A Framework for Presenting Energy Efficiency Feedback to Operators of Buildings

A. Craig Roussac

A thesis submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy

Sydney School of Architecture, Design and Planning
University of Sydney
2017
DECLARATION

This is to certify that to the best of my knowledge, the content of this thesis is my own work. This thesis has not been submitted for any degree or other purposes.

I certify that the intellectual content of this thesis is the product of my own work and that all the assistance received in preparing this thesis and sources have been acknowledged.

A. Craig Roussac
ACKNOWLEDGEMENTS

My supervisor at the University of Sydney, Professor Richard de Dear, for supervision, unwavering patience, reviewing every page and offering terrific guidance and support.

Professor Richard Hyde for encouragement and supervision during my first year.

Dr Gabrielle Kuiper for encouraging me to refocus my role in commercial real estate towards “action research”; a transition that led me to this project.

Investa Property Group, in particular Campbell Hanan, Executive Director, and Scott MacDonald, CEO and Chairman, for approving the research. Investa management for not interfering and all my colleagues who were open-minded and constructive participants.

Jesse Steinfeld and Baden Hughes, my colleagues and co-founders of Buildings Alive, for significant contributions to the development of the regression models and software application respectively.

Colleagues at Buildings Alive, most notably Dr Serguei Zavtrak for developing an alternative statistical model to support results verification, Dr Mitra Bahadorian for assistance reviewing and helping to optimise my data analysis code, and Cameron Roach for technical assistance with R programming challenges.

Staff at the Australian-American Fulbright Commission, in particular Dr Tangerine Holt, Dr Ruth Lee Martin and Dr Pablo Jiménez for assistance and support during the lead up to my time in the U.S.A.

Staff scientists at the Lawrence Berkeley National Laboratory who critiqued my initial results and analysis, in particular Dr Rick Diamond, my host at LBNL, for facilitating presentations and introductions.
Staff and students at UC Berkeley’s Center for the Built Environment for contributing ideas and suggestions and in particular CBE director Professor Ed Arens for hosting me during my time at UC Berkeley.

My parents, Russell and Judi, for giving me the confidence and drive to attempt something like this.

Most of all, my amazing wife Georgie for her boundless patience, love and support and our two boys, Charlie and Tom, for being such positive and inspiring kids.
ABSTRACT

Buildings account for more than half of the world’s total electricity consumption and almost one fifth of all energy-related greenhouse gas emissions. Between 1970 and 2010, indirect greenhouse gas emissions from commercial buildings more than quintupled globally and they continue to grow steadily.

The potential to reduce energy use—and consequently greenhouse gas emissions—from the operation of commercial buildings through non-technological (i.e. human behaviour-based) approaches is rarely assessed and poorly understood. This study, which has involved the facilities managers (i.e. operators) of thirty large Australian commercial office buildings, twenty-two of which were unchanged in terms of their technology and equipment for a period exceeding 220 business days, is longer in duration and larger in number of participants than any non-residential study reported in the literature previously.

Every weekday morning the facilities managers received a short email message containing a “judgement” about their building’s energy use for the previous day. The messages were generated by software that processed the preceding day’s weather and electricity meter data using a statistical model of each building’s operating patterns leading up to the commencement of the study. Messages conveyed either “good”, “bad” or “average” news in their subject line and there were also some short descriptive sentences and simple intra-day profile graphics provided within the body of the message. Together these pointed to performance anomalies and learning opportunities for the recipients to consider. Facilities managers had the option to discuss their experiences and feedback at a regular fortnightly discussion forum comprising their peers from nearby buildings and other observers, including the researcher.
The methodology adopted feedback design elements identified as important from residential studies but not previously applied in a non-residential context. Insights from the extensive literature on human motivation and behaviour, feedback methodologies and social learning theory were also applied.

No other studies reported in the literature have used a methodology of the type developed for this study and the social learning context which was adopted to support the participants, while based on established research and experience from other contexts, is a unique application. Furthermore, the methodology overcame difficulties in recruiting subjects and controlling for factors such as building size, use, design, and age, which have been found to limit studies of commercial building operations published to-date.

A significant divergence between actual and ‘predicted’ daily electricity consumption began to appear after about day 50 of the treatment and a mean improvement of greater than 6 percent was recorded over the entire intervention period. A saving of approximately 10 percent compared to the pre-intervention baseline was observed at the conclusion of the trial.

