 20 years. This could also be considered in terms of the number of extra workers' worth of work being completed by a certain number of people. For example, 1% increase would be roughly equivalent to 1 extra worker per 105 workers per year, 2% increase will be 1 extra worker per 52 workers per year and 5% will be 1 extra worker for 21 workers per year. Original figures for extra costs, savings for business in terms of percentage of average annual salary, and potential savings with energy and water savings are in Appendix B.

Table 10 Savings through increase in work performance only (not including potential savings from energy and water), considering extra costs per workstation for ESD features based on information collated in Tables 7 and 8.

	Years	1%	2%	5%	10%	15%	20%	40%
Low ESD	1	−1,780	−1,330	39	2,310	4,590	6,864	16,000
	5	39	2,310	9,140	20,500	31,900	43,264	88,800
	10	2,310	6,860	20,500	43,300	66,000	88,764	180,000
	20	6,860	16,000	43,300	88,800	134,000	179,764	362,000
	25	9,140	20,500	54,600	112,000	168,000	225,264	453,000
	50	20,500	43,300	112,000	225,000	339,000	452,764	908,000
High ESD	1	−3,500	−3,050	−1,680	594	2,870	5,140	14,200
	5	−1,680	594	7,420	18,800	30,200	41,500	87,000
	10	594	5,140	18,800	41,500	64,300	87,000	178,000
	20	5,140	14,200	41,500	87,000	1,330	178,000	360,000
	25	7,420	18,800	52,900	110,000	167,000	223,500	451,000
	50	18,800	41,500	110,000	224,000	337,000	451,000	906,000

The payback period and potential savings for the business per person due to increase in work performance (and increase in salary) of 5% is comparable to profit margins for most profitable industries. This allows an increase in work performance to be compared to profit margins, to visualise how an increase in work performance of 5% could be as significant as cost savings for companies. Also, comparisons with industries with high profit margins will show the minimum achievable increase in salary compared to profit margins. This is because 5% will be a significant difference for industries with low profit margins (example: for an industry with 5% profit margin, a 5% increase in profit equates to 100% of their profit margin) and will not be as significant for industries with high profit margins (example: for an industry with 20% profit margin, a 5% increase in profit equates to 25% of their profit margin).

Percentages for a potential increase in work performance could also be compared with the profit margins of office work industries. Biery (2016) reported the top 15 most profitable industries in 2016 with profit margins for each industry. For the industries related to office work, their profit margins are ranked thus: accounting, tax preparation, bookkeeping, payroll service with 18.3%; legal services with 17.4%; offices of real estate agents and brokers with 14.8%; offices of other health practitioners 14.2%; offices of dentists with 14.1%; specialised design services with 12.8%; and offices of physicians with 11.5%. Based on these profit margins, an increase of 5% in salaries equates to an increase in salary of 27.3% of their profit margin (for accounting, tax preparation, bookkeeping, payroll service of 18.3%) and could potentially be more for industries with lower profit margins.

2.3.4 Effect of Number of Environmental Stressors on Objective Measure of Work Performance

To determine how the self-reported and objective measures of work performance correlate with each other, one method is to observe the trend in how these two parameters change under different environmental conditions. Therefore, a minimum of two different environmental conditions will be required to create a trend and to observe the change. To stress the subject enough to change their performance for it to be observable, one environmental stressor could be sufficient, or it could require several stressors. Studies of environmental effects on measured performance, with a focus on a number of stressors and sizes of change in measure performance, have been summarised in Table 11.

The studies summarised in Table 11 show changes in measured performance to be under 15% for most cases, and few around 20%. Leaman and Bordass (1999) estimated that decreases or increases in turnover in typical office organisations of up to 15% could be affected by the design, management and use of the indoor environment, but they also mention that self-reported work performance shows differences of up to 25% between comfortable and uncomfortable states, with comfortable occupants showing consistently lower performance. Brill (1986, cited in Leaman and Bordass, 1999) and Vischer (1989, cited in Leaman and Bordass, 1999) also used an approximate figure of 15% as an increase or decrease in performance in their studies, and Lorsch and Abdou (1994, cited in Leaman and Bordass, 1999) discussed how improving the air quality increased work performance by 20%.

Table 11. List of studies with different environmental conditions and effect on measured performance.

Number of Environmental Stressors	Author	Type of Measured Performance/ Type of Environmental Stressor	Results
One	Boerstra et al. (2015)	Cognitive Environmental control / no control	From subjects having control over environmental conditions to having no control. Countering expectations, three objective measures of work performance show increase in performance for subjects with no control over their environmental conditions. ▲ 10.4% addition. ▲ 8.2% multiplication. ▲ 5.4% character typing.
	Hongisto et al. (2016)	Cognitive Sound	Experiment to examine how sound insulation and noise affect cognitive performance and acoustic satisfaction based on insulation guidelines. Serial recall and accuracy of mental arithmetic tasks are used for cognitive performance. To observe the greatest change in cognitive performance, the conditions with the lowest performance and the highest performance are compared. From Condition D (Sound Pressure Level of speech (SPL) 22dB, SPL of masking 42dB, SPL overall 42dB, Speech Transmission Index (STI) of 0.00) to Condition A (SPL of speech 30dB, SPL of masking 33dB, SPL overall 35dB, STI of 0.38). ▼ 8.1% for correct response % in series recall. From Condition A to Condition C (SPL speech 30dB, SPL of masking 42dB, SPL overall 42). ▼ 13.7% for error rate % in mental arithmetic task.
	Lan et al. (2009)	Cognitive Temperature	Tested four temperature conditions, 19 °C, 24 °C, 27 °C, and 32 °C. From best temperature to worst temperature condition. Changes are shown in decrease in reaction time (temperature with highest result − temperature with lowest result)/Decrease in accuracy (temperature with highest result − temperature with lowest result) To observe the greatest changes in cognitive performance, the differences between conditions with the lowest performance and the highest performance are compared. Letter Search 10.0 % (24 °C – 19 °C) / 1.7 % (24 °C – 32 °C) Overlapping 8.9 % (32 °C – 19 °C) / 0.2 % (27°C – 24 °C) Number Calculation 14.0 % (32 °C – 19 °C) / 4.3 % (24 °C – 27 °C) Conditional Reasoning 5.7 % (24 °C – 19 °C) / 25.0 % (24 °C – 19 °C) Spatial Image 12.8 % (24 °C – 19 °C) / 15.5 % (27 °C – 19 °C) Memory Span 1.8 % (24 °C – 32 °C) / No time recorded for memory span Picture Recognition 5.8 % (27 °C - 19 °C) / 5.5 % (24 °C – 27 °C) Symbol Digit Modalities Test 7.3 % (32 °C – 27 °C) / 5.8 % (24 °C – 32 °C) Visual Choice Reaction Time 1.3 % (32 °C – 19 °C) / 0.7 % (27 °C – 32 °C)

(continued)

Table 11. List of studies with different environmental conditions and effect on measured performance (continued).

Number of Environmental Stressors	Author	Type of Measured Performance/ Type of Environmental Stressor	Results
One	Lan, Wargocki and Lian (2014)	Cognitive and Simulated Office Work / Temperature	Study develops a method to integrate speed and accuracy into one metric of human performance. Subjects perform cognitive performance tests. The results show test result changes from 30 °C to 22 °C. ▼ 2% Text typing. ▲ 12% Addition. ▲ 10% troop. ▲ 5% Calculation.
	Lan and Lian (2010)	Cognitive / Temperature	Experiment carried out in field laboratory to investigate effect of indoor air temperature on productivity. Tested three temperature conditions, 17 °C, 21 °C, 28 °C. Changes are shown in decrease in reaction time (temperature with highest result – temperature with lowest result) / decrease in accuracy (temperature with highest result – temperature with lowest result) Letter Search 1.7% (21°C - 17°C) / 1.0% (17°C - 21°C) Overlapping 5.6% (21°C - 28°C) / 1.7% (21°C - 17°C) Memory Span 3.6% (17°C - 28°C) / No time recorded for memory span Picture Recognition 3.7% (21°C - 17°C) / 3.1% (17°C - 28°C) Number Calculation 5.8% (17°C - 28°C) / 0.4% (17°C – 21°C &28°C) Event Sequence 6.5% (21°C - 28°C) / 3.3% (17°C - 28°C) Conditional Reasoning 3.9% (17°C- 28°C) / 9.3% (28°C - 17°C) Reading Comprehension 8.7% (28°C - 17°C) / 6.4% (21°C - 17°C) Spatial Image 17.2% (21°C - 28°C) / 6.4% (21°C - 28°C) Graphic Abstracting 19.4% (21°C - 28°C) / 6.3% (28°C - 17°C) Visual Choice Reaction Time 4.0% (21°C - 28°C) / 1.1% (17°C - 28°C) Hand-eye Coordination 2.2% (21°C - 17°C) / 4.0% (21°C - 28°C)
	Niemela et al. (2002)	Actual performance (call centre) / Temperature	Investigates effect of air temperature on labour productivity in telecommunication offices. From 21.9 °C to 28.5 °C ▼ 11.9% (▼ 1.8% per ▲ 1 °C)

(continued)

Table11. List of studies with different environmental conditions and effect on measured performance (continued).

Number of Environmental Stressors	Author	Type of Measured Performance/ Type of Environmental Stressor	Results
One	Park and Yoon (2011)	Cognitive Ventilation	Evaluates the effect of ventilation rate on work performance. From 20 L s^{-1}person^{-1} to 5 L s^{-1}person^{-1} ▼ 5.1% addition ▼ 5.3% text typing ▼ 6.5% memorisation
	Seppanen and Fisk (2006)	Literature Review Temperature, Ventilation (separately)	Relative performance in correlation with temperature from 24 studies, where maximum performance is set at 1. 21–22°C is optimum temperature, and performance decreases as temperature decreases or increases from the 21–22°C range. ▼ 8–10% at 15 °C ▼ 12–25% at 35 °C Change in performance per 10 L s^{-1}person^{-1}. ▲ 2% at 10L/s/person 0% at 50L/s/person
	Tanabe et al. (2009)	Actual performance (call centre) Actual performance (call centre) for Temperature	Call response rate in correlation with air temperature in telecommunications. Increase in air temperature from 25°C to 26°C ▼ 1.9% (▼ 1.9% per ▲ 1°C)
	Valančius and Jurelionis (2013)	Simulated Office Work Temperature	Change in temperature from 22°C (set as optimum temperature). Remain at 22°C ▼ 0.2% text typing ▲ 1.0% arithmetic tasks ▲ 6.5% Tsai-Partington test ▲ 2.1% overall To 18°C ▼ 0.2% text typing ▲ 5.7% arithmetic tasks ▲ 12.9% Tsai-Partington test ▲ 5.4% overall To 26°C ▼ 0.2% text typing ▼ 2.3% arithmetic tasks ▲ 1.2% ▼ 0.6% overall

Table 11. List of studies with different environmental conditions and effect on measured performance (continued).

<table>
<tr><th>Number of Environmental Stressors</th><th>Author</th><th>Type of Measured Performance/ Type of Environmental Stressor</th><th>Results</th></tr>
<tr><td>One</td><td>Tham and Willem (2010)</td><td>Cognitive
Temperature</td><td>From 20°C to 26°C
▲ 42% Accuracy
▲ 7% Speed</td></tr>
<tr><td rowspan="3">Multiple</td><td>Balazova et al. (2008)</td><td>Simulated Office Work
Office type + temperature + sound</td><td>An experiment was conducted to examine the effect of office noise and temperature on human perception, comfort and office work performance.
Both temperature and noise had negative effects on tasks.
From conditions with best results to worst results.
From 23°C open office III (sound absorbent system, 49dB(A)) to 23°C cellular office (no sound played, 47.7dB(A))
▼ 13.4% test typing speed
From 28°C open office I (replicate of real office, 52dB(A)) to 23°C open office II (office with longer reverberation time, 54dB(A))
▼ 50.4% proof reading, detecting errors</td></tr>
<tr><td>Clausen and Wyon (2008)</td><td>Simulated Office Work
Temperature + sound + lighting + daylight + office layout + pollutants for linoleum</td><td>Conducted an experiment with different improvements to office indoor environments, with a set budget to improve indoor office quality. Indoor environmental condition to improve were surveyed by occupants, and test was conducted at 0% budget (standard, no changes to office), office environment with budget of 50% with no occupants' choice added in environment improvements, 50% with choice of improvements, and 100% budget. Simulated office work tests were conducted at different exposure times.

At 100% budget, 114minutes of exposure
▲ up to 7% simulated office work

At 50% budget with no choice, 114minutes of exposure
▼ 1% simulated office work</td></tr>
<tr><td>Tanabe et al. (2015)</td><td>Simulated Office Work</td><td><table>
<tr><th></th><th>25.5°C - Suit</th><th>28.5°C - Suit</th><th>28.5°C - Light clothes</th><th>28.5°C – Desk fan</th><th>28.5°C - All</th></tr>
<tr><td>Multiplication</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Accuracy</td><td>92%</td><td>92%</td><td>89%</td><td>92%</td><td>91%</td></tr>
<tr><td>Number of correct answers (answers/hr)</td><td>124</td><td>116</td><td>118</td><td>123</td><td>124</td></tr>
</table></td></tr>
</table>

(continued)

Table 11. List of studies with different environmental conditions and effect on measured performance (continued).

Number of Environmental Stressors	Author	Type of Measured Performance/ Type of Environmental Stressor	Results					
Multiple			**Proof reading**					
		Temperature, clothes, individual environmental control, Temperature + clothes + individual environmental control	Accuracy	74%	73%	72%	71%	74%
			Number of errors detected (detected/hr)	185	177	175	185	185
			Reading Speed (characters /hr)	23’384	22’482	22’785	23’836	23’269
			Creativity Thinking					
			Number of Ideas	28	30	27	28	28
	Tham (2004)	Actual performance (call centre) / Temperature and Ventilation	From 5L/s/person to 10L/s/person at 22.5°C ▼ 3.2% mean talk time From 5L/s/person to 10L/s/person at 24.5°C ▲ 17.7% mean talk time					

The studies show that there is no obvious evidence that multiple environmental condition stressors affect people’s work performance more than just applying one environmental stressor, with most changes in performance being below 10%, with some around 15% and a limited number with a change in performance of over 20%.

2.3.5 Laboratory vs Field Studies

Of the 32 studies in section 2.2.6, there is roughly the same number of experiments conducted in a laboratory environment as in a work environment. Laboratory experiments make the environmental conditions easier to control, therefore reducing the likelihood of other environmental conditions affecting people within the chamber. However, since it can be argued that the people are not in their comfortable usual working space, it is possible that the results are not directly applicable to how people are affected in their workspace by the same environmental condition. In contrast to the situation in laboratory studies, environmental conditions are harder to control in work environments for field studies. Thus, while it could be argued that the results from field experiments could be more applicable to real life situations as the people are in their usual working spaces, there are often also too many other factors in field experiments that could influence the environmental effects on work performance.

2.4 Methods of Gathering and Measuring Work Performance and Self-Reported Work Performance

2.4.1 Different Types of Self-Reported Work Performance Questions

Past studies examining the accuracy of self-reported work performance against an objective work performance measure used the direct approach in surveys to ask the building occupants to assess their work performance. This has been the major and common question used in past research and studies (Sullivan, Baird and Donn, 2013).

The other common question used in post-occupancy surveys is the environmental effect approach of asking the occupants to assess the effect of environmental conditions on their work performance. Surveys by Building Use Studies (BUS) and the Occupant Indoor Environmental Quality survey by the Center for the Built Environment (CBE) are major survey methods using the environmental effect approach.

There are arguments that state the difficulty for theparticipants to assess how their work performance is being affected by the environmental conditions (Sullivan, Baird and Donn, 2013), but the study by Clausen and Wyon (2008) also made the accuracy of the direct approach questionable.

2.4.1.1 Direct approach to assess work performance

Survey questions could simply ask participants to assess their own performance. An example where this has been used is in a survey by Toftum et al. (2010), where participants indicated their agreement with the statement “Right now, I am able to work on:” with a 0% to 100% line scale. This is a common method to obtain a measure of work performance, and has been used by Clements-Croome and Kaluarachchi (2000, cited in Sullivan et al. 2013), Clausen and Wyon (2008), Kekäläinen et al. (2010) and Cui et al. (2013).

Clausen and Wyon (2008) investigated the effect of temperature, air quality, noise and light on work performance in poor and improved environmental conditions. An imaginary “budget” was set up as a cost estimate for renovations and additions, judged by the users of the work environment, to improve the work environment, and the improved conditions were described as a work environment with the achievable improvement using 50% and 100% of this imagined budget. The investigation was completed by measuring simulated office work in comparison to direct approach self-reported work performance. The improvements made to the environment resulted in a decreased percentage of people dissatisfied, which increased the self-assessed work performance in all three improved conditions by 20–25%. However, the objective measurement of simulated office tasks results was lower, ranging slightly below and above 5% for two improved conditions and −1% for one improved condition. The comparisons suggest that direct approach questions could overestimate people's work performance. It is also potentially misleading, as people have assessed their performance to have increased by 20%, whereas the simulated office tasks results showed a 1% decrease in measured performance.

2.4.1.2 Assessing the effect of environmental conditions on work performance

The major surveys asking this type of questions include the Building Use Studies (BUS) and the Center for the Built Environment (CBE) surveys. The approach has also been used in studies by Frontczak (2011), Tanabe et al. (2015), Agha-Hossein et al. (2013), Humphreys and Nicol (2007) and Kim et al. (2016). The most commonly used BUS survey in New Zealand asks, "Please estimate how you think your productivity at work is increased or decreased by the environmental conditions in the building?" (Leaman and Bordass, 1999).

Sullivan, Baird and Donn (2013) suggest that assessing the environmental effect on work performance could be better than just assessing personal work performance. Ego causes people to want to have a positive self-image, making it more likely that they will externalise the cause of poor performance by blaming it on something or someone else (Hacker et al., 2000, cited in Sullivan, Baird and Donn, 2013; Ryvkin et al., 2012, cited in Sullivan, Baird and Donn, 2013). This implies that the environmental effect approach allows the occupants to externalise any problems, which could reduce bias in their responses (Sullivan, Baird and Donn, 2013). Another potential advantage of the environmental effect approach is due to people's poor metacognition. Kruger and Dunning (1999, cited in Sullivan, Baird and Donn, 2013) stated that skills required to critically assess their performance are the same as those required to perform well. Therefore, the environmental effect avoids people from critically assessing their work and just requires people to be able to assess how the environment is affecting them (Sullivan, Baird and Donn, 2013).

In support of this environmental effect approach, several studies have shown that both perceived work performance and objective work performance follow the same trend under the same environmental conditions (Balazova et al, 2008; Boerstra et al., 2015; Clausen and Wyon, 2008; Hongisto et al., 2016). However, it could be difficult for the subjects to assess how the environment is actually affecting their work performance. Leaman and Bordass (1999) also mention the difficulty for the building occupants to judge their work performance based on another reference point, such as what building or environment they should compare it with when judging how their current building or environment is having an effect on their work performance. Overall, since it is uncertain if any of the building occupants are aware of their environment and the effects it has on their work performance, it could be better to simply use the direct approach.

2.4.1.3 Other question types

Other types of questions that could be asked include asking the occupants to assess their work performance against others. An example of this is in the Health and Work Performance questionnaire by the World Health Organization, which asks "How often was your performance higher than most workers on your job?" A questionnaire used by Tanabe et al. (2013) asks the occupants to assess their work performance against their work performance the previous year, and de Vries et al. (2013) ask people to compare their performance with their lifetime best performance. Vaguely defined reference points could also be a problem for questions asking people to assess their performance in comparison to different periods. What people perceived as a satisfactory level, and when the top work happened could be different for different people (Sullivan, Baird and Donn, 2013).

There are also many problems with the questions asking people to assess their performance against other people. Burson et al. (2006) observed that the miscalibration in assessing relative work performance to other workers was mostly found in poor performance because of people's tendency to overestimate their skills and abilities. Burson et al. (2006) also noted that people with better performance tend to underestimate their performance slightly. As mentioned in the previous section, Kruger and Dunning (1999, cited in Sullivan, Baird and Donn, 2013) also suggested that there is a relationship between metacognition, work performance and judgemental accuracy. However, Burson et al. (2006) claimed that everyone makes errors when judging their relative performance, because individuals who are skilled at a certain task and an individual who is unskilled are similar in terms of skills to judge their relative performance, and people tend to think they are above average for tasks that seem easy and think they are below average for tasks that seem difficult (Kruger, 1999, cited in Burson et al., 2006).

2.4.2 Self-Reported Work Performance Scales

Work performance rating scales can be numerical or ordinal, depending on what is to be measured. Numerical scales show changes in work performance in percentage values whereas the ordinal scales show less precise changes in work performance. Numerical scales provide precise values and ordinal scales measure in less precise values of change in magnitude.

2.4.2.1 Ordinal Scales

Although used in surveys to measure the effect on work performance, ordinal scales are used differently from numerical scales. Ordinal scales could be in 5-point scales (Kekäläinen et al., 2010; Hongisto et al., 2016), 7-point scales (Lan and Lian, 2009) and some up to 11 points (Lamb and Kwok, 2015). The results from these scales provide less precise values on how small or large the changes in work performance were.

The main problem with the ordinal scale is that it does not provide a specific quantified measure of the effect on work performance, which is what is truly desired from the post-occupancy surveys that measure self-reported work performance. The lack of precision in the scale does avoid the misleading suggestion that a 10% change in self-reported work performance is a 10% change in actual work performance. However, because the scale points are not well defined, a change of 1 point could be interpreted differently by different people. Also, the scale points might not be equally spaced (Sullivan, Baird and Donn, 2013). Using the scale by Hongisto et al. (2016) as an example, it is questionable if the work performance difference between *3 – not well, nor badly* and *4 – quite well*, is the same as the difference between *4 – quite well* and *5 – very well*. Since different people could have different ideas and interpretations, this could influence the results and the data analysis (Marincic, 2011).

2.4.2.2 Numerical Scales

Surveys with numerical scales provide precise values and make it easier than with ordinal scales to obtain magnitudes of change in work performance. A common numerical scale in surveys ranges from 0 to 100% (Balazova et al., 2008; Clausen and Wyon, 2008; Tanabe, Haneda and Nishihara, 2015; Toftum et al., 2010). Another common approach is to use numerical scales that have both negative and positive figures of the same size. Scale sizes could range from the ±20% scale used by the CBE surveys (Centre for the Built Environment, 2013, cited in Sullivan, Baird and Donn, 2013), to the range of ±30% (Boerstra et al., 2015) and ±40% (Leaman and Bordass, 1999), and up to ±80% (Tanabe et al., 2013).

Since quantifying work performance is very complex, scales for occupants to report their work performance have not been validated against objective measurements of work performance. Therefore, it would be misleading to say that a 10% change in self-reported work performance is a 10% change in actual work performance. In addition, it could be argued that what can really be read from this scale is whether the people assessed their work performance to have a small change or a large change, which is basically the same as the ordinal scale. However, the scale points are more clearly defined with percentage values. Although the numerical scale might not be accurate, people could understand that +20% work performance is double of +10% work performance. The counter to this is that people do not have any method of estimating 10% of their work performance (Sullivan, Baird and Donn, 2013). If office work is too complicated to numerically measure or estimate, it could be argued that anyone working in a not-easily-quantifiable job will not have any idea of what 10% of their work performance is and they do not have any way to measure or estimate it. Therefore, the self-reported percentage changes in work performance should only be taken as an estimation of how the built environment had a small or a large effect (Sullivan, Baird and Donn, 2013).

2.4.3 Indicators of Work Performance

Since measuring work performance is complex, various aspects of work performance are measured to provide indications of how work performance has changed (Sullivan, Baird and Donn, 2013). Some common indicators that are used as a measure, or have shown correlation with the environmental effect on people's work performance include:

- Motivation.
- Job satisfaction.
- Absenteeism.
- Health-based measures such as measuring symptoms of sick building syndrome (SBS).
- Objective measures such as cognitive performance.

Correlations between indicators and work performance, as well as correlations with other self-reported indicators are observed in a number of studies. Jahncke and Halin (2012) observed that open-plan offices could have negative effects on work performance, fatigue and motivation. Studies by Lorsch and Ossama (1994, cited in Cui et al., 2013) and Lan et al. (2012, cited in Kershaw and Lash) noted that motivated people could maintain and improve work performance, but exerting more effort could result in lower motivation and cause fatigue. Negative effects on SBS symptoms such as fatigue may also negatively affect work performance (Kosonen and Tan, 2004; Tanabe, Nishihara and Haneda, 2007, cited in Boerstra et al., 2015; Tham and Willem, 2010; Wargocki et al., 1999; Vischer, 2008).

There are also studies that mention the positive relationship between job satisfaction and work performance (Al Horr et al., 2016; Choi, Loftness and Aziz, 2012; Clausen and Wyon, 2008; Kosonen and Tan, 2004b; Thomas, 2010; Vischer, 2008; Loftness et al., 2009; and Roelofsen, 2002). Work environments with better or improved indoor environmental quality are reported as reducing the number of complaints, staff turnover, and absenteeism, and enhancing work performance and satisfaction (Clements-Croome, 2000; McGraw Hill Construction, 2014; Tse and So, 2007). Tham (2016) proposes that indicators could be used for economic analysis when negative effects such as time lost through absenteeism and affected learning shows the need to improve the indoor environmental quality to maintain certain levels of attendance and learning.

Despite the correlation between actual work performance and other self-reported indicators of work performance, measures of change in work performance could be heavily influenced by questions and scales on the surveys.

2.5 Questions and Scales in Surveys Could Affect How People Assess their Work Performance

To measure the environmental effect on work performance, there is a consensus that obtaining self-reported work performance measures from surveys are a fast, easy and cheap method of collecting data. Although there is less consensus about the accuracy of the measures, the six studies that compare self-reported work performance against an objective measure of work performance mostly show the same trends.

Therefore, the focus of this research was narrowed to an aspect of the survey that could have an influence on the accuracy of work performance measures. Most studies focusing on the self-reported work performance, including the six studies that compared self-reported work performance with objective measures of work performance, have been conducted using the direct approach to obtain self-reported work performance. As already mentioned in section 2.4.1, although there are studies that argue the advantage of the environmental approach over the direct approach, there are currently limited studies studying how people assess themselves using the environmental approach.

As noted earlier, there are currently two major surveys in wide use internationally: the surveys by BUS and the CBE (Dykes, 2012). These two surveys both use a numerical scale to measure the effect of the built environment on work performance; these scales, however, have very different sizes.

2.5.1 BUS and CBE

BUS methodology surveys have assessed over 500 buildings, and the Occupant IEQ survey by CBE has assessed over 600 buildings around the world. The BUS methodology is used extensively throughout New Zealand with benchmarks available for New Zealand (Dykes, 2012). Both surveys use the same approach in measuring people's work performance in the built environment (Sullivan, Baird and Donn, 2013) using the question:

"Please estimate how you think your productivity at work is increased or decreased by the environmental conditions in the building?" (Leaman and Bordass, 1999).

The BUS survey presents a 9-point scale ranging from −40% to +40% with 10% intervals for their productivity question. The occupant IEQ survey developed by the CBE uses a smaller 7-point scale ranging from −20% to +20% (−20%, −10%, −5%, 0%, +5%, +10%, +20%) as shown in Figure 6 (Center for the Built Environment, 2013, cited in Sullivan, Baird and Donn, 2013).

It is already a known problem in questionnaire design that the size of the scale can change the results (Schwarz et al., 1985, cited in Sullivan, Baird and Donn, 2013), but would people also assess the effect on their productivity differently? Would the results obtained from CBE scales just be double what is obtained from using a BUS scale? The ranges of the scales used in the BUS survey questionnaire for perceived productivity is twice as large as the Occupant IEQ survey, and the minimum intervals are also larger. Therefore it is expected that using the BUS would get results with larger effects, but by how much?

This content is unavailable.

Please refer to the survey images by CBE and BUS. Alternatively, the survey images are also available in document by Sullivan, Baird and Donn, 2013)

Figure 6 BUS (above) and CBE (below) scales used to assess the environmental effect on people's work performance.

Self-reported work performance measures are used as a common method to quantify building occupants' work performance. As already mentioned in section 1.1, for businesses where increasing the office workers' work performance by improving the work environment could potentially be a beneficial investment (Feige et al., 2013; Lan, Wargocki and Lian 2011), change in work performance in numerical values is likely to be far more useful than ordinal scales. Since self-reported work performance measures are commonly used, it is important to examine how the scale sizes used in the surveys might affect how people assess their work performance. Assessing the two systems side by side could also make buildings assessed by the BUS and CBE comparable in terms of how the built environment affects work performance. In addition, if there are no differences between how people assess their performance using the different-sized scales, the results could suggest whether people are at least consistent when assessing how the indoor environmental conditions have an effect on their work performance.

2.5.2 Revised Thesis Question

“How do scale sizes influence how people assess the indoor environmental effect on their work performance?”

2.6 Chapter Summary

- Work performance is a better term used for how people perform in the built environment, than the term productivity. Productivity is a ratio of output/input, and therefore productivity is better used as an economic expression of human performance.
- A comfortable environment for people to work in refers to functional comfort, which is the environmental conditions that support people to finish their job-related tasks and activities.
- Six studies compare self-reported measures against objective measures of work performance. Although it is difficult to conclude that self-reported work performance is an accurate representation of “actual” work performance, most of the studies show trends in both self-reported and measured performance, which increase when the work environment improves, and decrease when environmental conditions are causing stress on the occupants.
- A number of environmental stressors do not have an obvious effect on work performance. The effects are mostly under 10%, with some around 15% and very few over 20%.
- Laboratory and field studies have different pros and cons that should be considered, depending on what is to be measured.
- There are currently two main types of questions and two main types of scale used in surveys to assess the environmental effect on people’s work performance through self-reported work performance. Many of the past studies use the direct approach, but it could be argued that the environmental approach reduces bias when measuring the environmental effect on people’s work performance. Of the ordinal and numerical scale, the numerical scale is more favourable, as it produces numerical figures that can be used in a predictive sense, making it easier to quantify the results as well as being easier for people to assess the environmental effect on their work performance.
- Indicators of work performance have been used to estimate the environmental effect on work performance. Common indicators used are motivation, job satisfaction, absenteeism, SBS symptoms and cognitive performance.
- Major surveys by BUS and CBE use the same environmental approach for their question but use different-sized survey scales.
- Does scale size affect survey results? If so, by how much?

3 Methodology: Designing a Survey and an Experiment

This chapter describes the methodology to examine how scale sizes in surveys affect how people assess the effect of the environment on their work performance. An experiment is developed to examine the difference in the results from using BUS and CBE surveys in correlation with other self-reported indicators of work performance and cognitive performance under different environmental conditions.

To examine how different-sized scales affect how occupants assess the change in their work performance, surveys were designed to be applied after working under different environmental conditions. For the experiment, participants were split into two groups of even numbers, the BUS group assessing their work performance using the ±40% scale and the CBE group assessing their work performance using the ±20% scale. To compare the trends in change in work performance between BUS and CBE groups rather than comparing two points on a graph, multiple environmental stress levels have been tested. Pink noise was used as the environmental stressor, to cause stress on the occupants and to create a change in work performance. Three different noise levels were tested to observe the change in work performance under different intensities of environmental stressors, because the results might not show a linear trend. Commonly used work performance indicators such as motivation, job satisfaction, fatigue, distraction and cognitive performance were also measured and surveyed to compare these indicators with how the participants assessed their work performance in different environmental conditions. Finally, as a cross-check of the self-reported survey responses, a number of cognitive tests were applied.

For the experiment, subjects were required to work for 60 minutes on their usual tasks under noise conditions, fill in the post-occupancy survey after the 60 minutes, and complete a set of cognitive performance tests. The majority of the subjects worked on their usual tasks under their usual work environment, therefore the subjects were familiar with their environment rather than being tested in an unfamiliar environment such as the laboratory. The subjects were not notified that it was an experiment about how different scale sizes affect how people assess their performance differently, but they were notified that it was an experiment to see how the environmental conditions (noise in particular for this experiment) affect the people in the work environment. Explanation about the scale sizes had been removed to avoid the subjects concentrating on their work environment conditions rather than their work. In addition to this, “dummy” questions that are not related to work performance were included to avoid the emphasis on the perceived productivity question.

The details of the methodology for this study are described in the following sections:

1) Section 3.1 describes what has been measured and obtained from the experiment.
2) Section 3.2 describes the environmental conditions used as environmental stressors for the experiment.
3) Section 3.3 outlines the overall process of the experiment.
4) Section 3.4 explains the number of participants required for the experiment.
5) Section 3.5 lists the potential limitations and issues with the methodology design.

3.1 Post-Occupancy Evaluation Survey

Refer to Section 6 Appendix C, for the surveys used in the experiment, with the BUS survey in Section 6 and the CBE survey in Section6.

The survey designed for the experiment covered questions about self-reported work performance and indicators of work performance such as motivation, satisfaction, fatigue and distraction as well as dummy questions. The survey was distributed to the occupants in paper form. BUS recommends the paper form rather than a web-based form, whereas CBE is web-based but does not state recommendations. The surveys were completed before the cognitive performance tests as the occupants might have assessed their work performance based on their cognitive performance tests. Wording and style of the scales are kept the same as the original questions and scales used in the BUS and CBE surveys.

3.1.1 Work Performance Question and Scale

The question used in the survey to measure the effect of the work environment on work performance was the same environmental approach question used in the BUS and CBE surveys: "Please estimate how you think your productivity at work is increased or decreased by the environmental conditions in the building?" (Leaman and Bordass, 1999). This question was chosen as it is used in both BUS and CBE surveys, but also to minimise subjects' assessing the change in their work performance based on their personal characteristics. Personal health, life events and other seasonal factors could affect an individual performance on different days. Therefore, for questions simply asking how an individual's work performance has changed, the subjects would take the above and potentially more personal situations into account.

All the other questions about indicators of work performance (Section 3.1.2) and dummy questions (Section 3.1.2.1) were kept consistent with the same 7-point ordinal scales with thumbs up on the right to report the good effects of the work environment and thumbs down on the left to report the bad effects of the work environment.

The exact scales from BUS and CBE surveys were used for subjects to assess the effect of environmental conditions on their work performance, following the same style scales used in both of the surveys, 9-point −40% to +40% for the BUS survey and a 7-point −20% to +20% scale for the CBE survey. Refer to Figure 6 for the scale used for the survey in the experiment.

A pilot study had been conducted previously, researching the relationship between cognitive performance and self-reported work performance using temperature as the environmental stressor. During the pilot test, a number of questions were raised as to what the subjects refer to when they assess their work performance changes. The main reference points were that the subjects were to base their work performance on their best work performance, worst work performance or usual work performance. Although this would affect how the subjects answered the self-reported work performance, the subjects were not advised on what reference point they were to use when they were assessing their work performance. This is because both the BUS and CBE surveys do not state how the subjects are to assess their work performance, which keeps the experiment survey similar to how both the surveys are assessed in the real world.

3.1.2 Comparing Indicators of Work Performance with Self-Reported Measures

For this study, the self-reported work performance results from using BUS scales were compared with results from CBE scales in correlation with the other self-reported indicators of work performance of motivation, job satisfaction, fatigue, distraction and cognitive performance as the objective measure of work performance. The choice of these four indicators was based on studies stating the strong relationship between work performance and one or more of the indicators (Al Horr et al., 2016; Lorsch and Ossama, 1994, cited in Cui et al., 2013; Lan et al., 2012, cited in Kershaw and Lash, 2013; Purdey, 2013; Roelofsen, 2002, cited in Purdey, 2013; Tham and Willem, 2010; Valančius and Jerelionis, 2013; Bitner, 1992, cited in Riratanaphong, 2014; Tanabe, Nishihara and Haneda, 2007, cited in Boerstra et al., 2015; Vischer, 2008; Thomas, 2010; Leaman and Bordass, 2007, cited in Agha-Hossein et al., 2013; Baird, Leaman and Thompson, 2012; Choi et al., 2012).

Questions from the BUS and CBE surveys were used for the questions about indicators used in this research, drawing on examples included in the studies by Bhawani (2010), Parkinson et al. (2017) and Roy (2008). 7-point ordinal scales were used for all of the questions for these indicators of work performance. The dummy questions, to prevent subjects from focusing too much on the environmental condition while working, were also based on questions from surveys from these studies.

3.1.2.1 Dummy questions

The dummy questions in the surveys are included because it was assumed that questions just about work performance and the indicators, could potentially influence the subjects to assess the environmental conditions on how they made them feel, rather than how the environmental condition had an effect on their work performance. The questions are about the overall comfort in the space, how the environment could affect their health and how they were satisfied with the acoustic qualities in the environment.

3.1.2.2 Cognitive performance test

The cognitive performance tests for this study are a series of nine different tests covering four different cognitive functions. They were developed and used by Lan et al. (2009). The older cognitive tests from Lan et al. (2009) have been used rather than the cognitive tests used in more recent studies by Lan, Wargocki and Lian (2011) and Lan, Wargocki and Lian (2014). This is because the test from Lan, Wargocki and Lian (2011) replaced some cognitive tests with simulated office tasks, and tests by Lan, Wargocki and Lian (2014) only included four cognitive tests, because the focus of the study was to integrate speed and accuracy into one metric of human performance. Cognitive performance tests were used rather than simulated office tasks, because results from cognitive performance tests are more applicable to a wide range of people and industries.

Simulated office work performance tasks are office tasks that have been simplified to a measurable state. However, the results may not necessarily reflect actual work performance because the tests are generally conducted for a shorter time than working in the office and the simplification of the office work implies that they are not the same as actual tasks in the office. Similar arguments could be made against the use of cognitive performance tests, but results from simulated office work performance cannot be applied to workers with different tasks, companies and industries whereas cognitive performance tests are designed to assess overall cognitive functions. In addition, simplifying office work into a test will also require quantifying quality aspects of office work, meaning that quality of office work is less considered. Fisk (2000) also argued that experimental studies are likely to cause overestimations in work performance reductions, and that measured results of cognitive performance do not accurately represent office work carried out in the office, because cognitive performance is observed as only one of the indicators of how objectively measured work performance is affected by the environmental conditions.

For this thesis experiment, other self-reported indicators such as motivation, fatigue and job satisfaction and distraction were also measured and compared to how people assessed their work performance with different scales.

All cognitive performance tests were developed in Microsoft Excel, from the description by Lan et al. (2009), with additional explanations listed in Table 12

Table 12. List of cognitive performance tests developed by Lan et al. (2009).

Test	Cognitive function	Test description
Letter search	Perception – visual search	• Identify presence or absence of a target letter from a string of ten letters, with a blank space between each letter. • Buttons *select* and *deselect* were located under each letter and each blank space; *deselect* is used to undo *select* when the subjects have mistakenly selected a letter that they want to deselect. • Blank spaces are selected when the target letter is missing from a string of letters that form a word • 50 strings per test. • Record: number of correctly selected letters and response time.
Overlapping	Perception – spatial orientation	• A photo of six paper cut-outs of geometrical shapes (circle, ellipse, triangle, square, rectangle, pentagon) in a pile was presented, subjects were to number the shapes from top to bottom with 1 being the top and 6 being the bottom. • Six trials per test. • Record: number of correct orders (not correct trials) and response time.

(continued)

Table 12. List of cognitive performance tests developed by Lan et al. (2009) (continued).

Test	**Cognitive function**	**Test description**
Memory span	Learning and memory – recall memory, verbal memory, attention	• Remember a string of numbers presented, one second per digit, and reproduce the string of numbers after they disappear. • Not allowed to use pen and paper. • Numbers start from four digits; every length is presented twice. • Test continues until it reaches second string of twelve digits or two mistakes are made consecutively. • Record: number of correct sets of strings and response time.
Picture recognition	Learning and memory – recognition memory, spatial memory, attention, response accuracy	• Identify the ten target pictures out of 20 pictures presented. • First, subjects were shown the ten target pictures (target stimuli, TS), one second per picture, the test includes a mix of ten TS and another ten more randomly chosen pictures (objective stimuli, OS). Subjects are to identify which of the pictures are TS and which of the pictures are OS. • Record: Number of correctly recognised pictures and response time.

(continued)

Table 12. List of cognitive performance tests developed by Lan et al. (2009) (continued).

Test	Cognitive function	• Test description
Symbol digit modalities test	Learning and memory – recall memory, verbal memory	• Ten symbols with a number that corresponds to the symbol were presented. Enter the number under each of the re-ordered corresponding symbols, remember them and enter the numbers again from memory for another reordered set of symbols. • Subjects cannot look back at the symbol and number pairs when they enter the numbers from memory. • Not allowed to use pen and paper. • Record: number of correctly recalled pairs and response time.
Number calculation	Thinking – mathematical procedures, response speed	• Add two three-digit numbers. • Random three-digit numbers (between 100 and 999) was generated, numbers with no 0 as any of the digit was chosen to avoid making some calculations easier than others. • 20 sets of calculations per test. • Record: number of correct sums and response time.
Conditional reasoning	Thinking – verbal reasoning	• One premise and four statements are presented. Choose one statement from four statements that can be concluded from the premise. • Allowed to use pen and paper. • Three sets of premise and statements per test. • Record: number of correct conclusions and response time.

(continued)

Table 12. List of cognitive performance tests developed by Lan et al. (2009) (continued).

Spatial image	Thinking – spatial reasoning, imagination	• One cuboid surface development (i.e. open box) and four cuboids are presented. Choose one cuboid images that could be made from the cuboid outspread. • Three trials per test. • Record: number of correct cuboid and outspread matches and response time.
Visual choice reaction time	Executive functions – response speed and accuracy	• One of four stimuli (left pointing arrow, right pointing arrow, triangle on left side of image, triangle on right side of image) is presented per trial. Identify and respond as to whether it is the *left* stimuli (left pointing arrow and triangle on left side of image) or the *right* stimuli (right pointing arrow and triangle on right side of image). • Respond to the *left* stimuli by keying in the command Ctrl + u and *right* stimuli by keying in Ctrl + p. • 50 trials (50 stimuli) per test. • Record: number of correct response and response time.

3.2 Environmental Conditions to Cause Change in Work Performance

Environmental stressors were added to cause a change in work performance. Three different environmental conditions were used to examine how different scales sizes affected the trends between the two groups.

To create an effect on work performance, the environmental conditions were altered enough to cause a stress on the subjects but were adjusted to cause no harm. The studies described in section 2.3.4 show that there is no obvious evidence that multiple environmental conditions cause a larger decrease in work performance. Therefore, only one environmental condition was altered to cause an effect large enough to be visible on the self-reported work performance scale.

Pink noise has been used in this research to cause stress on the subjects and affect their work performance through environmental stressors. Noise is known to affect work performance in office environments (Brocolini, Parizet and Chevret, 2016; Reinten et al., 2017) and pink noise, commonly used for sound masking for privacy and masking from sounds that cause greater annoyance, is reported to cause negative effects (Keighley and Parkin, 1979, cited in Al Horr et al., 2016). The experiment was conducted at weekends to avoid occupants who had not agreed to participate being exposed to the pink noise conditions. Noise was also more controllable at weekends because the air conditioning in the room was turned off. As the room was occupied by less than 10 people during the weekend, the room sound levels were naturally low. For these reasons, the noise was the easiest environmental stressor to control within the experiment room.

Room windows werekept open for ventilation during the test as air conditioning units do not operate during the weekends. Although temperature, relative humidity and lighting are affected by the weather, noise is not significantly affected by different weather conditions.

3.2.1 Sound/Noise Could Cause Annoyance to Affect Work Performance

Pink noise within the room was created using a pink noise generator and speakers. The pink noise generator was connected to an attenuator, to an amplifier and to the speakers. The level of noise from the speakers was adjusted each time with both the attenuator and the generator and measured to ensure the same noise levels for the same noise condition experiments on different days.

The environmental conditions that the room was set to:

1. No noise generated. (Assumed to be 40 dB(A))
2. Low noise level generated from the speaker (55 dB(A)).
3. Annoying noise level generated from the speaker (65 dB(A)).

For the *no noise* condition, it is assumed that the sound levels were 40 dB(A) as the sound meter was unable to measure sound levels under 40 dB(A). Sound levels within the room were consistently under 40 dB(A) during the *no noise* conditions, except for when doors opened and shut. The sound meters were tested with four other sound meters and the sound meter showing the median reading was chosen to measure the sound level in the experiment environment.

The selection of 55 dB(A) for the first annoyance noise level was the result of research from the World Health Organization (WHO) by Concha-Barrientos, Campbell-Lendrum and Steenland (2004). This research stated that with the source adapted by the Health Council of the Netherlands, HCN (1999, cited in Barrientos, Campbell-Lendrum and Steenland, 2004) and de Hollander et al. (2004, cited in Barrientos, Campbell-Lendrum and Steenland, 2004), occupants were observed to have reported annoyance at 55 dB(A) in office environments.

For the loudest noise level, 65 dB(A) was chosen as a noise level that is noticeably louder than the 55 dB(A) but lower than the suggested safe threshold by WHO (1999, cited by Hannah, Page and McLaren, 2016). The sound level of 65 dB(A) (Surprenant, 1999) and up to 75 dB(A) (Perham, Banbury and Jones, 2006) has been tested in past studies as an environmental stressor, therefore 65 dB(A) is a sound level safe enough to be tested on people while causing them stress and annoyance. The sound level of 75 dB(A) was not used, even though some studies state that 8-hour exposures of 75 dB(A) do not cause hearing impairment; this will be discussed in section 3.2.1.1.

3.2.1.1 Safe noise levels

Tested noise levels are significantly lower than those suggested to be a dangerous constant noise level in a work environment for an 8-hour working day. Hannah, Page and McLaren (2016) have also claimed that 70 dB(A) for extended periods are generally accepted by professionals as having no or little likelihood of any effect on long-term hearing acuity. WHO (1999, cited by Hannah, Page and McLaren, 2016) also claims that sound levels up to 75 dB(A) are not expected to cause hearing impairments even for a long noise exposure. However, Concha-Barrientos, Campbell-Lendrum and Steenland (2004) claim that the threshold for hearing loss for adults is 75 dB(A).

Government publications and guides state that for noise and sound levels above 85 dB(A), employers should take all reasonable steps to ensure no employees are exposed to noise above those levels during an 8-hour day (Hannah, Page and McLaren, 2016; Occupational Safety & Health Service, 1996a; Occupational Safety & Health Service, 1996b). Although tests have not been conducted on infants or pregnant women, Concha-Barrientos, Campbell-Lendrum and Steenland (2004) mention the threshold of hearing loss for unborn children to be 85 dB(A). Also, NZS 1269 predicts 95% of the people exposed to 85 dB(A) for 8 hours, will not develop a hearing loss that will exceed 10 dB(A) (mild hearing loss) over the lifetime of work (Occupational Safety & Health Service, 1996a).

3.2.2 Latin Square Method

Groups of people recruited for the experiment were split into two groups, one assessing their work performance with the BUS survey and the other assessing their work performance with the CBE survey. These two groups were further divided into three groups (Figure 7) to distribute the effect of becoming familiar with the loud noise level, the accumulation of annoyance of the loud noise levels, and the learning effect of the cognitive performance tests. The experiment was designed using the Latin square method as shown in Table 13to control these effects for this experiment.

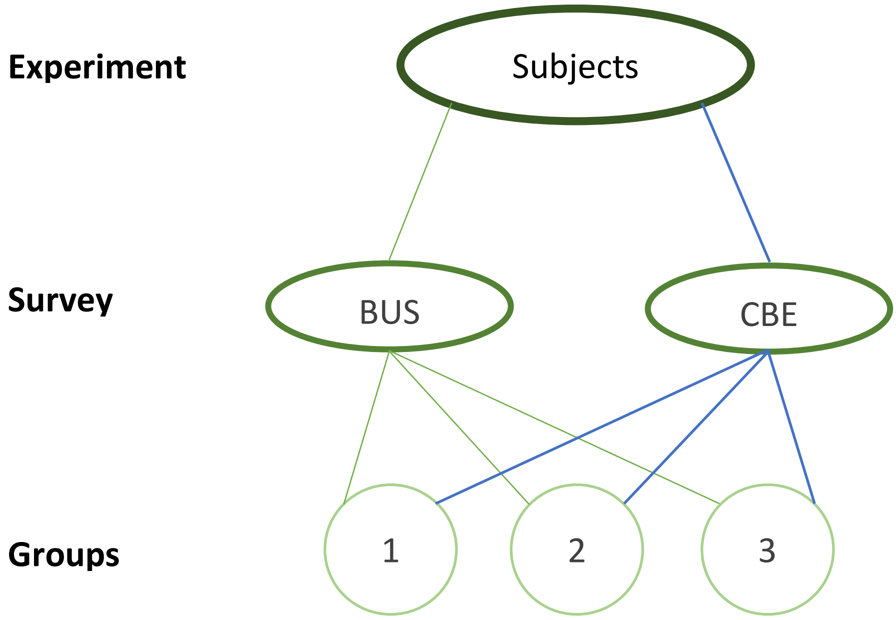


Figure 7 Groups of people divided for the experiment.