This study has implications for the design of programs aimed at reducing energy use from non-residential buildings, particularly in relation to the roles of motivation, goal setting and peer interaction, the development of skill and competency, and the permanency of savings. It also indicates the potential for achieving energy savings from operational buildings at minimal cost and without resorting to capital-intensive upgrade projects. Noting evidence that individual behaviours are heavily influenced by social and institutional contexts and reactions, and hence not easily generalised or up-scaled, specific recommendations for presenting energy efficiency feedback to building operators are offered.
CONTENTS

ACKNOWLEDGEMENTS ........................................................................................................... III
ABSTRACT ................................................................................................................................. V
LIST OF FIGURES ...................................................................................................................... XI
LIST OF TABLES ....................................................................................................................... XIV
LIST OF ABBREVIATIONS ....................................................................................................... XV

1. **Introduction** ..................................................................................................................... 1
   1.1. Problem and purpose .................................................................................................... 1
   1.2. Thesis structure .......................................................................................................... 4
   1.3. Delimitations .............................................................................................................. 6

PART A: BACKGROUND ........................................................................................................... 7

2. **Buildings Matter** ............................................................................................................. 7
   2.1. Impact of buildings ..................................................................................................... 7
   2.2. Energy use in buildings is growing ........................................................................... 9
   2.3. Large institutions keen to be green ......................................................................... 12
   2.4. Governments keen to be green ................................................................................. 13
   2.5. Design of buildings is a constraint ......................................................................... 14
   2.6. Factors explaining a building’s daily energy use ...................................................... 16
   2.7. Opportunities through technology and non-technological interventions .......... 18
   2.8. Comfort and IEQ considerations ............................................................................. 21
   2.9. Buildings getting greener, but more complex ......................................................... 23

3. **Research is required** ....................................................................................................... 26
   3.1. Previous studies focussed on residential ............................................................... 26
   3.2. Information alone is not enough ................................................................................. 27
   3.3. Some studies lack methodological strength .......................................................... 29
   3.4. Uncertainty about the drivers of energy behaviours ............................................. 31
   3.5. Opportunity .............................................................................................................. 32
   3.6. More data needed ..................................................................................................... 33

4. **Human motivation and behaviour** ................................................................................. 34
   4.1. Motivation drives behaviour .................................................................................... 34
   4.2. Behaviour theories – conditioning ......................................................................... 37
   4.3. Behaviour modification ......................................................................................... 40
4.4. Behaviourism provides an incomplete account ........................................40
4.5. Influence of values and ethics .................................................................44
4.6. Influence of habits ..................................................................................46
4.7. Triandis’ Theory of Interpersonal Behaviour ..........................................48
4.8. Stages of behavioural change ................................................................48
4.9. Limitations of behaviour-based approaches ...........................................53
5. Feedback .....................................................................................................56
5.1. What is feedback? .................................................................................56
5.2. Elements of highly effective feedback and assessment ........................59
5.3. Important contextual factors ..................................................................64
5.4. Insights from Residential feedback – HEMS .......................................66
5.5. Problems with feedback loops in buildings ............................................70
5.6. The boomerang effect ...........................................................................73
5.7. Feedback without goals ineffective? .........................................................75
5.8. Fatigue limitations ..................................................................................79
6. Social learning context ...............................................................................82
6.1. Complex behaviours require models ......................................................82
6.2. Social Learning Theory .........................................................................83
6.3. Social proof and Theory of Reasoned Action .......................................85
6.4. Social marketing and action ...................................................................87

PART B: RESEARCH DESIGN AND METHODOLOGY ......................................91
7. Questions and organising principles .........................................................91
7.1. Approach ...............................................................................................91
7.2. Model ....................................................................................................93
7.3. Research question and sub-problems ....................................................95
7.4. Causal hypothesis ..................................................................................98
8. Portfolio selection and preliminary review .............................................99
8.1. Portfolio selection ..................................................................................99
8.2. Investa Property Group portfolio ........................................................100
8.3. Pre-trial review of a subsample [need def of office somewhere] ..........103
8.3.1. Some cursory observations ...............................................................106
8.3.2. Building X, Sydney ..........................................................................107
8.3.3. Building AB, Sydney ........................................................................110
8.3.4. Relationships between energy intensity, occupant satisfaction, technology and behaviour ................................................................. 112

8.4. Initial prototype and its limitations .................................................. 113

9. Intervention design ........................................................................... 117

9.1. Automatic feedback ....................................................................... 117

9.1.1. Data sources ............................................................................. 118

9.1.2. Statistical model ........................................................................ 119

9.1.3. Processing and storage ............................................................... 125

9.1.4. Email messages ......................................................................... 127

9.2. Social learning environment ............................................................ 132

10. Approach to evaluation .................................................................... 135

10.1. Hypothesis is refutable .................................................................. 135

10.2. Additional model ......................................................................... 136

10.3. Statistical tests and reproducibility ............................................... 143

PART C: RESULTS .................................................................................. 145

11. Summary of results .......................................................................... 145

11.1. Summary of trial .......................................................................... 145

11.2. Regression discontinuity design .................................................... 152

11.3. Individual building results ............................................................ 154

PART D: DISCUSSION & CONCLUSIONS .............................................. 161

12. Comments on the interventions and significance of the study .............. 161

12.1. Significance of the study ............................................................... 161

12.2. Significant results ....................................................................... 162

12.3. Impact of removing eight buildings from the sample ..................... 164

12.4. Commonalities and differences between buildings and operators ...... 166

13. Limitations and strengths of the methodology .................................... 174

13.1. Multidisciplinary focus and design ................................................. 174

13.2. Stable conditions .......................................................................... 177

13.3. Quantitative study ....................................................................... 178

13.4. Trust and credibility was high ....................................................... 179

13.5. Limited interaction with tenants .................................................... 180

13.6. Factors impacting replicability of results ...................................... 181

13.6.1. Large modern Australian office buildings in CBD locations ...... 182
13.6.2. Market-leading energy performance prior to trial commencement...184
13.6.3. Professionally managed.................................................................185
13.6.4. Highly motivated owner-operator ...............................................186
13.6.5. Focus on ‘base building’ only (tenancies ignored).........................187
13.6.6. Researcher’s association with the sample ....................................188

14. Key findings ..........................................................................................191

14.1. Motivation and goals ........................................................................191
14.1.1. Awareness is necessary, but not sufficient ....................................192
14.1.2. No need for “rewards” or “incentives” ..........................................192
14.1.3. Explicit goals are unnecessary .......................................................193
14.2. The role of social learning ................................................................195
14.2.1. Cues and trust are important .........................................................196
14.2.2. Pressure from hierarchy not required ............................................198
14.2.3. Peers a good source of formative assessment ..............................199
14.3. Skill levels increased .........................................................................202
14.3.1. Evaluative expertise emerged .......................................................202
14.3.2. Skill transfer could have been faster ............................................205
14.4. Will savings persist? ........................................................................208
14.4.1. No boomerang effect ....................................................................208
14.4.2. Fatigue not an issue .....................................................................209
14.4.3. Energy saving becomes habit .......................................................210

15. Conclusions and recommendations ....................................................212

REFERENCES ............................................................................................219
APPENDIX A: APPROVAL TO CONDUCT AND PUBLISH RESEARCH ..........239
APPENDIX B: RDD SCATTER PLOTS FOR ALL 30 BUILDINGS .................241
LIST OF FIGURES

Figure 2-1. Regional direct and indirect CO2 emissions from the building subsectors (Lucon et al., 2014) ................................................................. 9

Figure 2-2. Weighted Average NABERS Rating & Total Rated Area (IPD, 2014) .............. 11

Figure 4-1. Operant conditioning diagram ........................................................... 39

Figure 4-2. The seven population segments (DEFRA, 2008) ....................................... 45

Figure 4-3. Triandis’ Theory of Interpersonal Behaviour (Triandis, 1977) ......................... 51

Figure 4-4. The flow of behavior change model (Geller, 2002) ...................................... 53

Figure 5-1. A typical feedback loop ........................................................................ 71

Figure 5-2. The effects of feedback intervention (FI)-induced attention on task- motivation processes and their consequences for performance (Kluger & DeNisi, 1996) ......................................................................................... 76

Figure 6-1. Action-reflection framework (Source: Adapted by Serrat, 2008; from Kolb, 1984) ......................................................................................... 90

Figure 7-1. Sequence of and interaction between variables in the research design .......... 92

Figure 7-2. Simplified version of Triandis’ Theory of Interpersonal Behaviour (Triandis, 1977) limited to key variables considered in this study. The complete model is presented in Figure 4-3. ................................................................. 94