Table 13. Latin Square Method used in this experiment.

		Sequences		
		1	**2**	**3**
Groups	**1**	No noise generated	55 dB(A)	65 dB(A)
	2	55 dB(A)	65 dB(A)	No noise generated
	3	65 dB(A)	No noise generated	55 dB(A)

3.2.3 Room

The room used for the experiment (Figure 8) was the usual working environment for most subjects and was a room within the same building for the other subjects who do not regularly work in the room. The possibility of having subjects from different rooms assessing the environmental effect on their work performance differently from subjects who were familiar with the room is discussed in Section 4.3. The experiment was conducted during the weekend when there were fewer occupants in the building. This creates a quiet condition of less than 40 dB(A), in a room where it is easy to control the sound level of the pink noise amplified from the speakers.

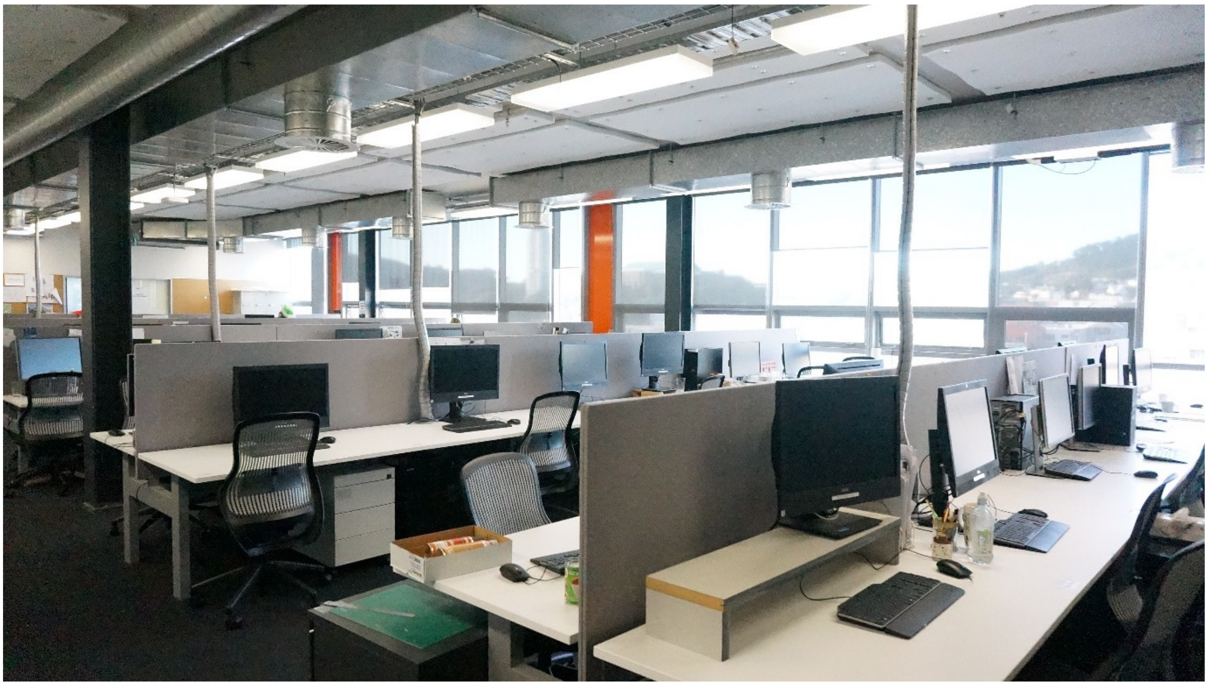


Figure 8 Room used for experiment with all artificial lights on. The room faces south with glazing covering ceiling to floor of the whole south facade.

Since the HVAC systems are turned off during the weekends due to their automatic settings, all windows were opened to bring in fresh air, and all artificial lights in the room were turned on for all of the conditions for consistency. Although the weather does affect light, temperature and humidity conditions in the room, there were no unusually hot or cold days, and no subjects reported any changes in work performance due to temperature, humidity or light conditions. The room faces south ensuring that the daylight variation varied between blue and cloudy sky lighting with no sun intrusion.

Figure 9 shows the sound levels at different areas of the room for the three conditions and the experiment room dimensions. No subjects were placed next to the speakers, where sound levels exceeded 75 dB(A) for some conditions. Although there are some inconsistencies in the sound levels at different distances from the speaker, the noise generated for the experiment was not to affect people's work performance at particular sound levels. The noise generated was for creating different levels of change in the work environment. Therefore, although the noise levels are not consistent throughout the room, the degree of change in the work environment stressors stays consistent for each subject because they sit in the same locations throughout the three sequences of the experiment.

The inconsistencies in the sound levels could have been minimised through the use of multiple speakers and the use of headphones (or earphones) to generate the noise. However, for a pilot project of this scale, developing an understanding of whether a more complex project was worthwhile, and understanding if the process could produce results these were excessive measures. In addition, allocating time to set up the two speakers to generate consistent noise throughout the room was not practicable within the 90 minutes per day during the weekend that the experiment was permitted and risked converting the normal work environment into a quasi-laboratory artificial work environment . Headphones to generate noise were also not practicable as these would have significantly altered the experiment from a normal workplace assessment to an artificial lab environment. This artificial lab impression would have been emphasised by the extra intrusion on the normal workplace of completing the required calibration of all equipment before every experiment such as setting the noise for each headphone.

To reduce the inconsistencies with the subjects experiencing different noise conditions in different areas of the room, the subjects were allocated to sit in the same seats throughout the three test conditions. With subjects present in all rows, all subjects will experience approximately 10dB(A) different between the three conditions: <40dB(A), 55dB(A) and 65dB(A). To further explain, the subjects in the row furthest away from the speaker will experience approximate 10dB(A) difference between <40dB(A), 50dB(A) and 61dB(A) and the subjects in the row closest to the speaker will also experience similar noise changes between <40dB(A), 58dB(A) and 68dB(A).

Although the windows were opened to compensate for the HVAC system not being turned on during the weekend, there were no recognisable noise differences due to the location of the experiment room in relation to the surrounding outdoor environment. The room is on the fifth floor above a quiet and narrow side street. Temperature could also have been an environmental stressor; however, turning on the HVAC was impracticable as this would require resetting the HVAC system for this experiment alone and risked adding fan noise in an uncontrollable manner to the workplace

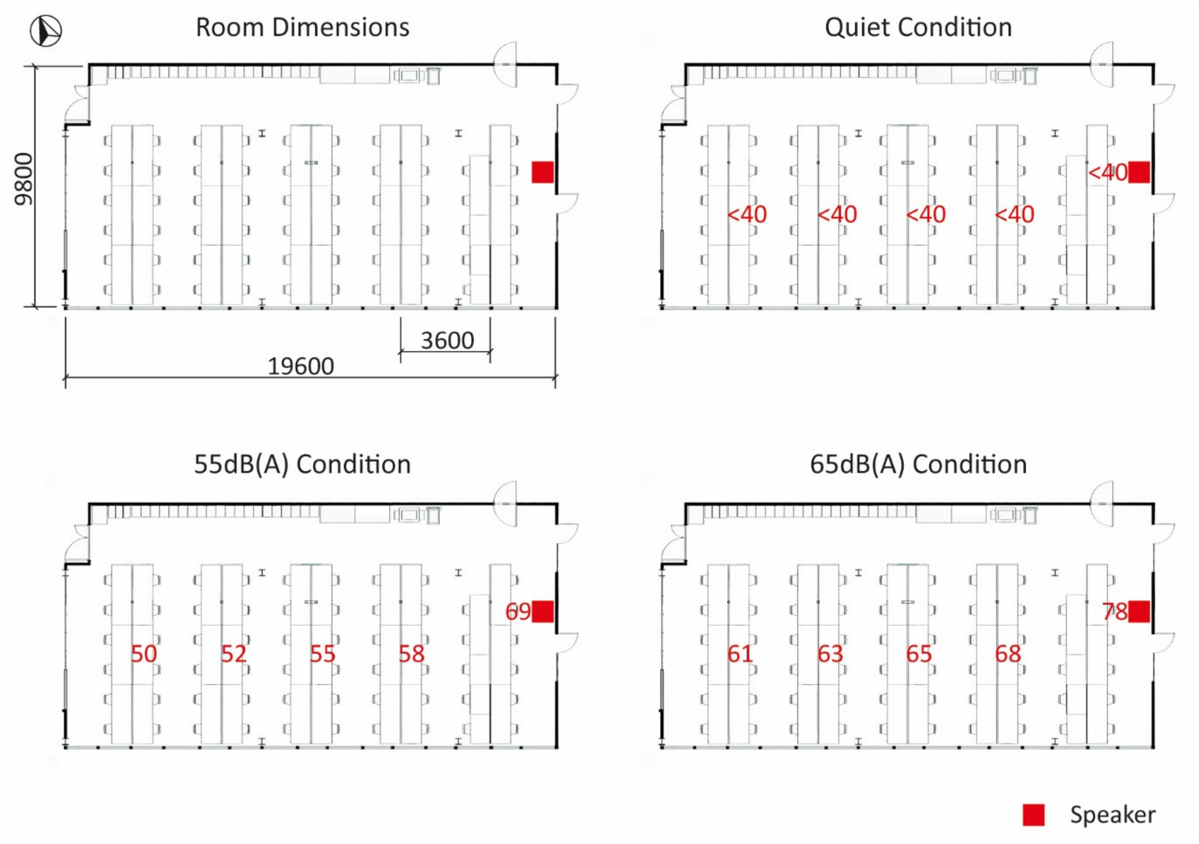


Figure 9 Dimensions of the experiment room with red numbers showing the sound levels at different areas of the room for different noise conditions. Room not to scale. Dimensions in mm and sound levels in dB(A).

3.2.4 Subjects

Subjects recruited for the experiment were mostly students who were familiar with the experiment room. Subjects who usually worked in this room had been working on their thesis for more than six months, and therefore the subjects could be considered to know the tasks they were needing to complete. However, not enough subjects could be recruited from just students within the room, nor from the other students working on their thesis in the same building. Therefore, some students who were working on university coursework, undergraduate students, were recruited for the experiment, which could have affected how they reported the environmental effect on their work performance. This is because students not familiar with the experiment room could be considered to have not known the environment well enough to make appropriate assessments, and students working on university coursework could be regarded as not knowing enough about their tasks or work to be able to make good assessments about how their work performance had changed. The three types of subjects that made up the total of 30 subjects for this experiment included:

- Students working on their thesis and familiar with the work environment (18 subjects)
- Students working on their thesis and unfamiliar with the work environment (6 subjects).
- Students working on coursework and unfamiliar with the work environment (6 subjects).

Students from the minority categories of working on their thesis and unfamiliar with the environment and the category of students working on coursework and being unfamiliar with the work environment were divided equally between the BUS and CBE groups.

Of the 30 students acquired for the experiment, 24 students (students working on thesis) were architecture students with no prior detailed knowledge about building science and the topic of work performance. Building science thesis students were not recruited because of the risk of having prior knowledge of the topic which might have affected how they perceived and answered the surveys. The remaining six students were project management students working on coursework who had not picked the topic of work performance for their studies.

3.3 Experiment Procedure

The experiment required the subjects to work on their usual work (complex tasks such as thesis writing and university coursework) for 60 minutes, fill out post-occupancy surveys and to complete the series of nine cognitive tests (simple tasks) from fill out post-occupancy surveys and to complete the series of nine cognitive tests (simple tasks) from Table 12. Before starting each sequence of the experiment, subjects were notified again about what they are expected to do in the experiment. For the subjects doing their first sequence of the experiment, a shortened version of the cognitive tests was recommended for the trial, so they were aware of the type of questions and tests and how to answer each test. The step-by-step procedure for the tests was:

1. Short briefing.
2. Working on usual work with the noise conditions for 60 minutes.
3. Survey.
4. Cognitive performance tests.

The experiment required the participants to come in on three different days for the three different environmental conditions. This was to avoid the accumulation of fatigue, which would affect the results differently between groups who started with a *comfortable* condition and an *uncomfortable* condition. From the pilot test conducted for a different study with temperature as the environmental stressor, the group of subjects who started the experiment with the uncomfortable condition reported fatigue after the first sequence. This could have caused the cognitive test results in the first sequence of the uncomfortable environment to not have a noticeable change compared with the second sequence, in the comfortable environment, even with the learning effect. The experimental setup was tested on different days, because people could get used to or accumulate annoyance for noise, which is also the reason that the Latin-square method has been used.

Cognitive performance tests were conducted at the end of each sequence to avoid the subjects being influenced by the cognitive performance tests when they were assessing the environmental effect on their work performance. Subjects who felt confident with a series of simple tasks in the cognitive performance test would perceive that their work performance is high, and people who were not as confident would perceive that the environment would have had a great impact on decreasing their work performance, as discussed in Section 1.3.

All the tests were conducted during the weekend at the same time of the day, on the same day (Saturday), keeping the circadian rhythm consistent so as not to cause a difference in results due to an accumulation of fatigue at different times of the day and at different times of the week.

3.4 Estimating Sample Size Required for the Experiment

For subject recruitment, a statistical simulation was conducted to estimate the number of subjects required for this experiment. Since the main focus was to examine the difference between self-reported work performance using the BUS and CBE scales, the number of subjects required for the experiment has been based on the margin of error of the difference between the two results. For example, if the difference between BUS and CBE results were 20, the results could range from 20 + x to 20 - x. The acceptable range for this simulation has been set to 5, as 5% of work performance could be considered significant, as discussed in section 2.3.3.2.

The simulation ran 10,000 calculations of the differences between two groups (BUS and CBE) for three trials (three different noise conditions) using standard deviation (SD) to present the spread of the data. The SD of the survey results was gathered from studies by Brown (2010), Baird, Leaman and Thompson (2012) and Parkinson et al. (2017) who have used the BUS surveys, and by Gou, Prasad and Lau (2013; 2014) who have used the CBE surveys. Studies with more people were prioritised for calculating the mean SD for running the simulation. SDs of each study and the calculations are included in Section 6 Appendix D.

Figure 10 shows the graph from the R statistical analysis program with possible simulation outcomes and the different confidence levels with estimated sample sizes. To achieve a margin of error of ±5% with a confidence level of 95% requires roughly 12 subjects in each group (24 subjects in total). However, since this is the minimum number of people required, 40 subjects were recruited for this experiment. Eventually, 10 subjects withdrew, leaving 30 subjects in total.

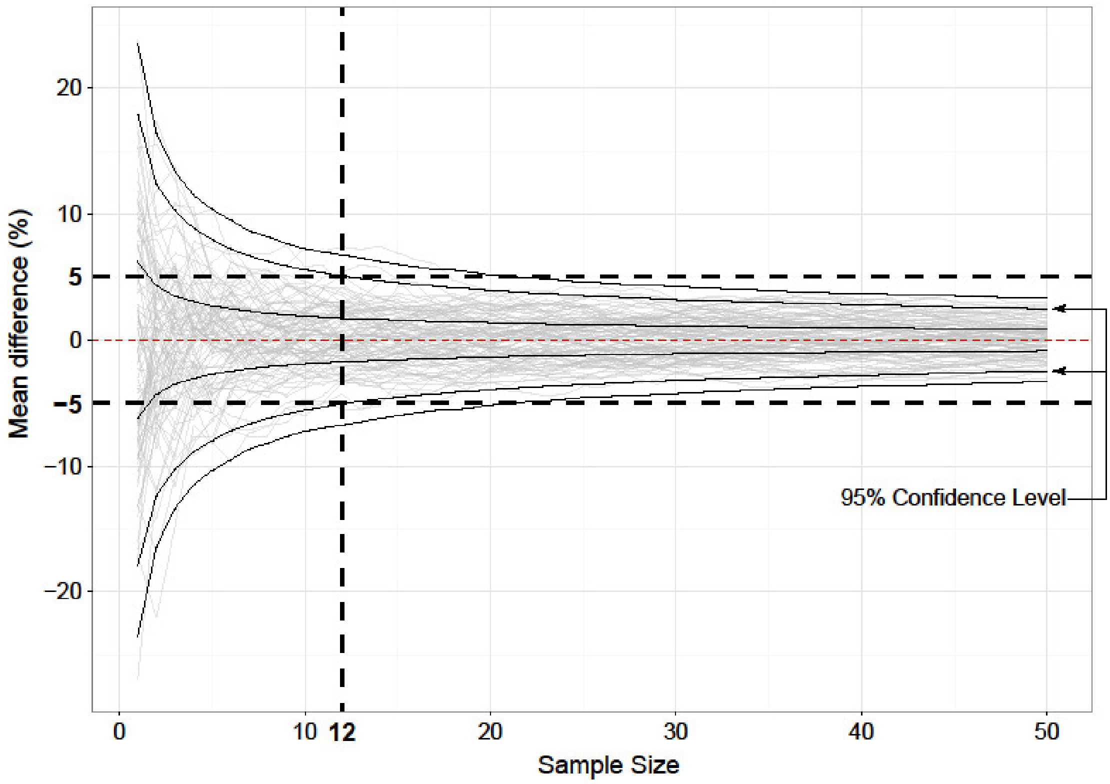


Figure 10 Graph showing 100 of the 10,000 simulations calculating the differences of the BUS and CBE means in grey lines, with inner black curves showing the 50% confidence levels, middle curves showing 95% confidence levels and outer curves showing the 99% confidence levels.

3.5 Limitations and Issues

3.5.1 HVAC Systems Turned Off During the Experiment Dates

The experiment was conducted during the weekends to cause less annoyance to building occupants who are not participating, but the HVAC system was turned off during the weekends. As mentioned in Section 3.2.3, the room condition was kept consistent in all conditions and no extreme or major temperatures were noticed and no comments on environmental stressors other than noise were made by the participants. Although kept consistent, small changes to environmental conditions such as temperature and humidity (as well as light, since weather does affect lighting conditions due to the large windows) could have had effects on the subjects' work performance. Therefore, it is not completely certain if subjects reported decreasing work performance and decreases in cognitive performance results were due to noise only, or were affected by other environmental stressors. This research did not focus on how all these environmental stressors affected the work performance. However, the variation in these other variables of temperature, humidity and light levels due to fluctuations in outdoor conditions were designed to be small because there is only one external, south facing, wall in the room and the windows were able to be closed had a cold Southerly wind seriously affected the indoor conditions. The noise conditions were designed to be sufficiently large that their potential effect relative to the effect of these other variables would also be relatively large.

3.5.2 Exposure Time of the Environmental Stressors On the Experiment Subjects

To maintain interest in the experiment, to avoid taking up too much time, and to avoid times of high stress, one sequence of the experiment was roughly 90 minutes in length. At recruitment stage, the majority of the students were not interested in the experiment after being notified that they would have to spend three sequences of experiments with noise for longer than 60 minutes each.

However, this is significantly shorter than the usual workday of 8 hours in the work environment. If the environmental conditions were to continue for 8 hours every day, fatigue and stress could increase and even people who find pink noise rather comfortable could also start to report annoyance and distraction from noise under long exposure times. People could assess the environmental effect on their work performance differently under 90 minutes of exposure and 480 minutes (8 hours) of exposure, and therefore the differences between BUS and CBE obtained from this experiment may not be a representation of how different the results would be in a real work situation.

4 Results: Did the Scale Sizes have an Effect?

To examine how different survey scale sizes cause a difference in how people assess the environmental effects on their work performance, several questions are asked in the following sections:

1) How have people assessed their work performance differently using the different-sized scales from BUS and CBE, correlating with the environmental conditions?
2) How have people assessed their work performance differently using the different-sized scales from BUS and CBE, correlating with the other self-reported indicators of work performance?
3) How do self-reported work performance results from using different scale sizes correlate with cognitive performance, an objective measure of work performance?
4) Did the subjects' characteristics such as not being familiar with the experiment room, having different tasks, having different perceptions about the survey and being aware of the noise conditions affect how they assessed the effect of the work environment on their work performance?
5) Considering the above, do scale sizes affect how people assess the effect of the work environment on their work performance?

4.1 How have People Assessed the Environmental Effect on their Work Performance Differently using Different Sized Scales?

To examine how survey scale sizes have affected how people assessed the effect of the work environment on their work performance, trend differences between BUS and CBE groups over the different environmental conditions have been analysed. Differences between the BUS and CBE results are not compared to each environmental condition, because looking at the ratio of the difference in each condition causes problems. For example, if results at one condition show values of 20 and 15 and the second condition shows 10 and 5, they both show differences of 5 but the difference between the two values would be 33% for the first condition ({5/15}*100) and 100% for the second condition ({5/5}*100).

4.1.1 Comparing the Trends in Different Environmental Conditions – How large are the differences?

The change in work performance between the quiet 40 dB(A) condition to the loudest condition of 65 dB(A) shows different trends between the two groups. The mean self-reported work performance for the BUS group decreased by 18.67% and the CBE group decreased by 11% (Table 14). There is a difference of 7.67% between the BUS and CBE groups' assessments of how the environment had an effect on their work performance between the comfortable and most uncomfortable conditions.

Although the differences between the results are not double (as the scale size difference between BUS and CBE surveys are) the difference of over 5% could be considered large enough to be significant. In this situation, the significance is not a statistical measure, but rather an observation that a productivity increase of just 5% is more than enough justification for a large investment of money as the return will be large. However, due to the large margin of error in all conditions, it could also be argued that the difference between BUS and CBE could be smaller and not considered to be significant.

Table 14 Self-reported work performance results from BUS and CBE groups in different noise conditions.

		Noise conditions (dB(A))		
Survey groups	**Results**	**40**	**55**	**65**
BUS	**Mean**	10.00	-3.33	−8.67
	Margin of error (95%)	10.25	10.40	10.65
	Sample size	15	15	15
	SD	18.52	18.77	19.22
CBE	**Mean**	6.33	−1.00	−4.67
	Margin of error (95%)	2.86	3.94	3.98
	Sample size	15	15	15
	SD	5.16	7.12	7.19

The overall trend and the gradient for the trend line are analysed in the graph shown in Figure 11**Error! Reference source not found.**. First, assuming that the no-noise-generated condition was 40 dB(A), the changes in work performance over the three environmental conditions line up close to the trend line. Although both groups have lower mean for 55 dB(A) than the trend line, the difference is within 5% and there are no obvious drops or rises in the mean self-reported work performance. The gradient for the BUS trend line is −0.7579 and −0.4439 for CBE, resulting in a gradient difference between the two trend lines of 0.31. The BUS slope gradient is 1.7 times the gradient of the CBE slope, and therefore, although not as large as double, the BUS slope is 70% greater than the CBE slope. This is assumed to be caused by the BUS survey having a wider scale of ±40% as opposed to the CBE scale having a narrower scale of ±20%.

4.1.1.1 Testing for Equivalence Between the BUS and CBE Result Trend Lines

An equivalence test is used to examine if the data from two groups are significantly different. An equivalence range with a mean of 0 is set, which indicates a range where the difference is considered to be not significant. Differences of the mean from two groups are plotted with error bars. The equivalence test is used to determine if the difference between two sets of data is small enough to be considered equivalent. Other possible outcomes of the equivalence test are shown in Figure 12.

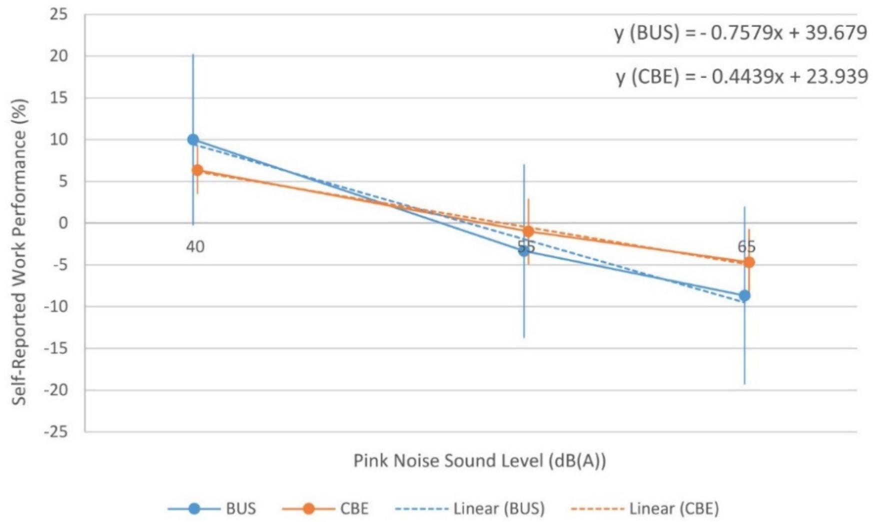


Figure 11 Results of self-reported work performance under different conditions for people who have assessed their work performance using the BUS scale (blue) and the CBE scale (orange).

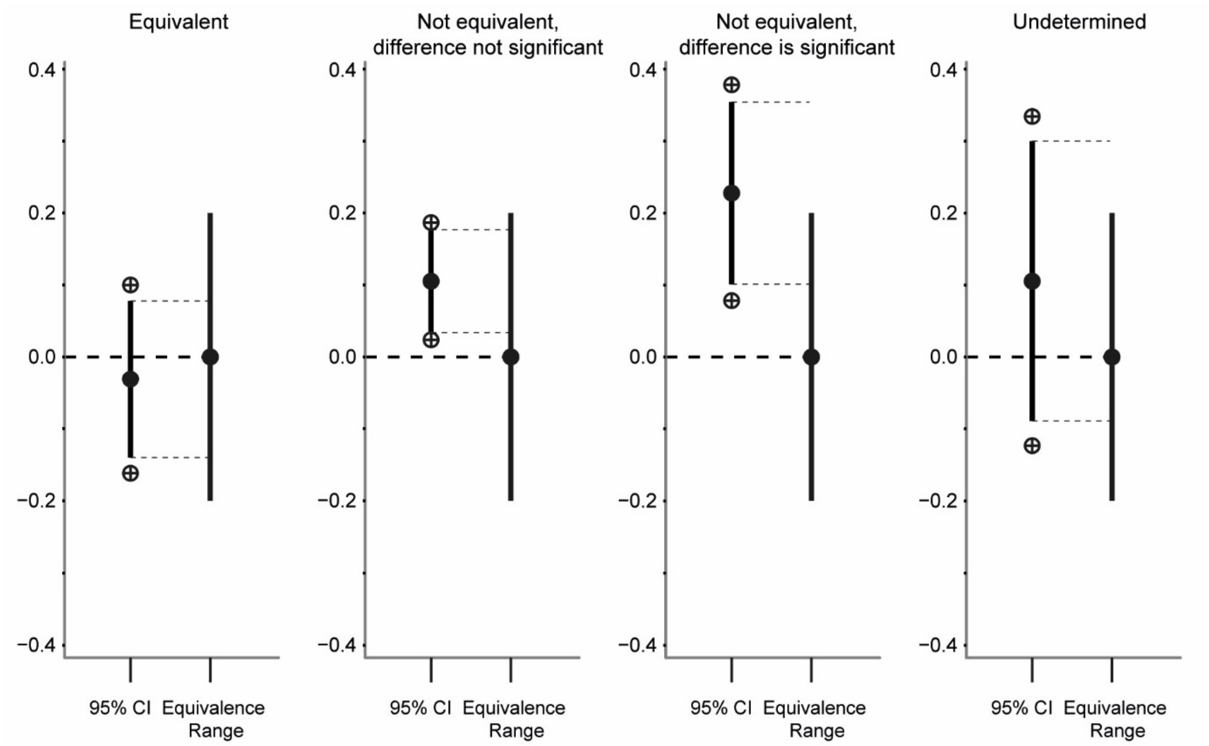


Figure 12 Example of equivalence test outcomes. The two data are considered equivalent if the mean difference and the error bars are within the equivalence range and the error bars overlap zero. The difference is not zero but is not considered

significant when mean difference and error bars are within the equivalence range, but the error bar does not overlap zero. The difference is considered a significant difference if the mean difference does not overlap with the equivalence and the error bars do not overlap with zero. The equivalence of the difference is undetermined if the mean difference is within the equivalence range, but the error bar is outside the equivalence range as well as overlapping with zero.

Figure 13 shows the equivalence testing for the gradient difference between the BUS and CBE. To determine what is considered a slope with the “same” gradient, an equivalence range of 0 ± 0.2 will be used. If the difference between two values is smaller than the equivalence range, the two are considered to be equivalent. The equivalence range of ±0.2 has been based on the significant difference of 5% (discussed in 2.3.3.2) divided by the differences in x-axis of 25 (65 dB(A) – 40 dB(A)).

Calculating the equivalence range assumes that the change in self-reported work performance is linear, and both lines (from BUS and CBE) are very close to being linear in this case. Assuming a straight slope from 40 dB(A) and 65 dB(A), the slope gradient will be −0.7467 for BUS and −0.44 for CBE. The gradient difference between this straight slope and the trend line is less than 0.012 (1.6% difference, 0.012/0.7467) for BUS and 0.004 (0.9% difference, 0.004/0.44) for CBE, and therefore could be considered a very similar slope. If the results between 40 dB(A) and 65 dB(A) vary by ±5%, the gradient will also vary by ±0.2 (5/{65-40}). Therefore, the range 0 ± 0.2 has been used as the equivalence range.

The gradient difference is calculated from the gradient of BUS minus CBE results ({− 0.7579}−{−0.4439}). The crossed circles in Figure 13 are used to mark the 97.5% confidence interval (CI), with the black lines showing the 95% CI (Lakens, 2016). This graph shows that the difference in the gradient does not fit within the equivalence range of 0 ± 0.2. However, since the positive error bars overlap 0 and also into the greater negative numbers, it is difficult to conclude that there is a difference between results produced by using the larger BUS scale and the smaller CBE scale, based on the trend differences over changing environmental conditions. The equivalence test shows that there could be a difference, but since the *p* value for the difference is 0.28, the results are not statistically significant suggesting that the experiment requires a larger sample size.

4.1.2 Estimating Number of Samples required using Results from experiment

Using the same method as that used to estimate the sample size in Section 3.4 and using a new SD calculated from the self-reported work performance results, a sample size required to obtain statistically significant results has been calculated. For the new simulation, the largest SDs from the three environmental conditions of 19.223 for BUS and 7.188 for CBE have been used. Figure 14 shows the result of simulation using a new set of SDs, suggesting that at minimum 22 samples in each group (44 in total) is required to achieve a margin of error of ±5% with a confidence level of 95%. Following the method in 3.4, more subjects than the minimum of 25 subjects in each group (50 in total) would be an adequate number for this experiment.

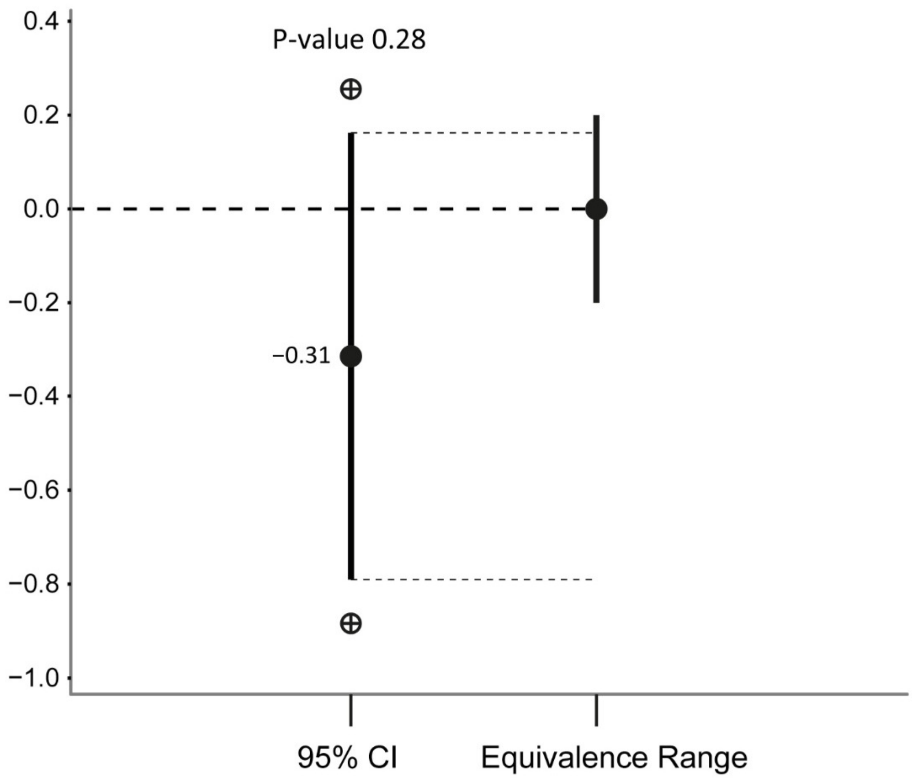


Figure 13 Equivalence test for the difference between the trend lines of self-reported work performance assessed using the BUS and CBE scales.

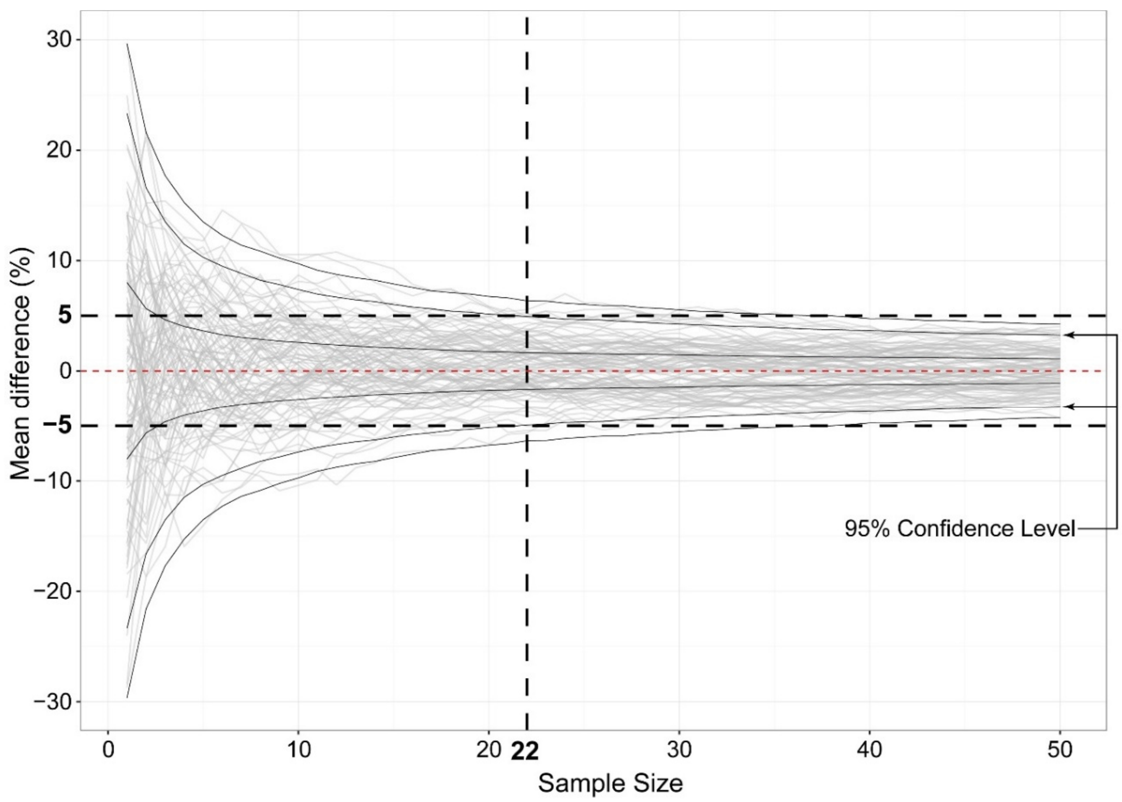


Figure 14 Graph showing the mean difference between two groups of data for each of the confidence levels with inner black curves showing the 50% confidence levels, middle curves showing 95% confidence levels and outer curves showing the 99% confidence levels to estimate the sample size using SD from results obtained from the experiment.

4.2 How do other Self-Reported Indicators of Work Performance Correlate with Self-Reported Work Performance Measures from Different Scales Sizes?

The differences observed in section 4.1 could be caused by one group reacting more to the environmental conditions than the other group. The other self-reported indicators of work performance are compared in Figure 15.

From these four graphs, each showing the other self-reported indicators of work performance, there is no obvious pattern in the trends to show that one group overreacted to the environmental conditions. Both groups show decreasing trends towards the 65 dB(A) environmental condition. A data summary of the other self-reported indicators of work performance obtained from the survey is in Appendix E.

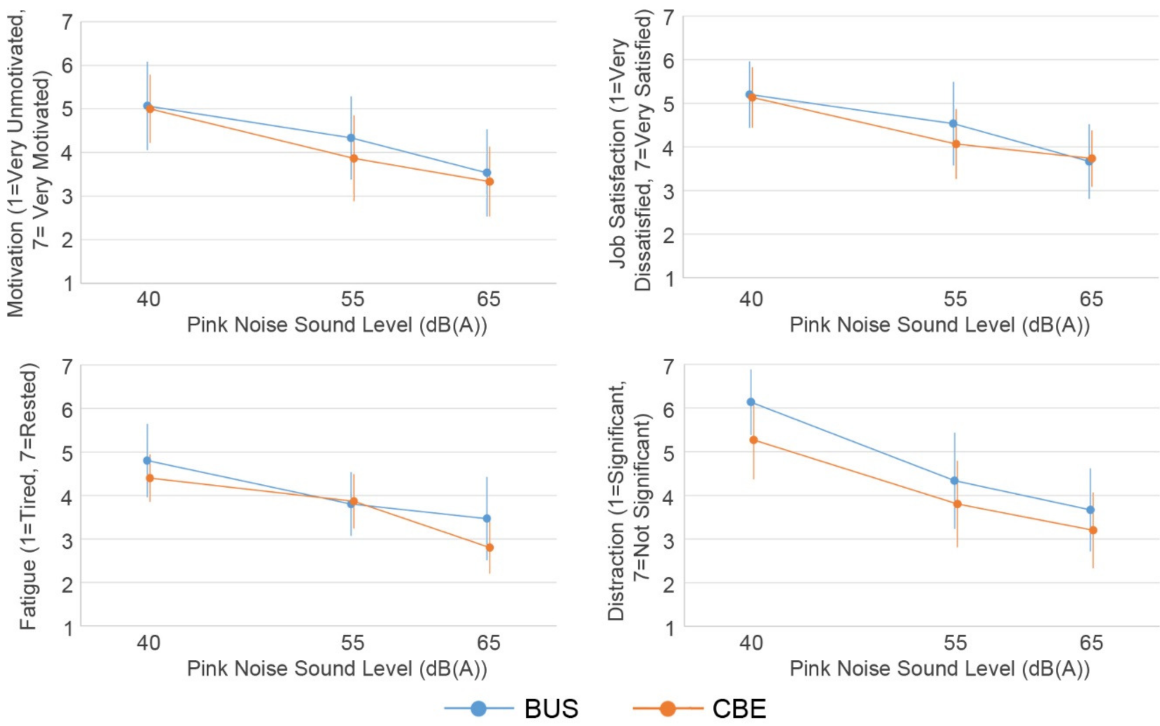


Figure 15 Trends of other self-reported indicators of work performance, motivation (top left), job satisfaction (top right), fatigue (bottom left), distraction (bottom right) for the BUS and CBE groups.

Table 15 shows the mean for other self-reported indicators of work performance with changes from 40 dB(A) to 65 dB(A) in percentages. Although other self-reported indicators were not assessed on a percentage scale, it is worth noting that these changes were observed in an environment with 65 dB(A) noise, which is close to most extreme conditions of noise that people would experience in offices. Since government-published guides advise employees to take all reasonable steps to ensure no employees are exposed to noise levels above 85 dB(A) (Hannah, Page and McLaren, 2016; Occupational Safety & Health Service, 1996a; Occupational Safety and Health Service. 1996b), a noise level of 65 dB(A) would be the highest reasonable noise levels that people would experience in office environments. The changes are roughly 30% with some exceeding 40%, which greatly exceeds the changes in the majority of the objective measures of work performance in Table 11where most changes did not exceed 20%.

Table 15 Mean points for other self-reported indicators of work performance at the highest (40 dB(A)) and lowest (65 dB(A)) points with percentage ratios showing changes in other self-reported indicators of work performance.

		40 dB(A)	65 dB(A)	Difference	Difference ratio (Difference/mean at 40 dB(A)
Motivation	**BUS**	5.1	3.5	1.6	31%
	CBE	5.0	3.3	1.7	34%
Job satisfaction	**BUS**	5.2	3.7	1.5	29%
	CBE	5.1	3.7	1.4	27%
Fatigue	**BUS**	4.8	3.5	1.3	27%
	CBE	4.4	2.8	1.6	36%
Distraction	**BUS**	6.1	3.7	2.4	39%
	CBE	5.3	3.2	2.1	40%

4.2.1 Comparing Indicators of Work Performance obtained from surveys with Self-Reported Work Performance

Other self-reported indicators of work performance were gathered through survey questions and have been compared with self-reported work performance measures. Since the other self-reported indicators of work performance are all reported on a 7-point scale, plotting the results from these other self-reported indicators against self-reported work performance shows how the people in the BUS group have assessed their work performance differently from the CBE group. Figure 12 shows the comparison of the BUS and CBE groups' self-reported work performance correlated with the four indicators of work performance, motivation, job satisfaction, fatigue and distraction, and their trend lines and equations. For all four comparisons, the BUS group's trend line shows a steeper gradient of over double the value compared with the results from the CBE group.

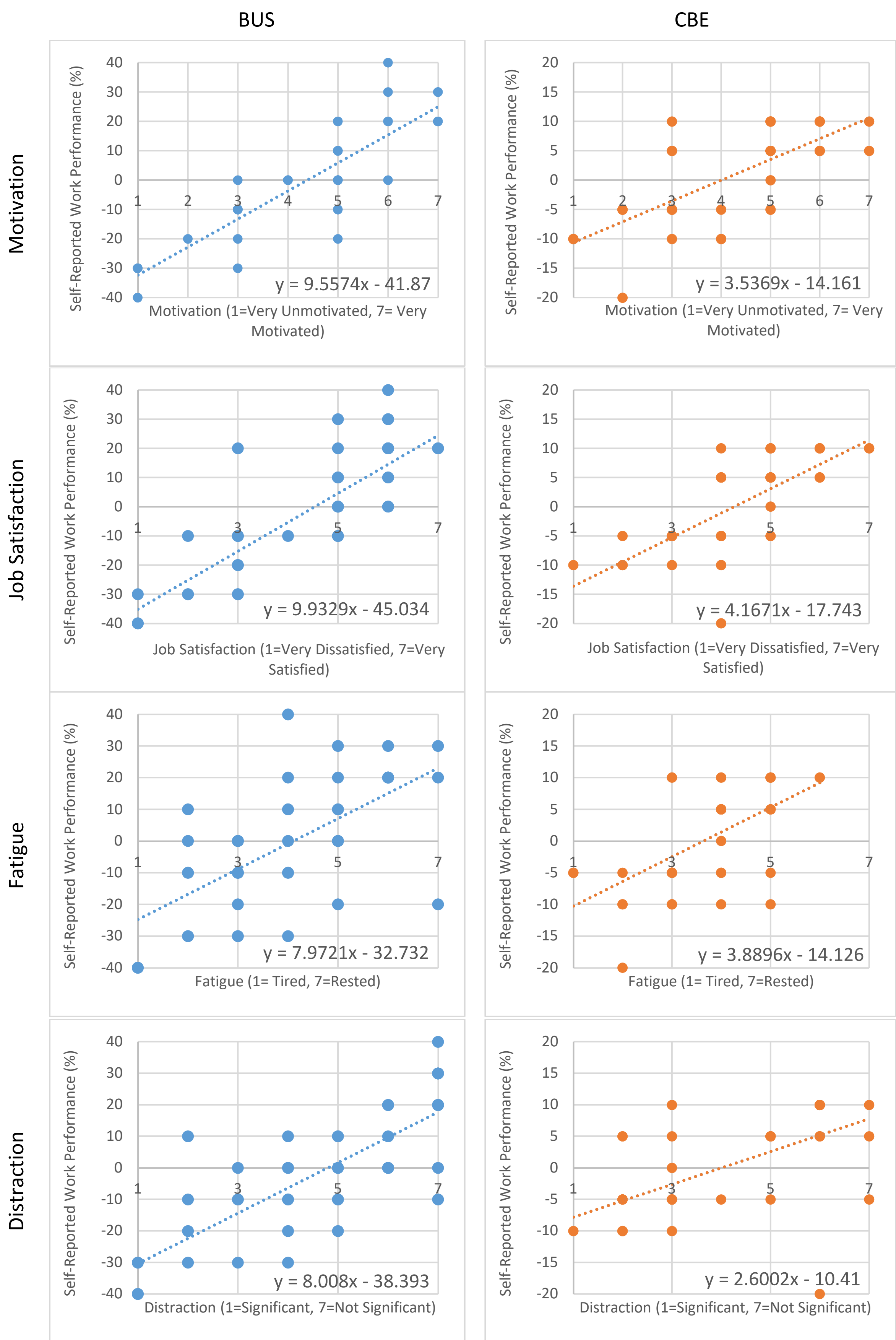


Figure 16 Self-reported work performance against the four indicators of work performance, the blue (left) showing the BUS and orange (right) showing CBE group.

The trend line gradient for BUS and CBE groups, their differences, the ratio of the differences, the margin of error (97.5%, one-tailed) and *p* value for the differences are shown in Table 16.

The ratio of the gradient differences shows ranges between 2.05 and 3.08, suggesting that the larger scale used for the BUS survey result in subjects assessing their work performance to be roughly double and up to three times larger than the CBE survey results. All the *p* values for the gradient differences are less than 0.05. Therefore the results are statistically significant, and it suggests that larger scale sizes could cause results to increase by two or even three times larger than when using scales that are half the size.

Table 16 Gradient for other self-reported indicators of work performance against self-reported work performance measures, with the differences, margin of error and p *value (all numbers rounded to 3 significant figures).*

	Gradient BUS	Gradient CBE	Gradient difference	BUS/CBE (Gradient)	Margin of error (97.5%)	*p* value
Motivation	9.56	3.53	6.02	2.71	1.80	5.54e-9
Job satisfaction	9.93	4.16	5.77	2.39	2.43	2.51e-5
Fatigue	7.97	3.89	4.08	2.05	3.30	0.0285
Distraction	8.00	2.60	5.41	3.08	2.16	4.83e-6

4.2.1.1 Testing for equivalence for each indicator of work performance

Figure 17 shows the equivalence testing for the gradient difference between the BUS and CBE groups' self-reported work performance results, correlated with the other self-reported indicators of work performance. The equivalence range used for this test was 0 ± 0.71, which is 5% (difference in work performance considered to be significant, refer to section 2.3.3.2) divided by 7 (7-point scale was used on the survey to assess and report changes of other self-reported indicators of work performance). For all tests, the gradient difference and the error bars do not fit within the equivalence range. These results suggest that the self-reported work performance assessments using the BUS and CBE scale sizes cause a significant difference in the results with the BUS scale generating larger reports of change in work performance than the CBE when correlated with other self-reported indicators of work performance.

From the equivalence test based on different environmental conditions described in Section 4.1.1.1, it is difficult to conclude that the BUS scales generates larger reports of change in work performance than the CBE survey scales. However, results from Figure 17 are sufficient to conclude that using BUS scales generates results that are significantly different to results from using CBE scales.

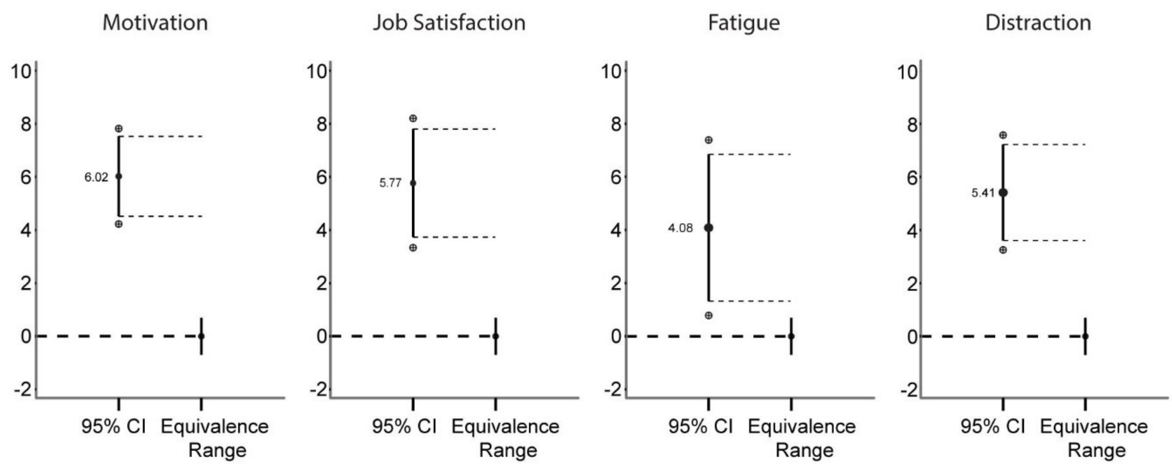


Figure 17 Equivalence testing for the gradient difference between BUS and CBE self-reported work performance measures, based on the other self-reported indicators of work performance. From left, motivation, job satisfaction, fatigue, and distraction.