Figure 8-1. Investa Property Group portfolio electricity intensity (Investa Property Group, 2010) ......................................................................................... 101

Figure 8-2. Base-building electricity consumption during the pre-trial period 2009-10 (Roussac, de Dear, & Hyde, 2011) ......................................................... 104

Figure 8-3. Energy consumption and air conditioning complaints at Building X, 2004–2008 (Roussac, 2009) ........................................................................... 108
Figure 11-5. Variance between ‘predicted’ and actual ‘work day’ electricity consumption.................................................................................................................................152

Figure 11-6. Regression discontinuity design (RDD) showing improvement trend prior and post intervention (vertical dashed line). .................................................................153

Figure 11-7. Distribution of calibration set gradients.................................................................155

Figure 11-8. Distribution of gradients for the trial period..........................................................155

Figure 11-9. Regression discontinuity design for each building in the analysis (and also BID0026). ...............................................................................................................................158

Figure 11-10. Pattern of improvement at each of the trial buildings in the analysis (and also BID0026). ...............................................................................................................................160

Figure 12-1. RDD showing trend for all 28 buildings in the study with sufficient data (none excluded). Vertical dashed line marks beginning of intervention. .................................................................................................................................165

Figure 12-2. Distribution of sample variances by social learning group..............................170

Figure 12-3. Distribution of sample variances by building size (blue bars exceed median).................................................................................................................................171

Figure 12-4. Distribution of sample variances by pre-trial NABERS Energy rating (lime green bars exceed mean).................................................................................................171

Figure 12-5. Comparison of improvement for 123 Australian office buildings categorised by pre-intervention rating on NABERS Energy (Steinfeld, 2016) .................................................................................................................................172

Figure 14-1. Investa hierarchy in place at the time of the study (active participants shaded). .................................................................................................................................198

Figure 14-2. Difference between ‘predicted’ and actual electricity consumption after implementing a similar feedback mechanism at 53 buildings on a university campus, but without the facilitated social learning.................................201
Figure 14-3. Screenshot of BID0001’s HVAC electricity submeter indicating a reduction in the morning ‘spike’ following a lowering of space temperature setpoint (Roussac & de Dear, 2012). ..................................................203

Figure 14-4. Impact of fine-tuning temperature setpoints at BID0001 on 31 August, 2011 and Friday 26 August, 2011, as illustrated in the ‘Pulse’ message. .................................................................204

Figure 14-5. Difference between ‘predicted’ and actual electricity consumption for all Investa buildings (including some not involved in the study). Post-study results (i.e. after ~250 days) were influenced by coaching and facilitation. All non-Investa buildings had coaching and facilitation from the outset. .................................................................207

LIST OF TABLES

Table 8-1. Sample of eleven of Investa’s commercial office buildings (Roussac, 2009) .................................................................................................................................105

Table 11-1. Summary of buildings included in the trial ........................................147

Table 11-2. Standard deviations of the two models .............................................148

Table 11-3. Trends for individual buildings, prior to and during the trial ..........156

Table 11-4. Individual building standard deviations and improvements (trial predictions minus actual) .................................................................159

Table 12-1. NABERS Energy ratings (0-6 star scale) and performance statistics prior to and during trial .................................................................168
LIST OF ABBREVIATIONS