4.2.2 Comparing Cognitive Performance with Self-Reported Work Performance

To find out whether the larger BUS scale and the smaller CBE scale produce a difference in 'productivity' results that is correlated with the objective measures of work performance, the two self-reported work performance results from BUS and CBE were compared with cognitive performance test results. Figure 18 shows the mean cognitive performance of all subjects for the nine tests conducted in the experiment. There is no consistency in the changes in cognitive performance, with some tests even showing increases in performance in the loudest noise environment (65 dB(A)). Tests such as letter search, overlapping, picture recognition, symbol digit modalities test and conditional reasoning show inconsistent trends over the three noise conditions. Memory Span, Spatial Image and Visual Choice Reaction Time show an increase, and Number Calculation shows a decreasing trend. Of the nine tests, Letter Search, Overlapping, Memory Span, Number Calculation, Conditional Reasoning and Visual Choice Reaction Time show no obvious change or difference in cognitive performance between some noise conditions. Results of cognitive performance test are summarised in Section 6 Appendix F.

The experiment was designed on the assumption that noise causes annoyance. Pink noise has been chosen as there is less high-frequency noise and therefore it is less high-pitched than white noise, which is known to cause some noise problems (Keighley and Parkin, 1979, cited in Al Horr et al., 2016). However, since properly designed pink noise is used for masking, sudden noises or other distracting noises could have been masked, which could have increased the cognitive performance in the noisy environment.

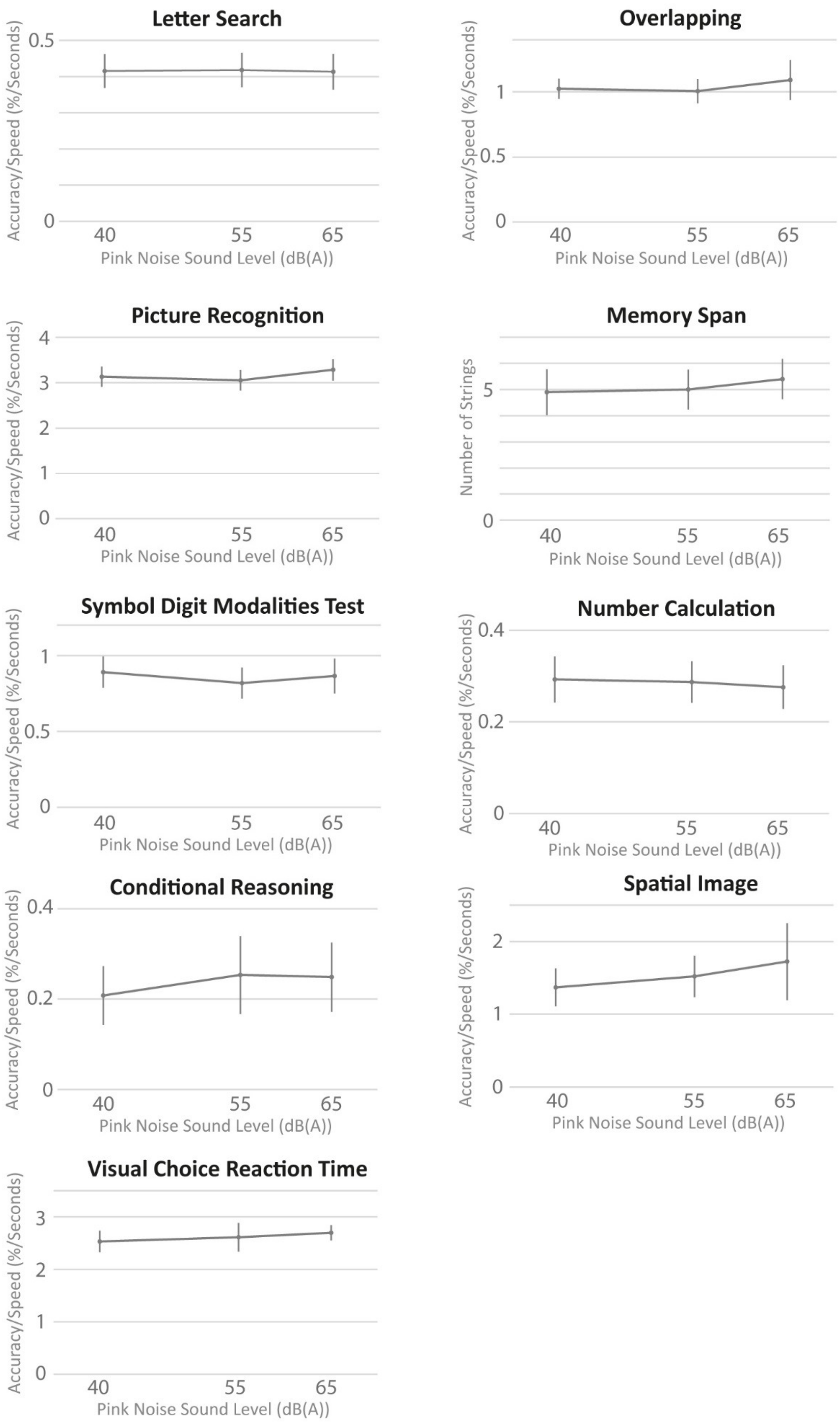


Figure 18 Mean cognitive performance of all subjects for all tests, over different noise level conditions.

4.2.3 Cognitive Performance against Self-Reported Work Performance

Figure 19 shows the cognitive performance results plotted against self-reported work performance for BUS and CBE survey groups. The positive and negative correlation between cognitive performance and self-reported work performance are not consistent between the two groups, and therefore the two trends cannot be analysed to conclude whether different survey scale sizes cause differences in the results.

There is some published evidence of a relationship between cognitive work performance and actual office work performance, as discussed in section 1.5. If cognitive performance represents a section of office work performance, and if self-reported work performance does not show a correlation with cognitive work performance, what are the subjects actually assessing when they are reporting their work performance?

What does self-reported work performance represent if this self-assessment does not correlate well with tests of cognitive performance, which are commonly assumed to be the human behavioural mechanism that are the drivers of work performance?

Of course, this assumes that there is a strong relationship between cognitive performance and office performance. The results could also suggest that either (1) cognitive performance tests do not represent office performance at all, or (2) that self-assessed work performance measures must bear little relationship to actual performance.

It is worth noting that for this particular test – the change in performance under pink noise – the masking effect could have had positive effects on simple tasks such as cognitive performance tests but cause negative effects for complex tasks such as report writing and designing. This may have caused the observed inconsistencies between self-reported work performance and cognitive performance.

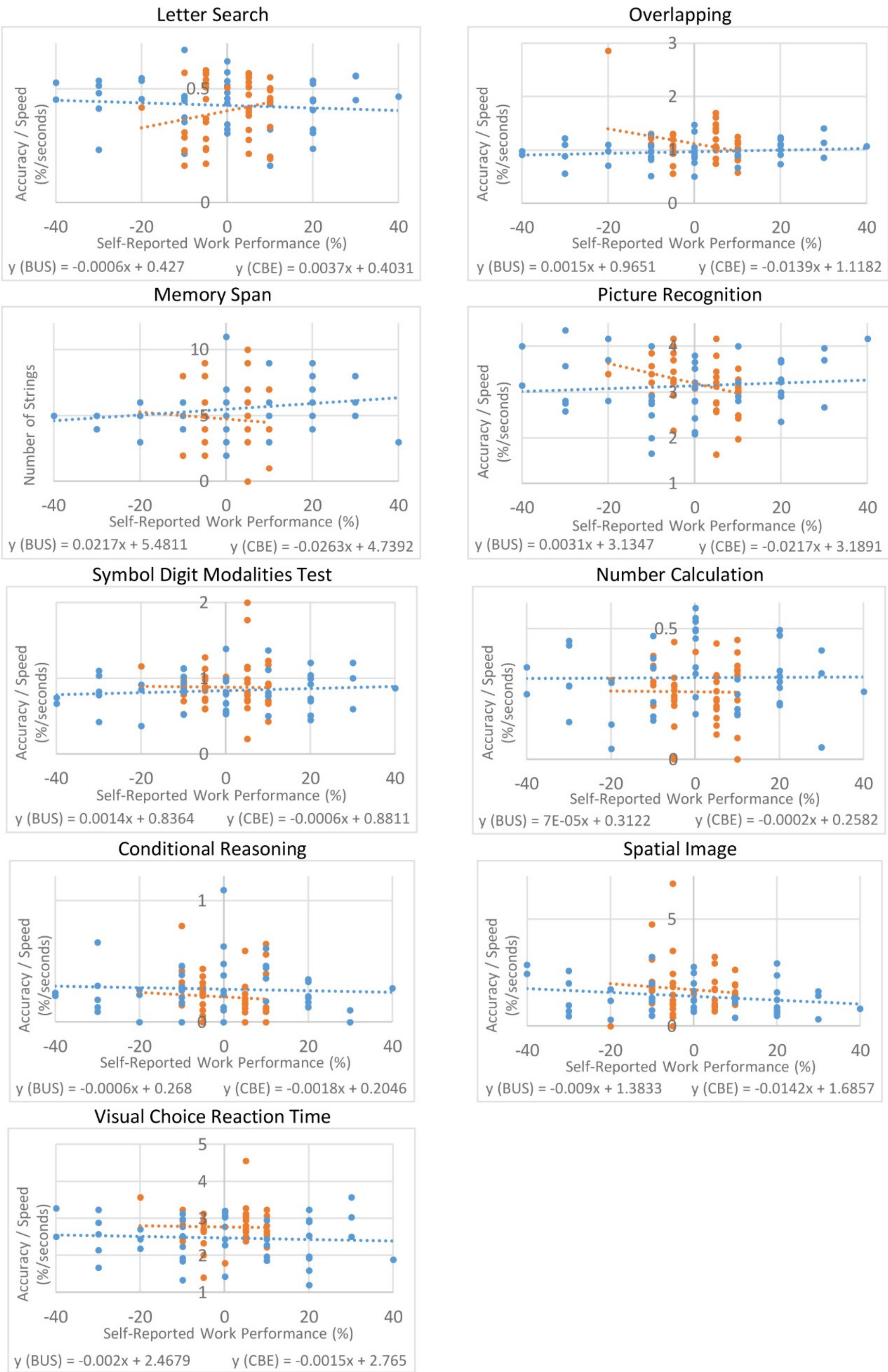


Figure 19 Cognitive performance results against self-reported work performance for each test with trend line for each survey group. Blue shows BUS and orange show CBE survey group results.

4.3 Subject Characteristics

4.3.1 Does Working in an Unfamiliar Work Environment Affect how People Assess the Environmental Conditions on their Work Performance?

As mentioned in Section 3.2.4, not all subjects usually worked in the room that was being used for the experiment. Therefore, people from other rooms could have had a different perception of the room in comparison to their usual working space, which could potentially have caused a difference in how they assess the environmental effect on their work performance. Sullivan, Baird and Donn (2013) mention the difficulty people have when assessing the environmental effect, rather than just assessing their work performance.

The question of the change in work performance with the noise level for subjects who are familiar and unfamiliar with the experiment room has been summarised in Table 17.

Although the gradients between participants who are familiar and unfamiliar with the experiment room shows some differences, all of the differences in numbers are not statistically significant. In the total of 30 subjects, 12 subjects were unfamiliar with the experiment room. An experiment which included sufficient numbers of subjects who are unfamiliar with the work environment could be used to assess whether it is difficult for people from different work environments to assess their work performance. The 12 subjects who are unfamiliar with the environment were divided equally across each of the BUS and CBE groups. The slight differences between the two groups indicate that the inclusion of these 12 in this manner is unlikely to have altered the basic conclusions, though it may have contributed to a slightly broader spread of response.

Table 17 Gradient differences for self-reported work performance, and other survey-obtained indicators of work performance between subjects who are familiar with the experiment room and subjects who are unfamiliar with the experiment room with the margin of error and p value (all numbers rounded to 3 significant figures).

	Gradient familiar	Gradient unfamiliar	Gradient difference	Margin of Error (97.5%)	*p* value
BUS Self-Reported	−0.848	−0.623	0.225	1.07	0.688
CBE Self-Reported	−0.450	−0.434	0.0161	0.400	0.935
Motivation	−0.0614	−0.0675	0.00614	0.0685	0.860
Job Satisfaction	−0.0687	−0.0434	0.00253	0.0590	0.394
Fatigue	−0.045	−0.0776	0.0326	0.0543	0.238
Distraction	−0.100	−0.0798	0.205	0.0700	0.566

4.3.2 Does Having Different Tasks Every Sequence Affect How People Assess the Effect of Environmental Condition on their Work Performance?

Also mentioned in Section 3.2.4, most subjects had been working on their thesis for around six months at the time when the experiment was conducted. However, there were a few subjects who were working on their university coursework during their experiment. If people require skills to perform their tasks well to critically assess their performance (Kruger and Dunning, 1999, cited in Sullivan, Baird and Donn) and if people tend to estimate their performance differently depending on their strengths (Sullivan, Baird and Donn, 2013), it seemed possible that subjects working on their thesis could assess their work performance differently from subjects working on their coursework.

It is debatable whether all subjects working on their thesis know about their work well enough to make critical assessments about their performance. However, they would have more knowledge about their work than subjects completing their coursework, who had been working on their courses for roughly one to two months when the experiment was conducted. Also, since there were several papers and assignments for students working on their coursework, their work could change between the three different experiment sequences. This could also have an effect on how the subject assessed the environmental effects on their work performance, as subjects could perceive their work performance to have increased for coursework they are good at and decreased for coursework they are not good at. The question about work performance asks the subjects to assess the environmental effect on their work performance rather than just their work performance, but how the subjects perceive their performance on different tasks could also affect how they assessed the environmental effect on their work performance.

This issue was tested by comparing the gradient of the change in work performance from subjects working on their thesis with that for subjects working on university coursework. The results of correlating these gradients with the noise levels have been summarised in Table 18. As with the comparisons in section 4.3.1, none of the differences show statistically significant results. There were six subjects working on coursework. Again it seems likely that more subjects would have to be tested if the analysis were to conclude whether people who are working on their work for longer assesses their work performance differently from people who have only been working on their tasks for one to two months. The six subjects working on their coursework have been equally divided between the BUS and CBE groups. Like the difference between subjects who were used to or new to the work environment, the effect of these 12 subjects may only have slight effects to the results, however, they could have also created a wider spread in the results.

Table 18 Gradient differences for self-reported work performance, and other survey-obtained indicators of work performance between subjects who are working on their thesis and subjects who working on their coursework with the margin of error and p value (all numbers rounded to 3 significant figures).

	Gradient Thesis	Gradient Course Work	Gradient difference	Margin of Error (97.5%)	*p* value
BUS self-reported	−0.838	−0.439	0.399	1.40	0.558
CBE self-reported	−0.263	−0.430	0.167	0.458	0.492
Motivation	−0.0689	−0.0439	0.0250	0.0820	0.557
Job satisfaction	−0.0601	−0.0526	0.00746	0.0852	0.838
Fatigue	−0.0539	−0.0746	0.0206	0.0702	0.543
Distraction	−0.0976	−0.0702	0.0274	0.0798	0.531

4.3.3 Perception of Survey Questions

After the experiment, comments were sought about how the subjects answered the survey questions. The subjects were confused as to how they were required to assess the change in their work performance due to environmental conditions, whether it should be based on best condition or worst condition or usual work environment. Subjects assessing the change in their work performance based on their best performance will assess their work performance to decrease or remain at 0%, whereas subjects assessing based on their worst performance will only report increases for the better environmental conditions. If subjects were assessing their work performance based on their usual environment, they would have assessed their work performance to have increased or decreased, but this assessment will also change according to where the subjects are usually working as some subjects do not usually work in the room the experiment was conducted in.

This problem refers to the difficulty for occupants to assess and relate their work performance based on other reference points (Leaman and Bordass, 1999), as already mentioned in Section 2.4.1.2. Most of the comments about the survey were made after or towards the end of all the experiment sequences, and therefore there were no consistencies with how subjects assessed their work performance. However, since the questions used in both BUS and CBE surveys do not state what the building occupants should base their reference points on, making a clear statement on how the subjects should report their work performance in the experiment would create consistency but differ from how environmental effects on work performance is currently assessed in real office buildings.

4.3.4 Some People Are Aware of the Order of the Noise Condition

Since each experiment sequence was conducted on different weekends over eight weeks in order to spread the environmental annoyance effect (refer to the Latin-square method in 3.2.2) and also due to recruitment problems, some subjects were aware of the different noise conditions they were about to be tested in. The subjects were also notified of the different noise conditions from the participant information sheet provided during recruitment. Therefore, a number of subjects who were aware of the different noise conditions have assessed their work performance by considering that there would be better or worse environmental conditions yet to be tested. This stopped the subjects from reporting their work performance to be a maximum decrease (−40% for BUS and −20% for CBE survey) even if they perceived their work performance to have decreased significantly, because the subjects knew that the noisiest environmental condition was yet to be tested.

This comment was also made towards the end of the experiment, and because the noise conditions to be tested were already notified in the information sheet, it is not known how many people assessed their work performance by considering that noisier environmental conditions were yet to be conducted. Also, because over half of the subjects usually work in the experiment room, it is difficult to prevent them from working in the room outside their experiment sequence times. Although this could have had some effect on the results, this could be seen as negligible, because only a limited number of subjects were in the experiment room at the weekend, outside the experiment times. However, if the experiment were to be conducted using usual workspaces during occupied days and hours, not notifying the degree of environmental stress and number of sequences could reduce this effect.

4.3.5 Annoyance of Pink Noise

Some subjects commented on how pink noise enhances their focus; some subjects reported annoyance with the pink noise and some commented how pink noise was annoying at the start but is easy to adjust to. To find out whether people who are comfortable and adjusted to pink noise have reported and performed better than people who have reported annoyance with the pink noise, distraction was compared with cognitive performance.

Self-reported work performance shows a positive correlation with distraction in Section 4.2.1 for both BUS and CBE groups. However, the correlation between distraction and cognitive performance shows opposing trends between the two groups for some tests as shown in Figure 20. Analysing this along with the mean cognitive performance for all the subjects on Figure 18, suggests that pink noise affects subjects’ cognitive performance differently between subjects and between tests. It shows increasing trends in one group or test while showing negative trends in other groups or tests.

However, self-reported work performance for their complex tasks consistently shows decreasing measurements in the louder noise conditions. Other environmental conditions that will have negative (or positive) effect on both complex and simple tasks could be considered to analyse the correlation between self-reported work performance against other subjective measures as well as providing an objective measure of work performance.

From a number of comments, pink noise has enhanced some subjects' work performance by masking out other occupants' conversations, but even so, some subjects still reported annoyance with pink noise. Controlling this would exclude people who are not annoyed with pink noise from the experiment. However, this could be argued to not represent real-life situations, as some people could be more resistant to certain environmental stressors. Also, since the sound level in the room is less than 40 dB(A) during the weekends, less noise from people walking, doors opening and closing, and clattering noise from keyboards could cause more annoyance for the occupants, which the pink noise was masking and so reducing the annoyance.

This could also link back to the physical and functional comfort discussed in section 2.3.2. For this experiment, the optimum comfort for cognitive performance could have been the noisy conditions. This could be caused by people working with greater focus in the noisy environment, which enhanced their cognitive performance. However, people report higher distraction in the noisier environment while some of the cognitive performance tests show higher results. This raises several questions, such as:

- Would people's cognitive performance also increase with other environmental stressors?
- Did cognitive performance increase for some tests and some subjects because they were exposed to the environment for only a short time?
- Is the optimum environment for conducting simple tasks not necessarily comfortable compared with the environment required for physical comfort?

Using other environmental stressors that will affect both complex and simple task performances negatively could potentially generate results that enable comparisons between self-reported work performance and cognitive performance. This comparison could be used to examine how survey scale sizes affect how people assess the environmental effect on their work performance, by using an objective measurement of work performance.

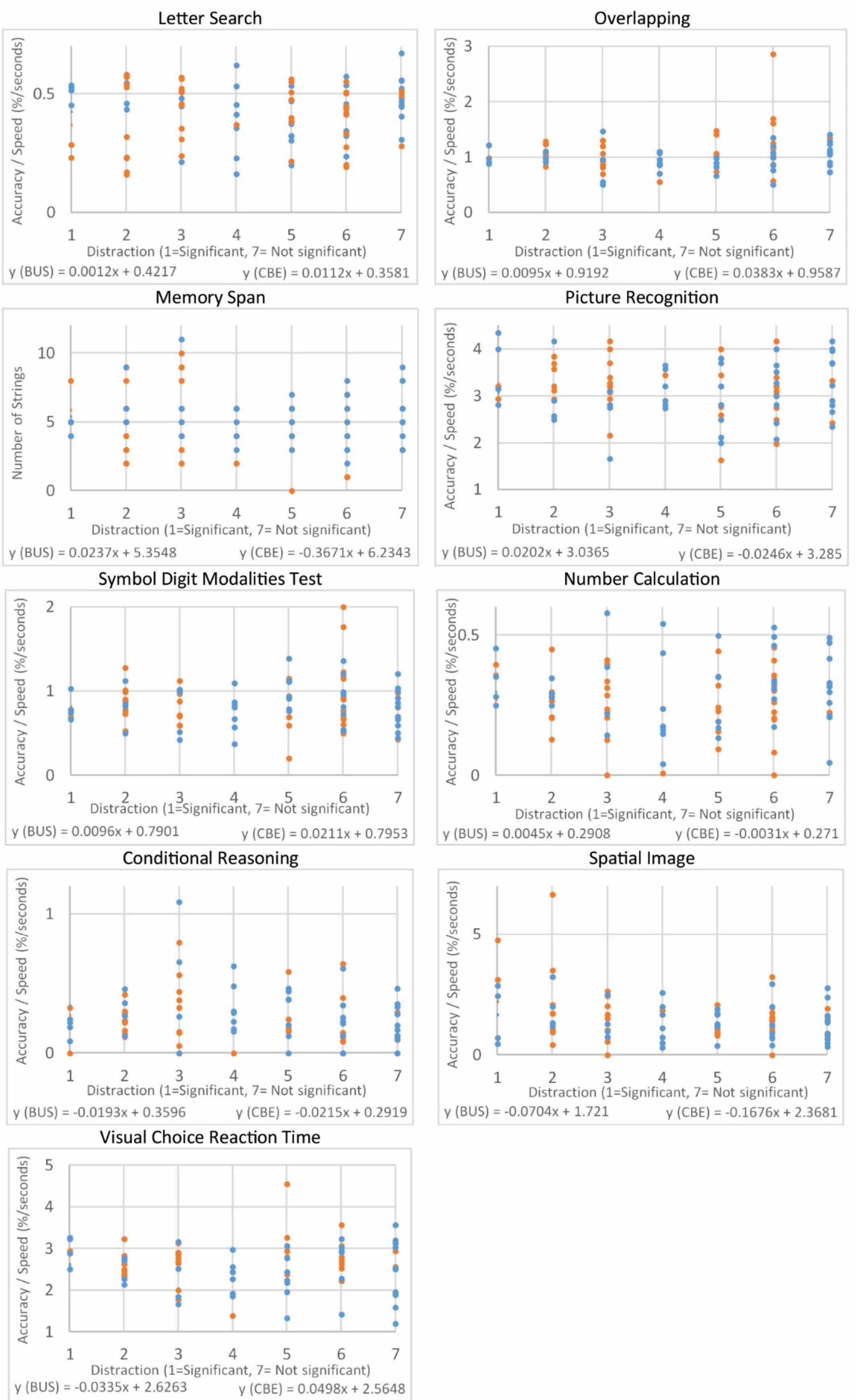


Figure 20 Cognitive performance against distraction for all nine cognitive performance tests with BUS group (blue) and CBE group (orange).

4.3.6 Summary: Does Controlling these Subject Characteristics Affect the Results?

To summarise the subject characteristics:

1. The subject characteristics are whether they are familiar or unfamiliar with the experiment room, and whether subjects are working on their thesis or on their coursework; it is difficult to conclude whether these factors caused a major difference in the results.
2. Setting a reference point for people having different perceptions about the survey question would create consistency and could change the results, but should not be stated, because BUS and CBE surveys do not set a certain reference point in their surveys.
3. Although some subjects were aware of the noise conditions, which led them to think about how to assess, due to the noise conditions yet to be tested, there was a very limited number of people who participated in the experiment that also came during the weekends outside their experiment times. Therefore, this had little impact on the results as not many people would have known the order of the noise levels that were to be tested.
4. Pink noise does seem to cause a distraction for most people during their complex tasks (thesis or university course work) but differs for their simple tasks (cognitive tests). It would be better in future follow up work to use environmental stressors that cause stress on both complex and simple tasks to be able to make comparisons between subjective and objective measurements.

It is difficult to conclude that these different subject characteristics caused a difference in results. As mentioned in section 3.2.4, the research design was intended to minimise this issue as students of different rooms and years were equally divided (point 1) between the BUS and CBE groups. Other characteristics are thought to have minimal effect, because setting a reference point will not represent the real-life situation (point 2), and the effect could be seen as negligible considering the number of people likely to experience e the same effect (point 3). For point 4, it would be better to use other environmental stressors. However, the only restriction from using pink noise for this experiment was the comparison between the subjective and objective measure of work performance which also restricts the assessment of how the different survey scale sizes had an effect.

4.4 Considering the Results, Do Scale Sizes Affect How People Assess the Environmental Conditions on their Work Performance?

To summarise:

1. Effect of scale size on how people assess the environmental effect on their work performance.

- The overall change in work performance between quiet condition (40 dB(A)) and loudest condition (65 dB(A)) is 18.67% for the BUS group and 11% for the CBE group. The difference of 7.67% is greater than ±5%, and therefore the difference is considered likely to lead to significantly different design decisions.
- The self-reported work performance trend over the range of noise levels shows that the BUS trend line is 70% greater than the CBE trend line.
- The trend-line gradient difference between BUS and CBE shows *p* value of 0.28, and therefore the differences were not statistically significant due in large part to the far wider spread of responses for the BUS group.

2. Correlation between other self-reported indicators of work performance and self-reported work performance results using different scale sizes.
    - Comparing the trends for self-reported work performance between the BUS and CBE scales shows that using BUS generates results at least twice and up to three times higher than using CBE scales when correlated with how subjects have assessed other self-reported indicators of work performance (motivation, job satisfaction, fatigue, distraction). This is a statistically significant result as the *t* test for all the comparisons made with other self-reported indicators shows *p* values of less than 0.05.
    - Trend-line gradient difference (equivalence test) between BUS and CBE shows BUS to be larger, with error bars showing no overlaps with zero. Therefore, using BUS scales generates results that are significantly larger than using CBE scales, when correlated with other self-reported indicators of work performance.
3. Using correlation with cognitive performance as an objective measure of work performance, to examine the effect of scale size differences.
    - Since pink noise is assumed to have both a positive and negative effect on work performance, results cannot be plotted against self-reported work performance as positive and negative gradients cannot be compared to examine how one is greater than the other.

The analysis from point 2 shows that using BUS scales generates larger results and analysis 1 also leans towards BUS trend lines, showing a greater trend than CBE trend lines. Analysis 1 would generate data on the magnitude of the difference between BUS and CBE, which would allow comparisons of past BUS and CBE surveys. Recruiting more subjects for the experiment would allow comparisons of past studies, which could provide information on designs and strategies to increase occupant work performance. Also, using environmental stressors to affect both complex and simple tasks, in the same manner, would allow self-reported work performance to be plotted against cognitive performance to examine the question using objective measures or work performance.

4.4.1 What Could Have Caused the Significant Differences in Self-reported Work Performance by Using Two Different Survey Scale Sizes?

The trend-line gradient for BUS is calculated to be 1.7 times the CBE trend line in 4.1, but the BUS group has assessed the change in their work performance to be at least 2× and up to 3× in section 4.2. One aspect of the survey scale that could explain the significantly larger results in BUS compared to CBE is the number of points on the scale. CBE survey scale size in terms of magnitude of change is half of the BUS survey scale but is only 7 points compared to the BUS 9-point scale. The BUS scales increases (or decreases into negative) in 10% increments whereas the CBE survey scale increases (or decreases) in 5% increments from −10% to +10%, but misses the 15% and jumps to 20%.

It seems likely that this last difference may affect how people assess the environmental effect on their work performance. Figure 21 shows the number of responses at each survey scale point for the BUS and CBE groups. This shows that the BUS groups have a wider range in how subjects have reported the environmental effect on their work performance than the CBE group, which has no responses of +20% and only 1 response at −20%. It seems from this that it is easier for people to report higher numbers on the scales if each point increases or decreases in consistent increments; along with this, it seems people tend to report a smaller number if there is a gap in the survey scale. The narrower range on the CBE scale seems implicated in the ×2 to ×3 difference in some comparisons. This suggests that the cause may not be the differing sizes of the two scales, but rather their structure.

As well as the survey scale having different points and having a different size, the minimum interval difference could also affect how people assess the change in their work performance. It seems likely that the interval size of the two scales could be an explanation of the finding that only one CBE response reported the change in their work performance to be 0%, whereas 10 responses of 0% change in work performance were reported in the BUS group. The CBE scale has a smaller minimum interval of ±5%, encouraging the reporting of small differences. The BUS minimum of ±10% appears likely to have inspired subjects to settle on 0% change in work performance as being closer to their reaction than ±10%.

In support of this interpretation, we have the observation that most responses from the CBE group have reported *at least some change* in work performance but there were fewer responses of *large changes* in work performance. Results therefore show BUS trend lines having a 1.7× greater gradient than CBE trend lines, because CBE subjects report at least some change in their work performance, but self-reported work performance correlated with other self-reported indicators shows the BUS trend line with ×2 to ×3 greater gradient because CBE subjects have not reported large changes in their work performance.

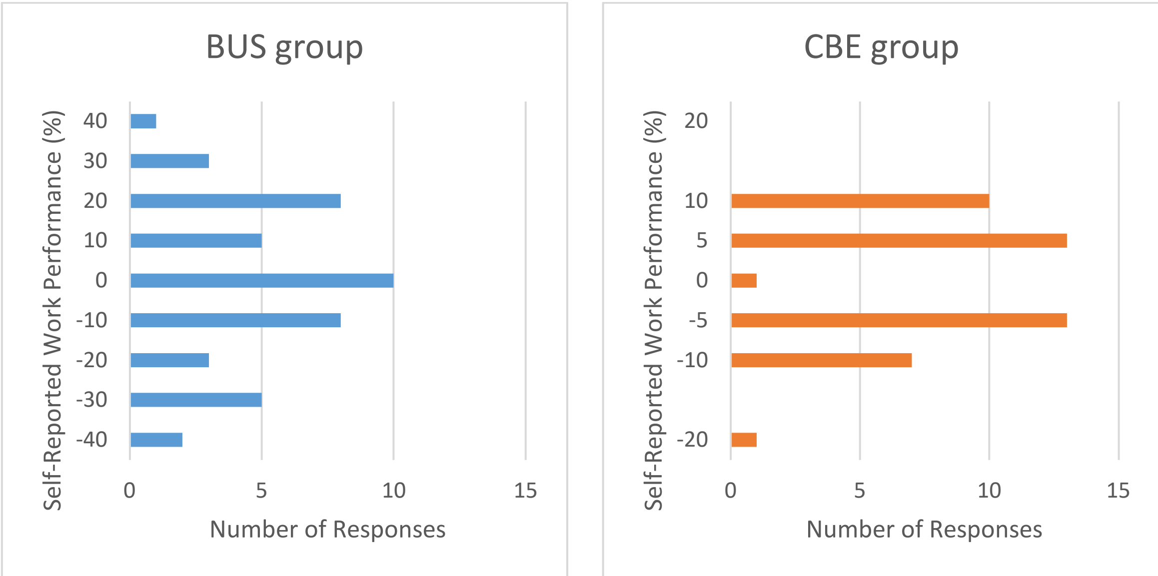


Figure 21 Number of responses at each point on the survey scale for BUS group (left) and CBE group (right)

5 Discussion and Conclusion

5.1 Discussion: Should surveys use the Larger BUS Scales or the Smaller CBE Scales to Measure the Environmental Effects on Work Performance?

Although it is difficult to conclude how much greater the measurements taken using BUS scales are than measurements taken using a CBE scale, there is enough data and results to conclude that these two scales do affect how people assess and report the effect of the environment on their work performance. For measurements of different survey results to be comparable, consistency between scales is required throughout the different types of surveys. However, the question remains: is the BUS scale or the CBE scale better for measuring the environmental effect on work performance? Unless this can be determined, then the common practice of estimating cost benefits from percentage productivity gains cannot be considered reliable.

5.1.1 BUS and CBE Survey Scale Measure Work Performance Differently

Both BUS and CBE scales do measure the environmental effect on people's work performance. However, these two scales measure work performance differently. Results from the BUS scales show larger changes in work performance, most likely not just because the scale size is double the size of CBE scale but also because each interval is in consistent 10% increments. The inconsistency in the intervals for the CBE survey could have reduced the number responses which assessed that the environment had the maximum effect on their work performance of ±20%, as discussed in section 4.4.1.

However, this does not necessarily mean that using the CBE scale generates smaller results for all situations. The number of responses showing a 0% change in work performance in the BUS group is 10 times higher than for CBE, but the CBE group does report at least some change in work performance. Therefore, CBE scales could generate larger results of a change in work performance than using BUS scales if the change in environmental conditions is small.

5.1.2 What Scale should be used to Measure how the Work Environment is Affecting Occupants' Work Performance?

Although there is still no general agreement about how self-assessed work performance measurements reflect changes in actual work performance, using consistent scales throughout the surveys could make results from different surveys comparable.

The results in section 4.2 from this particular study show that the scale range of ±40% could also be sensible, because changes in other self-reported indicators of work performance (Table 15) show changes of around 30–40%. However, a smaller scale could also be sensible due to the very small changes observed in cognitive performance. To create a new consistent scale, a ±20% scale size has been chosen rather than ±40% scale size, because studies from Table 11show the most changes in objective measures of work performance, such as simulated office tasks and cognitive performance of under 15%, a few around 20% and very limited number over 20%.

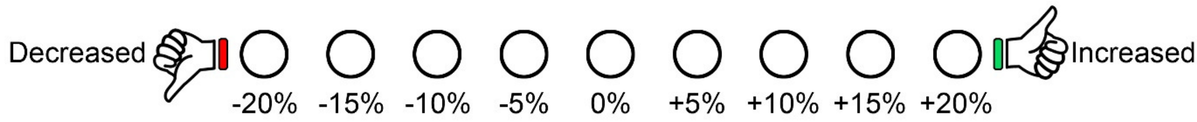


Figure 22 New 9-point scale proposed for consistency throughout the surveys and for results between different building surveys to be comparable.

The new scale, shown in Figure 22, combines the consistency of BUS with the size of CBE scale. Having a minimum interval of ±5% could make it easier for the occupants to assess *small* amounts of change in work performance. Also, including a ±15% that the CBE scale did not have, will allow the occupants to assess a change in work performance between 10% and 20%. It is unclear if all occupants are competent with their tasks to critically assess a *large change* in their work performance of ±20%. However, the new scale will create more consistency, as it seems sensible to have a mid-range step rather than requiring people to choose between exaggerating ±20% or minimising down to ±10%.

It is unclear if occupants are competent enough to critically assess the change in their work performance to be the maximum change of ±20%; however, it would create more consistency than people assessing a ±10% change just because ±15% does not exist and not assessing ±20% because the gap between ±10% and ±20% appears too large.

5.2 Conclusion

This study started with the research question of “How accurate is perceived work performance as an alternative measure of actual work performance?” and developed that research question through a literature review. The main research question was revised to “How do survey scale sizes influence how building occupants assess the indoor environmental effect on their work performance?”, to examine an aspect of perceived work performance that will affect its accuracy as an alternative measure of actual work performance.

An experiment has been conducted to examine how different self-assessment work performance scales from the BUS and CBE studies influence how people assess the environmental effect on their work performance. The experiment was conducted with pink noise as the main environmental stressor. The results show that subjects report a larger change in work performance due to environmental stressors using the larger BUS scale. However, the data does not show a statistically significant difference between the BUS and CBE scales that would support a conclusion that using larger scales necessarily generates greater changes in results. Comparison between the work performance scores from the BUS and CBE surveys when correlated to the other self-reported indicators of work performance shows that results from using BUS scales are at least 2× and up to 3× greater than results using the CBE scales, and these differences are statistically significant.

The number of responses at each point in the scale shows distributed responses in the consistent BUS scale, but there are fewer responses at the maximum change of ±20% and at 0% for the CBE scale. This is assumed to be caused by the CBE scale skipping ±15% and having a smaller minimum interval of 5%. Correlation with cognitive performance, to compare results from the BUS and CBE scales, could not be conducted. However, there is enough evidence to conclude that using BUS and CBE scales does generate very different results.

The new 9-pointscale from −20% to +20% proposed in this research could improve the measurement of the effect of a building's environment on work performance. Because research and case studies estimate potential benefits based on self-reported work performance (U.S. Environmental Protection Agency, 1989; Singh et al., 2010; Miller et al., 2009; Green Building Council of Australia, 2013), business decisions could be based on an increase in self-reported work performance of +10%. However, this could be as small as 3–5% if a smaller scale were used. Since different scale sizes are known to cause a difference in results, it is potentially incorrect to link self-reported work performance directly to actual work performance to calculate the hours of work and savings achieved through the increase in work performance.

The consistency of scales would be achieved by using one type of scale, but the problem still remains with inconsistencies between people on how they perceived the survey questions. Although it is difficult for people to assess their work performance based on other reference points (Leaman and Bordass, 1999), people will still use reference points to assess the changes in their work performance. However, people will have different reference points from each other (section 4.3.3), creating inconsistencies on how they assess the changes in their work performance.

Other inconsistencies include not being able to control other environmental stressors in the experiment room, exposure time of the subjects to the environmental stressors and the subjects' characteristics. Differences between the scales were observed; however, the relationship with cognitive performance remains unknown. The inconsistent changes in cognitive performance could potentially be caused by other environmental factors such as temperature, humidity, lighting and CO_2 levels, not just the noise. Also, as explained in section 4.3.5, the masking effect of pink noise could have increased the work performance and cognitive performance for a number of subjects, but this could show decreasing effects if the subjects were to be exposed for a longer time such as the eight hours of normal office hours.

In addition, other personal factors such as health, life events and other seasonal factors could affect how the subjects perform, which could lead to changes in how they perceive the environment to be affecting their work performance. However, it is impossible to control these factors and it would not be realistic to exclude subjects who are easily affected by personal situations. Effects from personal life events and seasonal factors could potentially be reduced by conducting the experiment in a shorter timeframe. However, this will lead to build-up of fatigue, and since the experiment room cannot be used during the weekdays, eight weeks were required to spread the learning effect (Latin square method). Also, as stated in Section 3.1.1, the phrasing of the questions could reduce the possibility of subjects being influenced by their personal situations when they assess the environmental effect on their work performance. The initial design of the experiment based on analysis of the standard deviations in the responses in previous experiments was intended to accommodate for this normal random distribution of other outside influences. The identification of the inadequacy of the sample size in this project suggests that these external influences are more important than the prior data indicated.

It is still unknown if self-reported work performance measures are an accurate measure of actual work performance. Also, knowing the ratio differences of the results from using BUS and CBE scales will allow comparison of the past survey results to study how building environments could improve work performance. However, the proposed new scale would at least provide some consistency in how occupants assess the environmental effect on their work performance.

5.3 Further Study

There are a number of questions still unanswered and improvements to the experiment that could be worthy of further research.

Recruiting more subjects for the experiment could prove the ratio difference between results from using BUS and CBE scales, as well as examining whether the subject characteristics have a significant effect on the results. The calculation for the number of subjects was based on SD values reported in studies by Brown et al. (2010), Baird, Leaman and Thompson (2012), Parkinson et al. (2017) Gou, Prasad and Lau (2013) and Gou, Prasad and Lau (2014), with simulations conducted to estimate the number of subjects required to obtain statistically significant data. To ensure that the data would be statistically significant, six more subjects than the minimum of 24 were recruited. All participants in the experiment environment were recruited by talking with them in person and handing out invitation cards, but only 18 agreed to participate. This required more people to participate from other rooms and years, which caused the range in the variety of people, possibly causing differences in the results. However, the SD showed greater variation in the results, particularly for this experiment, and therefore uncertainties with the result variations from different people could not be analysed. Although the experiment was designed to control these variations and obtain the data required to answer the main question of this thesis, for any new experiment based on this process the sample size simulation analysis suggests that around 44 subjects are required to calculate the ratio of BUS results against CBE results. To ensure that the results are statistically significant, a minimum of 50 or more subjects would be preferred to conduct further investigations.

Changing the environmental stressor would allow results from using different scale sizes in correlation with cognitive performance, an objective measure of work performance. People have reported high distraction for their usual tasks, but some cognitive performance has increased. Results from using pink noise also leave a number of questions that were not addressed in this research:

- Did the performance of usual tasks (complex tasks) really decrease?
- Would cognitive performance (simple tasks) decrease if subjects were exposed to pink noise for a longer time (e.g., full working hours of eight hours)?
- Would other environmental conditions also increase the performance of simple tasks if subjects were exposed for a short time?

Further investigation is required to observe how environmental stressors affect people's complex-task performance and simple-task performance, and whether short exposure to stressors could actually increase simple and possibly complex task performance. These questions will need to be answered before making a correlation between self-reported work performance and cognitive performance.

Although there are studies that argue the advantage of the environmental approach, the majority of the past studies have used the direct approach to obtain self-reported work performance measures (Sullivan, Baird and Donn, 2013). What is needed for making business decisions on improving the work environment is reliable data on how the environment is having an effect on occupants' work performance. The direct approach asks the occupants to assess their work performance rather than how the environment is having an effect, which could affect how occupants report changes in their work performance, as previously discussed. However, since the direct approach is already commonly used throughout the literature, it would be worth investigating if people would assess the change in their work performance to be different from using the two different questions. This could be conducted similarly but dividing the group of subjects into two with one group using the direct approach and one group using the environmental approach.

Finally, there is a need to investigate how self-reported work performance correlates with actual work performance. Using the same scale throughout the surveys will achieve consistency but it is not yet clear whether self-reported work performance measures should be directly linked to actual performance as a basis for investment decisions to improve the indoor environment. Studies are available that compare self-reported work performance to simulated office tasks and cognitive performance but it is arguable whether these measures really represent actual work performance. Studies comparing quantified actual work performance against self-reported work performance would be very useful, to observe how these two correlate. However, easily quantifiable call centres have simple tasks that could not be easily applicable to other offices with different work natures. It is difficult to quantify quality, which would also be vague, and to weigh the importance of work for the complex office buildings. Finding the correlation between self-reported work performance and actual performance will lead to improving building designs in general to suit the occupants and increase work performance.

6 References

Adamson, D. (2004). Design as a value generator for clients and society. *Building Research & Information*, [online] 32(3), p. 5. Available at http://www.tandfonline.com/doi/abs/10.1080/0961321042000221098 [Accessed 16 Aug. 2017].

Agha-Hossein, M.M. El-Jouzi, S. Elmualim, A.A. Ellis, J. and Williams, M. (2013). Post-occupancy studies of an office environment: Energy performance and occupants' satisfaction. *Building and Environment*, [online] 69, p. 10. Available at http://www.sciencedirect.com/science/article/pii/S0360132313002217?via%3Dihub [Accessed 26 May. 2017].

Al Horr, Y. Arif, M. Kaushik, A. Mozroei, A. Katafygiotou, M. and Elsarrag, E. (2016). Occupant productivity and office indoor environment quality: A review of the literature. *Building and Environment*, [online] 105, p. 21. Available at: http://www.sciencedirect.com/science/article/pii/S0360132316302001 [Accessed 10 Apr. 2017].

Baird, G. Leaman, A. and Thompson, J. (2012). A comparison of the performance of sustainable buildings with conventional buildings from the point of view of the users. *Architectural Science Review*, [online] 55(2), p. 10. Available at http://www.tandfonline.com/doi/abs/10.1080/00038628.2012.670699 [Accessed 26 May. 2017].

Balazova, I. Clausen, G. Rindel, J. Poulsen, T. and Wyon, D. (2008). Open-plan office environments: A laboratory experiment to examine the effect of office noise and temperature on human perception, comfort and office work performance. *Indoor Air*, [online] 703, p. 8. Available at: https://www.researchgate.net/profile/Geo_Clausen/publication/265991981_Open-plan_Office_Environments_A_Laboratory_Experiment_to_Examine_the_Effect_of_Office_Noise_and_Temperature_on_Human_Perception_Comfort_and_Office_Work_Performance/links/54b51e660cf28ebe92e4c291.pdf [Accessed 2 May 2017].

Bhawani, S. (2010). *Post occupancy evaluation: Development of an instrument and a process to access occupant satisfaction in renovated university office settings: A case study approach.* MSc. Michigan State University.

Biery, M. (2016). The Most Profitable Industries in 2016. [online] Forbes. Available at https://www.forbes.com/sites/sageworks/2016/08/06/the-most-profitable-industries-in-2016/#1ecc54ea2557 [Accessed 6 Dec. 2017].

Boerstra, A. Kulve, M. Toftum, J. Loomans, M. Olesen, B. and Hensen, J. (2015). Comfort and performance impact of personal control over thermal environment in summer: Results from a laboratory study. *Building and Environment*, [online] 87, p.12. Available at: http://dx.doi.org/10.1080/00038628.2012.744298 [Accessed 8 Apr. 2016].

Brocolini, L. Parizet, E. and Chevret, P. (2016). Effect of masking noise on cognitive performance and annoyance in open plan offices. *Applied Acoustics*, [online] 114, p. 12. Available at http://www.sciencedirect.com/science/article/pii/S0003682X16302067 [Accessed 12 Jun. 2017].

Brown, Z. Inalhan, G. Cole, R. J. Robinson, J. and Dowlatabadi, H. (2010). Evaluating user experience in green buildings in relation to workplace culture and context. *Facilities*, [online] 28(3/4), p. 14. Available at http://www.emeraldinsight.com/doi/full/10.1108/02632771011023168 [Accessed 17 Jun. 2017].

Burson , K. A. Larrick, R. P. and Klayman, J. (2006). Skilled or unskilled, but still unaware of it: How perceptions of difficulty drive miscalibration in relative comparisons. *Journal of Personality and Social Psychology*, [online] 90(1), p. 18. Available at https://jdc325.files.wordpress.com/2009/06/bursonlarrickklayman.pdf [Accessed 2 Aug. 2017].

Byrd, H. and Rasheed, E. (2016). The Productivity Paradox in Green Buildings. *Sustainability; Basel*, [online] 8(4), p. 13. Available at: http://search.proquest.com/docview/1779945776/abstract/E88CFE3445F94F90PQ/3 [Accessed 10 Apr. 2017].

Choi, J. Loftness, V. and Aziz, A. (2012). Post-occupancy evaluation of 20 office buildings as basis for future IEQ standards and guidelines. *Energy and Buildings*, [online] 46, p. 9. Available at: http://www.sciencedirect.com/science/article/pii/S0378778811003434 [Accessed 10 Apr. 2017].

Clausen, G. and Wyon, D. (2008). The Combined Effects of Many Different Indoor Environmental Factors on Acceptability and Office Work Performance. *HVAC&R Research*, [online] 14(1), p.11. Available at: http://www.tandfonline.com/doi/abs/10.1080/10789669.2008.10390996 [Accessed 2 May. 2016].

Clements-Croome, D. (2000). *Creating the Productive Workplace*. London: E&FN Spon, p. 384.

Clements-Croome, D. (2005). Designing the Indoor Environment for People. *Architectural Engineering and Design Management*, [online] 1(1), p.11. Available at: http://dx.doi.org/10.1080/17452007.2005.9684583 [Accessed 10 Apr. 2017].

Clements-Croome, D. (2011a). The Interaction Between the Physical Environment and People. [online] In: S.A. Abdul-Wahab, ed., *Sick Building Syndrome in Public Buildings and Workplaces*. Berlin: Springer-Verlag, pp. 239-259. Available at https://link.springer.com/content/pdf/10.1007/978-3-642-17919-8_13.pdf [Accessed 16 Aug. 2017].

Clements-Croome, D. (2011b). Sustainable intelligent buildings for people: A review. *Intelligent Buildings International*, [online] 3(2), p. 20. Available at http://www.tandfonline.com/doi/abs/10.1080/17508975.2011.582313 [Accessed 16 Aug. 2017].

Colliers International (2016). *Workplace Report // 2016*. [online] Auckland: Colliers International New Zealand, p. 16. Available at http://www.colliers.co.nz/find%20research/specialty%20reports/new%20zealand%20workplace%20report%202016/ [Accessed 30 Nov. 2017].

Concha-Barrientos, M. Compbell-Lendrum, D. and Steenland, K. (2004). Occupational noise. *Environmental Burden of Disease*, Geneva: World Health Organisation, [online] 9, p. 41. Available at http://www.who.int/quantifying_ehimpacts/publications/en/ebd9.pdf [Accessed 16 Jun. 2017].