Advanced Metering Infrastructure (AMI)
Air Handling Unit (AHU)
Assessment Report Five (AR5): IPCC Fifth Assessment Report
Assessment Report Four (AR4): IPCC Fourth Assessment Report
Australian Securities Exchange (ASX)
Australian Sustainable Built Environment Council (ASBEC)
Australian Bureau of Agriculture and Resource Economics (ABARE)
Australian Council of Superannuation Investors (ACSI)
Australian Energy Market Operator (AEMO)
Building Energy Efficiency Certificate (BEEC)
Building Management & Control System (BMCS)
Bureau of Meteorology (BoM)
Central Business District (CBD)
Centre for International Economics (CIE)
Commercial Building Disclosure Program (CBD Program)
Council of Australian Governments (COAG)
Department for Environment, Food and Rural Affairs (DEFRA),
Energy Information System (EIS)
Energy Use Intensity (EUI)
Facilities Manager (FM)
Facilities Resource Management (FRM)
File Transfer Protocol (FTP)
Gross Domestic Product (GDP)
Heating, Ventilation, and Air Conditioning (HVAC)
Home Energy Management System (HEMS)
In-Home Display (IHD)
Indoor Environmental Quality (IEQ)
Information Technology (IT)
Intergovernmental Panel on Climate Change (IPCC)
International Energy Agency (IEA)
Investment Property Databank Ltd (IPD)
Knowledge of Performance (KP)
Knowledge of Results (KR)
Leadership in Energy and Environmental Design (LEED)
Low Energy High Rise (LEHR) project
National Australian Built Environment Rating System (NABERS)
National Electricity Market (NEM)
Net Lettable Area (NLA)
New South Wales (NSW)
Organisation for Economic Co-operation and Development (OECD)
Property Council of Australia (PCA)
Randomised Control Trial (RCT)
Real Estate Investment Trust (REIT)
Regression Discontinuity Design (RDD)
Secure Shell File Transfer Protocol (S/FTP)
Simple Mail Transfer Protocol (SMTP)
Theory of Interpersonal Behaviour (TIB)
Theory of Planned Behaviour (TPB)
Variable Speed Drive (VSD)
World Business Council for Sustainable Development (WBCSD)
1. INTRODUCTION

1.1. Problem and purpose

The operations and maintenance of large, complex, non-residential buildings are usually attended to by specialist technicians overseen by a facilities manager (FM). A person can enter the facilities management profession from a wide variety of backgrounds and there is no particular level of training or experience that qualifies someone to be an FM for a large, complex, non-residential building. It is not uncommon for an FM to have no background in energy management whatsoever. Despite this, technicians generally defer to FMs (or their equivalent role, as the use of titles varies) for decisions about operational matters and FMs, with reference to the priorities of building owners, have a key role in determining how budgets are spent and how building systems are operated.

The FM’s decisions about the way a building is operated have direct consequences for energy use, and the difference between the amount of energy used by an efficiently operated building and one that is inefficient, aside from the influence of the building’s underlying design and technology, can be substantial. The management of energy use in buildings is therefore a significant environmental, social and economic concern because the potential (discussed in chapter 2) for non-residential buildings to contribute to large-scale and low-cost reductions in energy use—and energy-related greenhouse gas emissions—is widely acknowledged. Yet non-technological (i.e. human behaviour-based) approaches to saving energy in building operations are rarely assessed in the literature. As a consequence, there is still much to be learned about the application of behaviour-based interventions for unlocking energy savings, both in terms of the scale of the savings that might arise and the techniques required to effectively address them.
Almost all of the literature that does relate to energy consuming behaviours in buildings comes from the residential sector, and particularly from studies involving the occupants of residential dwellings. This literature provides valuable insights to consider in the context of non-residential buildings. However, unlike the operators of non-residential buildings, the subjects of these studies (referred to typically as “occupants”, “residents” or “customers” of utility companies) deal with relatively uncomplicated “on / off” choices and rarely face the sorts of technical challenges involved in the optimisation of complex building systems.

Likewise, the extensive literature on human motivation and behaviour, application of feedback methodologies and the transfer of knowledge through social interactions spans many fields, and much of it is directly relevant to the operation of buildings. Yet few published studies have applied knowledge from the literature of behaviour science to the questions of how to either influence or explain the patterns of energy use in non-residential buildings, and none have sought to apply quantitative methods to assess the reduction potential as this thesis does.

A researcher must overcome an array of challenges in order to explore the influence that human motivation and competence have over the amount of energy used to operate non-residential buildings. An assessment of the reduction potential introduces even more. This may, in part, help to explain why so little research has been undertaken in this field despite its environmental, social and economic importance. First of all, a statistically significant sample of buildings needs to be recruited in order to provide a sound basis for evaluation. Differences in building size, use, mechanical systems design, plant and equipment, architectural design, and age, must therefore be understood and normalised. FMs, too, come from different backgrounds, have different levels of experience and
education, possess a range of beliefs and attitudes, etc. which must either be accommodated or accounted for. Furthermore, the amount of energy used to operate a non-residential building from one day to the next is influenced by a variety of factors aside from the competence and motivation of its operators. Weather conditions, occupancy patterns and user demands can all have a profound influence on energy use. All of these are not just challenges for researchers attempting to unpack the relationships between cause and effect; they are also challenges for the FMs who inevitably find it very difficult to isolate, measure and evaluate the impact of their actions. FMs regularly find that they are unable to measure energy savings from tuning their buildings’ operations because they are offset by factors outside their control. And without unambiguous and timely feedback, how can an FM be expected to experiment, learn and improve?

The study presented in this thesis focusses on a portfolio of thirty commercial office buildings from the climatically dissimilar Australian cities of Melbourne, Sydney, and Brisbane. Commercial buildings, like industrial, educational and health buildings, are a classification within the non-residential building category (OECD, 2001). Commercial office buildings are a typology within the commercial category, as are warehouses and retail shops. In Australia, a building is defined as a commercial office building if its dominant use is for “administrative, clerical, professional or similar information-based activities, including any support facilities for those activities” (Department of Industry, 2014a, p. 1). The thesis builds on the hypothesis that FMs operating Australian commercial office buildings will reduce energy usage if: a) they are provided with automated and unambiguous daily energy performance feedback that accounts for factors outside their control; and b) they have the opportunity to share and discuss this feedback with peers.
1.2. Thesis structure

The thesis is comprised of four parts, aside from this introductory content and the references and appendices at the end. Part A, which has five chapters, provides a broad overview of the theory and subject matter that was taken into account when designing the study. It starts, in chapter 2, by providing an overview of the impact that buildings have on people and the environment, energy usage trends, and some of the factors driving them. Chapter 3 then looks at significant research conducted in the field, including some of the challenges this research has encountered and also what has been learned in other contexts that can be applied in a study involving operators of large commercial buildings. Chapter 4 explores the theory of human motivation and behaviour including the major influences on the field of behaviour science. It introduces Triandis' Theory of Interpersonal Behaviour (TIB) which has been used as a theoretical basis for the intervention design and analysis, and also discusses various influences on behaviour and decision making that are particularly relevant for the operations of commercial buildings as well as some of the limitations of behaviour-based approaches. Chapter 5 then introduces “feedback”, from the relatively recent origins of the term through to the ingredients that are now widely recognised as necessary for it to be effective as a device to influence behaviour change. It also describes many of the limitations and challenges inherent in feedback mechanisms, both in buildings and through application more generally. The background section then concludes with an overview of Social Learning Theory and its application in chapter 6.

The focus then shifts in Part B to the research design and methodology. Chapter 7 shows how the literature discussed in Part A was used to inform the methodology, in particular the application of the TIB as a framework for analysing the results of the intervention. Chapter 8 then introduces the thirty commercial office buildings that were involved in the
study, including the pre-trial analyses and prototypes considered. The specifics of the intervention design are described in chapter 9, notably the ‘feedback’ component, including how it was generated and communicated and the ‘social learning environment’ that formed around the FMs. Chapter 10 explains the approach adopted for evaluating results, including how the commonly accepted conditions for validation were met, and the construction of an independent statistical model.

The results of the study are presented in Part C which is comprised only of chapter 11.

Part D is made up of four chapters: three discussion chapters and a short concluding chapter with recommendations. It begins with Chapter 12 which provides some context for the study in terms of its significance within the literature. Chapter 12 also considers in detail the characteristics of the sample and individual buildings in an evaluation of notable influences on the results. Chapter 13 then describes the various limitations and strengths of the study and considers whether or not the findings might be replicable in other settings. Chapter 14 gives a summary of my research’s main findings by positioning insights from the intervention within the literature. It addresses four main questions relating to the results: 1) the importance of motivation and goal-setting, 2) the role of ‘social learning’, 3) whether the skill-level of FMs actually increased as a consequence of the intervention and 4) whether savings are likely to persist over the longer term. This last question, while difficult to address in a closed-ended study, is perhaps the most crucial of all when considering the policy implication of this research. Chapter 15 then concludes the thesis with a series of recommendations for applying the research and further study.
1.3. Delimitations

The objective of this thesis is to contribute to a deeper understanding of the potential for behaviour-based interventions to be a source of energy savings and greenhouse gas emission reductions from buildings. The method adopted is quantitative and it focusses primarily on the energy performance data derived from the participating buildings. The personal characteristics, competence and motivations of the FMs are not evaluated directly. Rather, they are treated as intervening variables revealed in the relationship between independent and dependent variables. Likewise, an evaluation of specific organisational factors associated with the owner of the trial buildings—such as its policies, systems and procedures, culture and technology platforms—is outside the scope of this study.
2. BUILDINGS MATTER