Cui, W. Cao, G. Park, J. Ouyang, Q. and Zhu, Y. (2013). Influence of indoor air temperature on human thermal comfort, motivation and performance. *Building and Environment*, [online] 68, p. 9. Available at: http://www.sciencedirect.com/science/article/pii/S036013231300190X [Accessed 16 Apr. 2016].

Dykes, C. (2012). *Use Perception Benchmarks for Commercial and Institutional Buildings in New Zealand*. MBSc. Victoria University of Wellington.

Fang, L. Wyon, D. Clausen, G. and Fanger, P. (2004). Impact of indoor air temperature and humidity in an office on perceived air quality, SBS symptoms and performance. *Indoor Air*, [online] 14, p. 8. Available at: http://onlinelibrary.wiley.com/doi/10.1111/j.1600-0668.2004.00276.x/abstract [Accessed 20 Apr. 2017].

Feige, A. Wallbaum, H. Janser, M. and Windlinger, L. (2013). Impact of sustainable office buildings on occupant's comfort and productivity. *Journal of Corporate Real Estate*, [online] 15(1), p.28. Available at: http://search.proquest.com.helicon.vuw.ac.nz/docview/1366685087/abstract/F2624E59FED74EA2PQ/3 [Accessed 8 Apr. 2016].

Frontczak, M. (2011). *Human comfort and self-estimated performance in relation to indoor environmental parameters and building features*. PhD. Technical University of Denmark.

Fullbrook, D. and Jackson, Q. (2005). *Value Case for Sustainable Building in New Zealand.* Wellington: Ministry for the Environment, p. 80.

Gou, Z. Prasad, D. and Lau, S. (2013). Are green buildings more satisfactory and comfortable? *Habitat International*, [online] 39, p. 6. Available at http://www.sciencedirect.com/science/article/pii/S0197397513000064 [Accessed 17 Jun. 2017].

Gou, Z. Prasad, D. and Lau, S. (2014). Impacts of green certifications, ventilation and office types on occupant satisfaction with indoor environmental quality. *Architectural Science Review*, [online] 57(3), p. 11. Available at http://www.tandfonline.com/doi/full/10.1080/00038628.2014.908113 [Accessed 17 Jun. 2017].

Green Building Council of Australia. (2013. *THE BUSINESS CASE FOR GREEN BUILDING.* [pdf] New South Wales: Green Building Council of Australia. Available at https://www.gbca.org.au/uploads/63/34623/Evolution_2013_Business_Case_for_Green_Building.pdf [Accessed 17 Jan. 2018].

Hannah, L. Page, W. and McLaren, S. (2016). A Review of Occupational Noise in New Zealand. *New Zealand Acoustics*, [online] 29(1), p. 18. Available at http://www.acoustics.org.nz/journal/pdfs/Hannah_L_ANZ2016.pdf [Accessed 16 Jun. 2017].

Haynes, B. (2007). An evaluation of office productivity measurement. *Journal of Corportate Real Estate; Bingley*, [online] 9(3), p. 12. Available at: http://search.proquest.com/docview/233418280/abstract/E88CFE3445F94F90PQ/20 [Accessed 8 Apr. 2016].

Haynes, B. (2008). The impact of office comfort on productivity. *Journal of Facilities Management*, [online] 6(1), p. 15. Available at: http://search.proquest.com.helicon.vuw.ac.nz/docview/218926803/abstract/F2624E59FED74EA2PQ/1 [Accessed 8 Apr. 2016].

Hongisto, V. Varjo, J. Leppamaki, H. Oliva, D. and Hyona, J. (2016). Work performance in private office rooms: The effects of sound insulation and sound masking. *Building and Environment*, [online] 104, p. 12. Available at: http://www.sciencedirect.com/science/article/pii/S0360132316301433 [Accessed 17 Apr. 2017].

Hudson, (2015). *Hudson Salary Guides 2015*. [online] Available at http://nz.hudson.com/portals/nz/documents/salary%20guides/2015/SalaryTables2015-NZ-OS-preview.pdf [Accessed 29 Nov. 2017].

Humphreys, M. and Nicol, F. (2007). Self-Assessed Productivity and the Office Environment: Monthly Surveys in Five European Countries. *ASHRAE Transactions; Atlanta*, [online] 113, p. 17. Available at: http://search.proquest.com/docview/192515114/citation/E88CFE3445F94F90PQ/1 [Accessed 17 Apr. 2017].

Jahnke, H. and Halin, N. (2012). Performance, fatigue and stress in open-plan offices: The effects of noise and restoration on hearing impaired and normal hearing individuals. *Noise & Health*, [online[14(60), p. 13. Available at http://www.noiseandhealth.org/article.asp?issn=1463-1741;year=2012;volume=14;issue=60;spage=260;epage=272;aulast=Jahncke [Accessed 15 Aug. 2017].

Kekäläinen, P. Niemela, R. Tuomainen, M. Kemppila, S. Palonen, J. Riuttala, H. Nykyri, E. Seppanen, O. and Reijula, K. (2010). Effect of reduced summer indoor temperature on symptoms, perceived work environment and productivity in office work: An intervention study. *Intelligent Buildings International*, [online] 2(4), p.16. Available at: http://search.proquest.com.helicon.vuw.ac.nz/docview/762485880/abstract/1234A3DD35D3421CPQ/1 [Accessed 16 Apr. 2016].

Kershaw, T. and Lash, D. (2013). Investigating the productivity of office workers to quantify the effectiveness of climate change adaptation measures. *Building and Environment*, [online] 69, p. 9. Available at http://www.sciencedirect.com/science/article/pii/S0360132313002072?via%3Dihub [Accessed 26 May. 2017].

Kosonen, R. and Tan, F. (2004a). Assessment of productivity loss in air-conditioned buildings using PMV index. *Energy and Buildings*, [online] 36(10), p. 7. Available at: http://www.sciencedirect.com/science/article/pii/S0378778804001835 [Accessed 10 Apr. 2016].

Kosonen, R. and Tan, F. (2004b). The effect of perceived indoor air quality on productivity loss. *Energy and Buildings*, [online] 36(10), p. 6. Available at https://search.proquest.com/docview/16187982/E57B84576E074595PQ/1 [Accessed 15 Aug. 2017].

Lakens, D. (2016). Absence of evidence is not evidence of absence: Testing for equivalence. [online] The 20% Statistician. Available at http://daniellakens.blogspot.co.nz/2016/05/absence-of-evidence-is-not-evidence-of.html [Accessed 11 Dec. 2017].

Lamb, S. and Kwok, K.C.S. (2015). A longitudinal investigation of work environment stressors on the performance and wellbeing of office workers. *Applied Ergonomics*, [online] 52, p. 8. Available at http://www.sciencedirect.com/science/article/pii/S0003687015300375 [Accessed 20 May. 2017].

Lan, L. and Lian, Z. (2009). Use of neurobehavioral tests to evaluate the effects of indoor environment quality on productivity. *Building and Environment*, [online] 44(11), p. 10. Available at: http://www.sciencedirect.com/science/article/pii/S0360132309000390 [Accessed 10 Apr. 2016].

Lan, L. Lian, Z. Pan, L. and Ye, Q. (2009). Neurobehavioral approach for evaluation of office workers' productivity: The effects of room temperature. *Building and Environment*, [online] 44(8), p. 11. Available at: http://www.sciencedirect.com/science/article/pii/S0360132308002394 [Accessed 8 Apr. 2016].

Lan, L. Wargocki, P. and Lian, Z. (2011). Quantitative measurement of productivity loss due to thermal discomfort. *Energy and BuildingsI,* [online] 43(5), p. 6. Available at http://www.sciencedirect.com/science/article/pii/S0378778810003117 [Accessed 8 Apr. 2016].

Lan, L. Wargocki, P. and Lian, Z. (2014). Thermal effects on human performance in office environment measured by integrating task speed and accuracy. *Applied Ergonomics*, [online] 45(3), p. 6. Available at: http://www.sciencedirect.com/science/article/pii/S0003687013001348 [Accessed 20 Apr. 2017].

Leaman, A. (1990). Productivity and office quality. *Facilities*, [online] 8(4), p. 3. Available at: http://www.emeraldinsight.com/doi/pdfplus/10.1108/EUM0000000002104 [Accessed 17 Apr. 2017].

Leaman, A. and Bordass, B. (1999). Productivity in buildings: the 'killer' variables. *Building Research & Information*, [online] 27(1), p.16. Available at: http://www.tandfonline.com/doi/abs/10.1080/096132199369615 [Accessed 8 Apr. 2016].

Leyten, J. Kurvers, S. and Raue, A. (2013). Temperature, thermal sensation and workers' performance in air conditioned and free-running environments. *Architectural Science Review*, [online] 56(1), p. 8. Available at: http://dx.doi.org/10.1080/00038628.2012.745391 [Accessed 20 Apr. 2017].

Loftness, V. Aziz, A. Choi, J. Kampschroer, K. Powell, K. Atkinson, M. Heerwagen, J. (2009). The calue of post-occupancy evaluation for building occupants and facility managers. *Intelligent Buildings International*, [online] 1(4), p. 20. Available at http://www.tandfonline.com/doi/abs/10.3763/inbi.2009.SI04 [Accessed 15 Aug. 2017].

Macmillan, S. (2006). Added value of good design. *Building Research & Information*, [online] 32(3), p. 15. Available at http://www.tandfonline.com/doi/abs/10.1080/09613210600590074 [Accessed 16 Aug. 2017].

Mallawaarachchi, H. De Silva, L. and Rameezdeen, R. (2016). Indoor environmental quality and occupants' productivity. *Built Environment Project and Asset Management*, [online] 6(5), p. 16. Available at: http://www.emeraldinsight.com/doi/full/10.1108/BEPAM-09-2015-0046 [Accessed 17 Apr. 2017].

Mallawaarachchi, H. De Silva, L. and Rameezden, R. (2017). Modelling the relationship between green built environment and occupants' productivity. *Facilities*, [online] 35(3/4), p. 18. Available at: http://www.emeraldinsight.com/doi/full/10.1108/F-07-2015-0052 [Accessed 17 Apr. 2017].

Marincic, J. L. (2011). *Vague quantifiers of behavioural frequency: An investigation of the nature and consequences of differences in interpretation.* Ph.D. The University of Nebraska.

Mcgraw Hill Construction. (2014). *The Drive toward Healthier Buildings: the Market Drivers and Impact of Building Design and Construction on Occupant Health Well-being and Productivity.* [online] Bedford: Mcgraw Hill Construction, p. 104. Available at: http://www.worldgbc.org/sites/default/files/Drive_Toward_Healthier_Buildings_SMR_2014.pdf [Accessed 15 Aug. 2017].

Miller, Norm G. Pogue, Dave. Gough, Quiana D. and Davis, Susan M. (2009). Green Buildings and Productivity. *Journal of Sustainable Real Esetate*, [online] 1(1), p. 31. Available at http://catcher.sandiego.edu/items/business/Productivity_paper_with_CBRE_and_USD_Aug_2009-Miller_Pogue.pdf [Accessed 17 Jan. 2018].

Niemela, R. Hannula, M. Rautio, S. Reijula, K. and Railio, J. (2002). The effect of air temperature on labour productivity in call centres – a case study. *Energy and Buildings*, [online] 34(8), p. 6. Available at: http://www.sciencedirect.com/science/article/pii/S0378778802000944 [Accessed 10 Apr. 2016].

Nishihara, N. Wargocki, P. and Tanabe, S. (2014). Cerebral blood flow, fatigue, mental effort, and task performance in offices with two different pollution loads. *Building and Environment*, [online] 71, p. 10. Available at: http://www.sciencedirect.com/science/article/pii/S0360132313002825 [Accessed 20 Apr. 2017].

Occupational Safety & Health Service, (1996a). *Management of Noise in the Workplace.* Wellington: Department of Labour, p. 67.

Occupational Safety & Health Service, (1996b). *Management of Noise at Work.* Wellington: Department of Labour, p. 216.

Park, J.S. and Yoon, C. H. (2011). The effects of outdoor air supply rate on work performance during 8-h work period. *Indoor Air*, [online] 21(4), p. 7. Available at http://onlinelibrary.wiley.com/doi/10.1111/j.1600-0668.2010.00700.x/abstract [Accessed 17 Apr. 2017].

Parke, W.G. and Taylor, R. (2014). *Construction Clients' Protocols – Best Practice Guide 5: Facilities Management*. [online] Auckland: Construction Clients Group, p. 4. Available at https://www.constructing.co.nz/files/file/258/220685%20CCG-Best%20Practice%20Guide%20%235-WEB.pdf [Accessed 16 Aug. 2017].

Parkinson, A. Reid, R. McKerrow, H. and Darren, W. (2017). Evaluating positivist theories of occupant satisfaction: a statistical analysis. *Building Research & Information*, [online] 43(3), p. 15. Available at http://www.tandfonline.com/doi/full/10.1080/09613218.2017.1314148 [Accessed 17 Jun. 2017].

Payscale, (2017). *Office Administrator Salary (New Zealand).* [online] Available at https://www.payscale.com/research/NZ/Job=Office_Administrator/Hourly_Rate [Accessed 29 Nov. 2017].

Perham, N. Banbury, S. and Jones, D. M. (2006). Do realistic reverberation levels reduce auditory distraction? *Applied Cognitive Psychology*, [online] 21(7), p. 9. Available at http://onlinelibrary.wiley.com/doi/10.1002/acp..1300/abstract;jsessionid=CE332DB884465C630DB649C51B4619F8.f04t04 [Accessed 12 Jun. 2017].

Purdey, B. and Leifer, D. (2012). A preliminary study of cognitive failures in open plan offices. *Facilities; Bradford*, [online] 30(11/12), p. 17. Available at: http://search.proquest.com/docview/1033808609/abstract/85EB3E96437D49B2PQ/4 [Accessed 17 Apr. 2017].

Purdey, B. (2013). Occupant stimulus response workplace productivity and the vexed question of measurement. *Facilities*, [online] 31(11/12), p. 16. Available at http://www.emeraldinsight.com/doi/full/10.1108/F-03-2012-0021 [Accessed 26 May. 2017].

Reinten, J. Braat-Eggen, P. E. Hornikx, M. Kort, H. S. M. and Kohlrausch, A. (2017). The indoor sound environment and human task performance: A literature review on the role of room acoustics. *Building and Environment*, [online] 123, p. 18. Available at http://www.sciencedirect.com/science/article/pii/S0360132317302913 [Accessed 12 Jun. 2017].

Riratanaphong, C. (2014). Performance measurement of workplace change: in two different cultural contexts. *A+BE: Architecture and the Built Environment*, [online] 4(2), p. 378. Available at http://journals.library.tudelft.nl/index.php/faculty-architecture/article/view/Riratanaphong/pdf_Riratanaphong [Accessed 26 May. 2017].

Roelofsen, P. (2002). The impact of office environments on employee performance: The design of the workplace as a strategy for productivity enhancement. *Journal of Facilities Management*, [online] 1(3), p. 18. Available at https://search.proquest.com/docview/218906517/abstract/1A72F859EDFE4ACBPQ/4 [Accessed 15 Aug. 2017].

Roy, K. (2008). Acoustic Design for Green Buildings. *Environmental Design + Construction*, [online] 11(3), p. 2. Available at https://search.proquest.com/docview/235037029?accountid=14782 [Accessed 17 Jun. 2017].

Saari, A. and Aalto, L. (2006). Indoor environment quality contracts in building projects. *Building Research & Information*, [online] 34(1), p. 9. Available at: http://www.tandfonline.com/doi/full/10.1080/09613210500279596 [Accessed 20 Apr. 2017].

Saxon, R. (2005). *Be Valuable – A guide to creating value in the built environment.* [online] London: Constructing Excellence, p. 56. Available at http://constructingexcellence.org.uk/wp-content/uploads/2014/10/BeValuable.pdf [Accessed 16 Aug. 2017].

Seppanen, O. and Fisk, W. (2006). Some Quantitative Relations between Indoor Environmental Quality and Work Performance or Health. *HVAC&R Research; Atlanta*, [online] 12(4), p. 17. Available at: http://search.proquest.com/docview/213361895/citation/B70BCA92D7374C93PQ/1 [Accessed 10 Apr. 2016].

Silva, L. Fialho, F. Coutinho, A. Botelho, L. and Augusto, X. (2000). Analysis of the Correlation among thermal Dissatisfaction and Productivity in the Indoor Environment with VDT. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, [online] 44(8), p.4. Available at: http://pro.sagepub.com.helicon.vuw.ac.nz/content/44/8/12 [Accessed 8 Apr. 2016].

Singh, Amanjeet. Syal, Matt. Grady, Sue C. and Korkmaz, Sinem. (2010). Effects of Green Buildings on Employee Health and Productivity. *Am J Public Health*, [online] 100(9), p.4. Available at https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2920980/ [Accessed 17 Jan. 2018].

Sullivan, J. Baird, G. and Donn, M. (2013). *Measuring Productivity in the Office Workplace*. Wellington: Centre for Building Performance Research, p. 75.

Surprenant, A. M. (1999). The Effect of Noise on Memory for Spoken Syllables. *International Journal of Psychology*, [online] 34(5-6), p. 6. Available at http://onlinelibrary.wiley.com/doi/10.1080/002075999399648/abstract [Accessed 12 Jun. 2017].

Tanabe, S. Nishihara, N. and Haneda, M. (2007). Performance evaluation measures for workplace productivity. *Proceedings – 6th International Conference on Indoor Air Quality, Ventilation and Energy Conservation in Buildings: Sustainable Built Environment*, [online] 2, p. 8. Available at: http://www.inive.org/members_area/medias/pdf/Inive%5CIAQVEC2007%5CTanabe.pdf [Accessed 20 Apr. 2017].

Tanabe, S. Kobayashi, K. Kiyota, O. Nishihara, N. and Haneda, M. (2009). The effect of indoor thermal environment on productivity by a year-long survey of a call centre. *Intelligent Buildings International*, [online] 1(3), p.11. Available at: http://search.proquest.com.helicon.vuw.ac.nz/docview/194683808/abstract/F2624E59FED74EA2PQ/8 [Accessed 17 Apr. 2017].

Tanabe, S. Iwahashi, Y. Tsushima, S. and Nishihara, N. (2013). Thermal comfort and productivity in offices under mandatory electricity savings after the Great East Japan earthquake. *Architectural Science Review*, [online] 56(1), p. 10. Available at: http://dx.doi.org/10.1080/00038628.2012.744296 [Accessed 20 Apr. 2016].

Tanabe, S. Haneda, M. and Nishihara, N. (2015). Workplace productivity and individual thermal satisfaction. *Building and Environment*, [online] 91, p. 9. Available at http://www.sciencedirect.com/science/article/pii/S036013231500089X [Accessed 10 Apr. 2016].

Tham, K. (2004). Effects of temperature and outdoor air supply rate on the performance of call center operators in the tropics. *Indoor Air*, [online] 14(7), p. 7. Available at http://onlinelibrary.wiley.com/doi/10.1111/j.1600-0668.2004.00280.x/abstract [Accessed 24 Apr. 2016].

Tham, K. and Willem, H. (2010). Room air temperature affects occupants' physiology, perceptions and mental alertness. *Building and Environment*, [online] 45(1), p.5. Available at: http://www.sciencedirect.com/science/article/pii/S0360132309000985 [Accessed 24 Apr. 2016].

Tham, K. Wargocki, P. and Tan, Y. (2015). Indoor environmental quality, occupant perception, prevalence of sick building syndrome symptoms, and sick leave in a Green Mark Platinum-rated versus a non-Green Mark-rated building: A case study. *Science and Technology for the Built Environment*, [online] 21(1), p. 10. Available at http://www.tandfonline.com/doi/abs/10.1080/10789669.2014.967164 [Accessed 15 Aug. 2017].

Tham, K. (2016). Indoor air quality and its effects on humans – A review of challenges and developments in the last 30 years. *Energy and Buildings*, [online] 130, p. 14. Available at http://www.sciencedirect.com/science/article/pii/S0378778816307812 [Accessed 15 Aug. 2017].

Thomas, L. (2010). Evaluating design strategies, performance and occupant satisfaction: a low carbon office refurbishment. *Building Research & Information*, [online] 38(6), p. 15. Available at http://www.tandfonline.com/doi/abs/10.1080/09613218.2010.501654 [Accessed 26 May. 2017].

Trademe, (2017). *Office & administration salary information*. [online] Available at https://www.trademe.co.nz/jobs/salary-guide/office-administration.htm [Accessed 29 Nov. 2017].

Toftum, J. Kolarik, J. Belkowska, D. and Olesen, B. (2010). Influence on Occupant Responses of Behavioural Modification of Clothing Insulation in Nonsteady Thermal Environments. *HVAC&R Journal*, [online] 16(1), p. 16. Available at http://www.tandfonline.com/doi/abs/10.1080/10789669.2010.10390892 [Accessed 10 Apr. 2016].

Tse, W. L. and So, A. (2007). The importance of Human Productivity to Air-Conditioning Control in Office Environments. *HVAC&R Research*, [online] 13(1), p. 19. Available at http://www.tandfonline.com/doi/abs/10.1080/10789669.2007.10390941 [Accessed 15 Aug. 2017].

Tsutsumi, H. Tanabe, S. Harigaya, J. Iguchi, Y. and Nakamura, G. (2007). Effect of humidity on human comfort and productivity after step changes from warm and humid environment. *Building and Environment*, [online] 42(12), p. 9. Available at http://www.sciencedirect.com/science/article/pii/S0360132306003751 [Accessed 20 Apr. 2017].

U.S. Environmental protection Agency. (1989). *Report to Congress on Indoor Air Quality; Volume II: Assessment and Control of Indoor Air Pollution.* Washington DC: US EPA, p. 250.

Vischer, J. (2008). Towards an Environmental Psychology of Workspace: How People are affected by Environments for Work. *Architectural Science Review*, [online] 51(2), p. 12. Available at: http://www.tandfonline.com/doi/abs/10.3763/asre.2008.5114 [Accessed 17 Apr. 2017].

Valančius, R. and Jurelionis, A. (2013). Influence of indoor air temperature variation on office work performance. *Journal of Environmental Engineering and Landscape Management*, [online] 21(1), p. 7. Available at http://www.tandfonline.com/doi/abs/10.3846/16486897.2012.721371 [Accessed 26 May. 2017].

Wargocki, P. Wyon, D.P. Baik, Y.K. Clausen, G. and Fanger, P.O. (1999). Perceived Air Quality, Sick Building Syndrome (SBS) Symptoms and Productivity in an Office with Two Different Pollution Loads. *Indoor Air,* [online] 9, p. 15. Available at http://onlinelibrary.wiley.com/doi/10.1111/j.1600-0668.1999.t01-1-00003.x/abstract [Accessed 15 Aug. 2017].

Wargocki, P. and Djukanovic, R. (2003). Estimate of an econominc benefit from investment in improved indoor air quality in an office building. *Proceedings of Healthy Buildings 2003*, [online] 3,

p. 6. Available at
https://www.researchgate.net/publication/260386817_Estimate_of_an_economic_benefit_from_investment_in_improved_indoor_air_quality_in_an_office_building [Accessed 13 Sep. 2017].

Zhao, J. Zhu, N. and Lu, S. (2009). Productivity model in hot and humid environment based on heat tolerance time analysis. *Building and Environment*, [online] 44(11), p. 6. Available at: http://www.sciencedirect.com/science/article/pii/S0360132309000146 [Accessed 20 Apr. 2017].

Appendix A: Systematic Literature Review - Search Process

Table A 1 shows the process for filtering relevant information from each database. The databases are ProQuest (PQ), ScienceDirect(SciDi), Sage, Taylor and Francis (T&F), Google Scholar(GS), Avery Index to Architectural Periodicals(AIAP) and JSTOR.

Table A 1 Process for filtering and refining relevant information for research from databases.

Criteria	Fields	PQ	SciDi	Sage	T&F	GS	AIAP	JSTOR	Total
Initial Search ("Work performance" OR productivity) AND building AND (perceived OR "self estimated" OR "self assessed") AND NOT agriculture AND NOT soil AND NOT water AND NOT construction AND NOT machine AND NOT manufacture AND NOT production AND NOT specie AND NOT product AND NOT safety AND NOT land AND NOT residential AND NOT disease	All	2,109	795	1,490	261,563	84,200	1	6	350,164
Subject and Publication filtering	All	153	64	35	529	276	1	0	1,058
Title Screening	Title	15	21	11	23	14	1	0	85
Abstract Screening	Abstract	5	15	5	14	4	1	0	44
Protocol Test "work performance" OR "perceived productivity" OR "self assessed productivity" OR "self estimated productivity" OR "task performance" OR "occupant performance" OR "worker performance" OR "human performance"	All	+20							64
Refinement	Full Text	8	12	0	8	4	0	0	32

Appendix B: Justifying what is a significant difference, comparing average NZ workers Salary with cost of ESD buildings – Original Numbers

The following tables present the original numbers used to analyse how the 5% of increase in work performance is a "significant" increase, as described in Section 2.3.3.1. Numbers in Section 2.3.3.1 have been rounded to three significant figures for simplicity, and this section presents the original numbers used for the analysis.