2.1. Impact of buildings

According to the most recent assessment of the Intergovernmental Panel on Climate Change (IPCC), in 2010 buildings accounted for 51 percent of global electricity consumption, 32 percent of total global final energy use (i.e., energy supplied to the final consumer) and 19 percent of energy-related greenhouse gas emissions (Lucon et al., 2014). Greenhouse gas emissions from the buildings sector more than doubled during the period 1970 to 2010 (IEA, 2012, cited in Lucon et al., 2014) and, by most estimations, these figures are set to grow further. Projections from the World Business Council for Sustainable Development (WBCSD) suggest that worldwide energy consumption for buildings will grow by 45 percent from 2002 to 2025 and that demand from buildings will drive about half of all energy supply investments to 2030 (WBCSD, 2009).

Of the 32.7 PWh of energy used by buildings globally in 2010, one quarter (8.4 PWh) went to commercial buildings with the remaining three quarters (24.3 PWh) supplied to residential buildings (IEA, 2013, cited in Lucon et al., 2014). Commercial buildings\(^1\) account for a greater proportion of final energy demand in OECD countries where, according to the WBCSD, approximately one third of the typically 30-40 percent of total demand is for commercial buildings and the other two-thirds is for residential (WBCSD, 2009). The IPCC has projected that “without action, global building final energy use may

---

\(^1\) The IPCC and WBCSD place non-residential building types, including industrial, educational, health and all other commercial uses under the category “commercial”. In this chapter the generic “commercial” label is used when referring to non-residential buildings.
double or potentially even triple by mid-century, but with ambitious action it can possibly stabilize or decline as compared to its present levels” (Lucon et al., 2014, p. 57).

The scale of these impacts and opportunities vary from country to country depending on the energy mix and climate. In the U.S., buildings account for 40 percent of total greenhouse gas emissions and 70 percent of all electricity consumption (Hendricks et al., 2009). In Australia, a study commissioned by the Australian Sustainable Built Environment Council (ASBEC) estimated that 23 percent of the country’s greenhouse gas emissions are attributable to buildings, of which 43 percent (10 percent of the nation’s total) are attributable to commercial buildings (Allen Consulting Group, 2010). This is slightly at odds with IEA (2013) data cited by the IPCC (2014), showing that in the Pacific OECD region (which is comprised mainly of Australia and New Zealand), the commercial buildings sub-sector in 2010 accounted for 53 percent of the final energy used in buildings, up from 49 percent in 1990 – a relatively high proportion when compared with North America (44 percent) and Western Europe (33 percent). However, it appears that the overall usage in Australian commercial buildings is less compared with the UK, for example, where approximately 18 percent of emissions come from the commercial buildings sub-sector (Carbon Trust, 2009).

By far, the greatest proportion of building energy use occurs during the operational phase (UNEP, 2009). Though figures vary widely from building to building, and between building typologies, studies suggest that over 80 percent of greenhouse gas emissions are attributable to the provision of building services including heating, ventilation, and air conditioning (HVAC), water heating, lighting, entertainment and telecommunications (e.g. Adalberth et al., 2001; Junnila et al., 2006; Suzuki & Oka, 1998). Typically about 10
to 20 percent of energy is consumed in materials manufacturing and transport, construction, maintenance and demolition (Junnila et al., 2006; Suzuki & Oka, 1998).

**2.2. Energy use in buildings is growing**

The proportion of total primary energy (i.e. energy contained in raw fuels) supplied to commercial buildings in developed countries is significant and growing steadily (Allen Consulting Group, 2010; Lucon et al., 2014). But operational building energy use in developing countries, particularly in Asia, is growing at an alarming rate, as illustrated in Figure 2-1 below.

**Figure 2-1. Regional direct and indirect CO2 emissions from the building subsectors (Lucon et al., 2014)**

Between 1970 and 2010, indirect greenhouse gas emissions from commercial buildings (primarily from electricity use) more than quintupled globally from 0.48 to 2.52 GtCO2-eq/yr. Indirect emissions from residential buildings also grew significantly (from 0.80 to 3.50 GtCO2-eq/yr), but at a slightly slower rate (Lucon et al., 2014).
The IPCC does not separate out statistics by country, but a 2008 study commissioned by ASBEC found that the total energy demand from the Australian commercial buildings sector had “tripled since the 1970s” (CIE, 2008). That study suggested commercial property sector emissions are tied to GDP and would, therefore, continue growing at a compound annual rate of 2.1 percent per year until 2030, taking into account the projected growth in stock and Australian Bureau of Agriculture and Resource Economics (ABARE) projections that the rate of end use energy efficiency improvement will be maintained at 0.5 percent per annum (CIE, 2008; Syed et al., 2007).