Table B 1 shows the original figures of extra costs for upgrading a typical office building in New Zealand to a low and high ESD building with estimated potential savings (Fullbrook and Jackson, 2005), and Table B 2 adjusts these figures to per person by multiplying by 17.2, the area per person in m^2 (Colliers International, 2016).

Table B 1 Original figures of extra costs of upgrading to low or high green office buildings and estimated savings per m2 (Fullbrook and Jackson, 2005).

Building costs and savings $/m^2	Benchmark building capital cost	ESD building capital cost	$ difference	$ difference in %	Annual Savings (Energy) $/m^2	Annual Savings (Water) $/m^2	Total Annual Savings (Energy and Water) $/m^2	Simple Payback Years
Office: low-medium ESD	2,000	2,130	130	6.5%	11	0.3	11.3	11.50
Office: medium-high ESD	2,000	2,230	230	11.5%	17	0.6	17.6	13.07

Table B 2 Extra costs for low and high green office buildings adjusted to cost per work station of 17.2m2.

Building costs and savings $/Workstation ($ /17.2m^2)	Benchmark building capital cost	ESD building capital cost	$ difference	$ difference in %	Annual Savings (Energy)	Annual Savings (Water)	Total Annual Savings (Energy and Water)	Simple Payback Years
Office: low-medium ESD	34,400	36,636	2,236	6.5%	189.2	5.16	194.36	11.50
Office: medium-high ESD	34,400	38,356	3,956	11.5%	292.4	10.32	302.72	13.07

Average New Zealand salaries of office workers were obtained from Hudson (2015), Payscale (2017) and Trademe (2017). The accumulation of the percentages of the average salaries, with some percentages based on BUS and CBE scales, are used to calculate potential savings achievable through increase in work performance in Table B 4 and Table B 6, and potential savings in comparison to annual salary in

Table B 5 and

Table B 7.

Table B 3 Percentage of average New Zealand salary with accumulated totals for each percentage.

		Salary	1%	2%	5%	10%	15%	20%	40%
Years	**1**	**45,500**	455	910	2,275	4,550	6,825	9,100	18,200
	5	**227,500**	2,275	4,550	11,375	22,750	34,125	45,500	91,000
	10	**455,000**	4,550	9,100	22,750	45,500	68,250	91,000	182,000
	20	**910,000**	9,100	18,200	45,500	91,000	136,500	182,000	364,000
	25	**1,137,500**	11,375	22,750	56,875	113,750	170,625	227,500	455,000
	50	**2,275,000**	22,750	45,500	113,750	227,500	341,250	455,000	910,000

Table B 4 Total savings in NZD achieved per work station (or per employee) for each of the percentage for savings through increase in work performance only, considering extra costs per workstation for ESD features.

	Years	1%	2%	5%	10%	15%	20%	40%
Low ESD	**1**	−1,781	−1,326	39	2,314	4,589	6,864	15,964
	5	39	2,314	9,139	20,514	31,889	43,264	88,764
	10	2,314	6,864	20,514	43,264	66,014	88,764	179,764
	20	6,864	15,964	43,264	88,764	134,264	179,764	361,764
	25	9,139	20,514	54,639	111,514	168,389	225,264	452,764
	50	20,514	43,264	111,514	225,264	339,014	452,764	907,764
High ESD	**1**	−3,501	−3,046	−1,681	594	2,869	5,144	14,244
	5	−1,681	594	7,419	18,794	30,169	41,544	87,044
	10	594	5,144	18,794	41,544	64,294	87,044	178,044
	20	5,144	14,244	41,544	87,044	132,544	178,044	360,044
	25	7,419	18,794	52,919	109,794	166,669	223,544	451,044
	50	18,794	41,544	109,794	223,544	337,294	451,044	906,044

Table B 5 Percentage of savings in Table 23 in comparison to the average New Zealand salary, for savings through increase in work performance only.

	Years	1%	2%	5%	10%	15%	20%	40%
Low ESD	1	−3.9%	−2.9%	0.1%	5.1%	10.1%	15.1%	35.1%
	5	0.1%	5.1%	20.1%	45.1%	70.1%	95.1%	195.1%
	10	5.1%	15.1%	45.1%	95.1%	145.1%	195.1%	395.1%
	20	15.1%	35.1%	95.1%	195.1%	295.1%	395.1%	795.1%
	25	20.1%	45.1%	120.1%	245.1%	370.1%	495.1%	995.1%
	50	45.1%	95.1%	245.1%	495.1%	745.1%	995.1%	1995.1%
High ESD	1	−7.7%	−6.7%	−3.7%	1.3%	6.3%	11.3%	31.3%
	5	−3.7%	1.3%	16.3%	41.3%	66.3%	91.3%	191.3%
	10	1.3%	11.3%	41.3%	91.3%	141.3%	191.3%	391.3%
	20	11.3%	31.3%	91.3%	191.3%	291.3%	391.3%	791.3%
	25	16.3%	41.3%	116.3%	241.3%	366.3%	491.3%	991.3%
	50	41.3%	91.3%	241.3%	491.3%	741.3%	991.3%	1991.3%

Table B 6 Total savings in NZD achieved per work station (or per employee) for each of the percentage for savings through increase in work performance with energy and water savings, considering extra costs per workstation for ESD features.

	Years	1%	2%	5%	10%	15%	20%	40%
Low ESD	1	−1587	−1132	233	2508	4783	7058	16158
	5	233	2508	9333	20708	32083	43458	88958
	10	2508	7058	20708	43458	66208	88958	179958
	20	7058	16158	43458	88958	134458	179958	361958
	25	9333	20708	54833	111708	168583	225458	452958
	50	20708	43458	111708	225458	339208	452958	907958
High ESD	1	−3198	−2743	−1378	897	3172	5447	14547
	5	−1378	897	7722	19097	30472	41847	87347
	10	897	5447	19097	41847	64597	87347	178347
	20	5447	14547	41847	87347	132847	178347	360347
	25	7722	19097	53222	110097	166972	223847	451347
	50	19097	41847	110097	223847	337597	451347	906347

Table B 7 Percentage of savings in Table 25 in comparison to the average New Zealand salary, for savings through increase in work performance with energy and water savings.

	Years	1%	2%	5%	10%	15%	20%	40%
Low ESD	1	−3.5%	−2.5%	0.5%	5.5%	10.5%	15.5%	35.5%
	5	0.5%	5.5%	20.5%	45.5%	70.5%	95.5%	195.5%
	10	5.5%	15.5%	45.5%	95.5%	145.5%	195.5%	395.5%
	20	15.5%	35.5%	95.5%	195.5%	295.5%	395.5%	795.5%
	25	20.5%	45.5%	120.5%	245.5%	370.5%	495.5%	995.5%
	50	45.5%	95.5%	245.5%	495.5%	745.5%	995.5%	1995.5%
High ESD	1	−7.0%	−6.0%	−3.0%	2.0%	7.0%	12.0%	32.0%
	5	−3.0%	2.0%	17.0%	42.0%	67.0%	92.0%	192.0%
	10	2.0%	12.0%	42.0%	92.0%	142.0%	192.0%	392.0%
	20	12.0%	32.0%	92.0%	192.0%	292.0%	392.0%	792.0%
	25	17.0%	42.0%	117.0%	242.0%	367.0%	492.0%	992.0%
	50	42.0%	92.0%	242.0%	492.0%	742.0%	992.0%	1992.0%

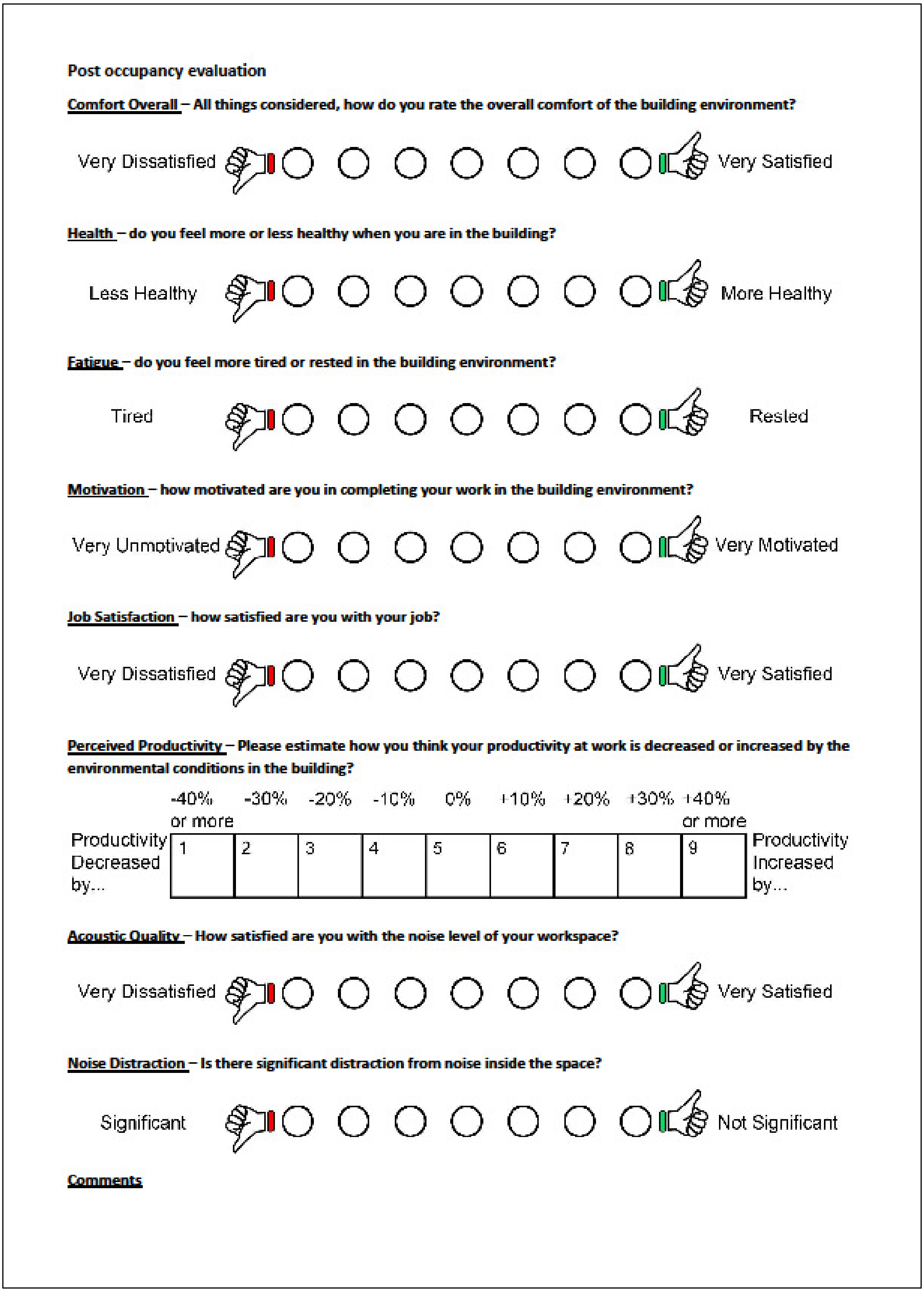
The payback period for the low- and high-ESD buildings shown in Table B 8 suggest that increase in work performance change in the first 1–5% shows a great reduction in payback periods, and large increases in work performance reduces payback period significantly, but the changes in the higher percentages are not as significant as changes in the lower percentage increase in work performance.

Table B 8 Estimated payback period by considering increase in work performance savings and other savings from upgrading to a green office building.

Payback Period (years)		1%	2%	5%	10%	15%	20%	40%
Increase work performance savings	Low ESD	4.9	2.5	1.0	0.5	0.3	0.3	0.1
	High ESD	8.7	4.4	1.7	0.9	0.6	0.4	0.2
Increase in work performance + energy + water savings	Low ESD	3.4	2.0	0.9	0.5	0.3	0.2	0.1
	High ESD	5.2	3.3	1.5	0.8	0.6	0.4	0.2

Appendix C: Survey

BUS Survey


Post occupancy evaluation

Comfort Overall – All things considered, how do you rate the overall comfort of the building environment?

Very Dissatisfied ○ ○ ○ ○ ○ ○ ○ Very Satisfied

Health – do you feel more or less healthy when you are in the building?

Less Healthy ○ ○ ○ ○ ○ ○ ○ More Healthy

Fatigue – do you feel more tired or rested in the building environment?

Tired ○ ○ ○ ○ ○ ○ ○ Rested

Motivation – how motivated are you in completing your work in the building environment?

Very Unmotivated ○ ○ ○ ○ ○ ○ ○ Very Motivated

Job Satisfaction – how satisfied are you with your job?

Very Dissatisfied ○ ○ ○ ○ ○ ○ ○ Very Satisfied

Perceived Productivity – Please estimate how you think your productivity at work is decreased or increased by the environmental conditions in the building?

	-40% or more	-30%	-20%	-10%	0%	+10%	+20%	+30%	+40% or more	
Productivity Decreased by...	1	2	3	4	5	6	7	8	9	Productivity Increased by...

Acoustic Quality – How satisfied are you with the noise level of your workspace?

Very Dissatisfied ○ ○ ○ ○ ○ ○ ○ Very Satisfied

Noise Distraction – Is there significant distraction from noise inside the space?

Significant ○ ○ ○ ○ ○ ○ ○ Not Significant

Comments

Figure C 1 Post-occupancy survey used for the experiment with the BUS scale.

CBE Survey

Post occupancy evaluation

Comfort Overall – All things considered, how do you rate the overall comfort of the building environment?

Very Dissatisfied ○ ○ ○ ○ ○ ○ ○ Very Satisfied

Health – do you feel more or less healthy when you are in the building?

Less Healthy ○ ○ ○ ○ ○ ○ ○ More Healthy

Fatigue – do you feel more tired or rested in the building environment?

Tired ○ ○ ○ ○ ○ ○ ○ Rested

Motivation – how motivated are you in completing your work in the building environment?

Very Unmotivated ○ ○ ○ ○ ○ ○ ○ Very Motivated

Job Satisfaction – how satisfied are you with your job?

Very Dissatisfied ○ ○ ○ ○ ○ ○ ○ Very Satisfied

Perceived Productivity – Please estimate how you think your productivity at work is decreased or increased by the environmental conditions in the building?

Decreased ○ -20% ○ -10% ○ -5% ○ 0% ○ +5% ○ +10% ○ +20% Increased

Acoustic Quality – How satisfied are you with the noise level of your workspace?

Very Dissatisfied ○ ○ ○ ○ ○ ○ ○ Very Satisfied

Noise Distraction – Is there significant distraction from noise inside the space?

Significant ○ ○ ○ ○ ○ ○ ○ Not Significant

Comments

Figure C 2 Post-occupancy survey used for the experiment with the CBE scale.

Appendix D: BUS and CBE Case Studies with Standard Deviations

All numbers from the studies are converted to percentage values, and data from Brown et al. (2010) has been obtained through the graph presenting the percentage of people assessing their work performance at each point on the scale. Although all the studies listed use numerical scales, some studies report the changes on ordinal scales (e.g., CBE is a 7-point scale from −20% to +20%, where *1* is −20% and *7* is +20%). For calculating the mean standard deviation for each group, the sum of standard deviation multiplied by number of subjects in each of the studies, was divided by the total number of subjects in the group.

Table D 1 Studies presenting the standard deviation and number of responses in the surveys with average standard deviations for BUS and CBE surveys.

Type of Survey	Author	Number of Responses	Standard Deviation
BUS	Brown et al. (2010)	HQ1 = 145 HQ2 =164	HQ1 = 13.26 HQ2 = 16.23
	Baird, Leaman and Thompson (2012)	2035	10.02
	Parkinson et al. (2017)	2700	16
	Mean SD for BUS = 13.52		
CBE	Gou, Prasad and Lau (2013)	Green building = 774 non-green building = 477	Green building = 7.56 non-green building = 8.31
	Gou, Prasad and Lau (2014)	Highly rated green building = 593 low-rated green building = 181 non-green building = 477	Highly rated green building = 7.69 low-rated green building = 8.42 non-green building = 7.76
	Mean SD for CBE = 7.83		

Appendix E: Data Summary for Other Self-reported Indicators of Work Performance

Table E 1 shows the changes in other self-reported indicators of work performance, motivation, job satisfaction, fatigue and distraction obtained from the surveys, over the three pink noise conditions. The overall mean, margin of error, sample size and SD is not included, because analysis from other self-reported indicators will correlate with the corresponding self-reported work performance surveys measures.

Table E 1 Summary of other self-reported indicators of work performance gathered from survey.

			Noise conditions (dB(A))		
	Group	**Results**	**40**	**55**	**65**
Motivation	**BUS**	Mean	5.07	4.33	3.53
		Margin of Error (95%)	1.01	0.95	1.00
		Sample Size	15	15	15
		SD	1.83	1.72	1.81
	CBE	Mean	5.00	3.87	3.33
		Margin of Error (95%)	0.78	0.98	0.80
		Sample Size	15	15	15
		SD	1.41	1.76	1.45
Job Satisfaction	**BUS**	Mean	5.20	4.53	3.67
		Margin of Error (95%)	0.76	0.96	0.85
		Sample Size	15	15	15
		SD	1.37	1.73	1.54
	CBE	Mean	5.13	4.67	3.73
		Margin of Error (95%)	0.69	0.80	0.64
		Sample Size	15	15	15
		SD	1.25	1.44	1.16

(continued)

Table E 2 Summary of other self-reported indicators of work performance gathered from survey (continued).

			Noise conditions (dB(A))		
	Group	**Results**	**40**	**55**	**65**
Fatigue	**BUS**	Mean	4.80	3.80	3.47
		Margin of Error (95%)	0.84	0.73	0.96
		Sample Size	15	15	15
		SD	1.52	1.32	1.73
	CBE	Mean	4.40	3.87	2.80
		Margin of Error (95%)	0.55	0.62	0.60
		Sample Size	15	15	15
		SD	0.99	1.12	1.08
Distraction	**BUS**	Mean	6.13	4.33	3.67
		Margin of Error (95%)	0.75	1.10	0.95
		Sample Size	15	15	15
		SD	1.36	1.99	1.72
	CBE	Mean	5.27	3.80	3.20
		Margin of Error (95%)	0.90	0.99	0.87
		Sample Size	15	15	15
		SD	1.62	1.78	1.57

Appendix F: Cognitive Performance Test Results

Table F 1shows the cognitive performance test results, where performance is calculated by accuracy/speed (correct percentage /time in seconds) except for memory span where performance is number of correct strings, for each noise condition. Changes in cognitive performance over the noise conditions are not consistent, as some show increasing performance with the noisiest 65 dB(A) conditions, some show a decrease and some show best performance in the 55 dB(A) conditions. As well as changes in cognitive performance being inconsistent across the noise conditions, the changes in work performance differ between BUS and CBE groups. Inconsistent changes do not allow comparisons between the data to analyse differences in self-reported work performance measures from BUS and CBE and their correlation with cognitive performance.

Table F 1 Summary of cognitive performance test results for each noise condition.

			Noise Conditions (dB(A))		
	Group	**Results**	**40**	**55**	**65**
Letter Search	**BUS**	Mean	0.41	0.44	0.43
		Margin of Error (95%)	0.07	0.06	0.07
		Sample Size	15	15	15
		SD	0.13	0.11	0.13
	CBE	Mean	0.42	0.40	0.39
		Margin of Error (95%)	0.07	0.08	0.08
		Sample Size	15	15	15
		SD	0.12	0.15	0.14
	All	Mean	0.42	0.42	0.41
		Margin of Error (95%)	0.05	0.05	0.05
		Sample Size	30	30	30
		SD	0.13	0.13	0.13
Overlapping	**BUS**	Mean	0.99	0.97	0.93
		Margin of Error (95%)	0.09	0.15	0.11
		Sample Size	15	15	15
		SD	0.17	0.28	0.20
	CBE	Mean	1.05	1.04	1.25
		Margin of Error (95%)	0.14	0.12	0.28
		Sample Size	15	15	15
		SD	0.24	0.22	0.50

	All	Mean	1.02	1.00	1.09
		Margin of Error (95%)	0.08	0.09	0.15
		Sample Size	30	30	30
		SD	0.21	0.25	0.41
Memory Span	BUS	Mean	5.33	5.80	5.27
		Margin of Error (95%)	0.88	1.21	0.87
		Sample Size	15	15	15
		SD	1.59	2.18	1.58
	CBE	Mean	4.47	4.20	5.53
		Margin of Error (95%)	1.61	0.87	1.39
		Sample Size	15	15	15
		SD	2.90	1.57	2.50
	All	Mean	4.90	5.00	5.40
		Margin of Error (95%)	0.87	0.76	0.77
		Sample Size	30	30	30
		SD	2.34	2.03	2.06
Picture Recognition	BUS	Mean	3.21	2.98	3.19
		Margin of Error (95%)	0.37	0.35	0.38
		Sample Size	15	14	15
		SD	0.67	0.61	0.69
	CBE	Mean	3.05	3.12	3.38
		Margin of Error (95%)	0.29	0.32	0.31
		Sample Size	15	15	15
		SD	0.52	0.59	0.57
	All	Mean	3.13	3.05	3.29
		Margin of Error (95%)	0.22	0.23	0.24
		Sample Size	30	29	30
		SD	0.59	0.59	0.60
Symbol Digit Modalities Test	BUS	Mean	0.88	0.80	0.82
		Margin of Error (95%)	0.12	0.16	0.13
		Sample Size	15	15	15
		SD	0.22	0.29	0.24
	CBE	Mean	0.90	0.83	0.91

		Margin of Error (95%)	0.19	0.15	0.20
		Sample Size	14	15	15
		SD	0.33	0.27	0.37
	All	Mean	0.89	0.82	0.87
		Margin of Error (95%)	0.10	0.10	0.12
		Sample Size	29	30	30
		SD	0.27	0.27	0.31
Number calculation	**BUS**	Mean	0.33	0.31	0.30
		Margin of Error (95%)	0.07	0.07	0.08
		Sample Size	15	15	15
		SD	0.13	0.13	0.14
	CBE	Mean	0.25	0.27	0.25
		Margin of Error (95%)	0.07	0.06	0.06
		Sample Size	15	15	15
		SD	0.13	0.11	0.11
	All	Mean	0.29	0.29	0.28
		Margin of Error (95%)	0.05	0.05	0.05
		Sample Size	30	30	30
		SD	0.13	0.12	0.13
Conditional Reasoning	**BUS**	Mean	0.24	0.30	0.26
		Margin of Error (95%)	0.10	0.15	0.10
		Sample Size	15	15	15
		SD	0.18	0.27	0.18
	CBE	Mean	0.17	0.21	0.23
		Margin of Error (95%)	0.09	0.10	0.13
		Sample Size	14	15	15
		SD	0.16	0.18	0.24
	All	Mean	0.21	0.25	0.25
		Margin of Error (95%)	0.06	0.09	0.08
		Sample Size	29	30	30
		SD	0.17	0.23	0.21
Spatial	**BUS**	Mean	1.27	1.55	1.35
		Margin of Error (95%)	0.39	0.44	0.49
		Sample Size	15	15	15

		SD	0.71	0.79	0.88
	CBE	Mean	1.47	1.49	2.09
		Margin of Error (95%)	0.38	0.42	0.98
		Sample Size	15	15	15
		SD	0.69	0.76	1.77
	All	Mean	1.37	1.52	1.72
		Margin of Error (95%)	0.26	0.28	0.53
		Sample Size	30	30	30
		SD	0.69	0.76	1.42
Visual Choice Reaction Time	BUS	Mean	2.36	2.43	2.62
		Margin of Error (95%)	0.34	0.41	0.23
		Sample Size	15	14	15
		SD	0.62	0.71	0.41
	CBE	Mean	2.73	2.78	2.77
		Margin of Error (95%)	0.19	0.39	0.20
		Sample Size	13	15	15
		SD	0.32	0.71	0.36
	All	Mean	2.53	2.61	2.70
		Margin of Error (95%)	0.21	0.27	0.15
		Sample Size	28	29	30
		SD	0.53	0.72	0.39