Recent evidence suggests that Australian building operators are adopting energy efficiency programs and changing their usage patterns in response to rising electricity prices and other factors (AEMO, 2013; IPD, 2014) at a faster rate than was being projected only a few years ago. This is particularly evident in large commercial office buildings covered by the Commercial Building Disclosure Program (CBDP) which was established by the Building Energy Efficiency Disclosure Act 2010. The CBDP requires a current NABERS rating\(^2\) be displayed to prospective purchasers or tenants when commercial office space of 2,000 square metres or more of Net Lettable Area (NLA) is offered for sale or lease. According to a recent analysis commissioned from IPD (2014) by the Australian Government Department of Industry, introduction of the CBDP was associated with an immediate drop in the average NABERS Office Energy rating of offices within the IPD Property Database (comprised mainly of large institution-owned buildings), as previously un-rated buildings were forced for the first time to disclose

---

\(^2\) The National Australian Built Environment Rating System (NABERS) is a government-operated scheme that evaluates the environmental performance of Australian buildings against a set of benchmarks developed using national building performance data. The NABERS suite of rating tools measures performance on a 6 star scale, with 2.5 to 3 stars representing average performance. A 6 star rating demonstrates market leading performance, while a 1 star rating means the building is performing well below average market practice and has considerable scope for improvement. For more information, visit http://www.nabers.gov.au.
performance. Subsequently there was a steady improvement in ratings as performance of rated buildings improved, as can be seen from the following chart (Figure 2-2).

![Graph showing NABERS Energy rating improvement](image)

**Figure 2-2. Weighted Average NABERS Rating & Total Rated Area (IPD, 2014)**

The NLA-weighted average NABERS Energy rating of office buildings within the IPD Property Database improved by 1.2 stars during the period June 2006 to June 2014 (IPD, 2014), corresponding to a reduction in annual energy use of approximately 25 percent over the period despite aggregate energy consumption across the commercial building sector increasing. These results are obviously significant as they challenge the projections made by CIE (2008) and Syed et al. (2007) and appear to support the most recent assessment by the IPCC in its Fifth Assessment Report (AR5) that “final energy use may stay constant or even decline by mid-century, as compared to today's levels, if today's cost-effective best practices and technologies are broadly diffused (medium evidence, high agreement)” (Lucon et al., 2014, p. 4).

This IPD / Department of Industry data accords with selected observations from other jurisdictions and building types, as noted by the IPCC:
“The significant advances in building codes and appliance standards in some jurisdictions over the last decade already demonstrated that they were able to reverse total building energy use trends in developed countries to its stagnation or reduction. However, in order to reach ambitious climate goals, these need to be substantially up-scaled to further jurisdictions, building types, and vintages.”

(Lucon et al., 2014, p. 5)

One of the key challenges for comparing data between buildings and jurisdictions noted by the IPCC in AR5 was that actual greenhouse gas emission values vary significantly depending on the emission factors of the energy sources. This is particularly the case for indirect emissions in the sector, notably electricity, which account for the majority of emissions. Even within Australian jurisdictions covered by the research presented here, electricity emission factors (kgCO₂-eq/kWh) vary widely: NSW – 0.86, Victoria – 1.18, Queensland – 0.81 (Department of the Environment, 2014). Accordingly, this thesis adopts the approach of the IPCC and also focuses on final energy use, rather than emissions, as a measure of environmental impact and reduction opportunity.

2.3. Large institutions keen to be green

Australia has over 22 million square metres of commercial office accommodation (PCA, 2009). The majority of this is located in the mainland capital city Central Business Districts (CBD) in buildings controlled by institutional investors operating listed (i.e. quoted on the Australian Securities Exchange) and unlisted property trusts. Even though only five percent of commercial facilities in major office markets are larger than 5,000 square metres, they represent more than 50 percent of total NLA (MacDonald & Bray, 2011). The majority of these are owned by the large institutions. In Sydney CBD, for example, the Council of the City of Sydney estimates that 60 percent of office accommodation is controlled by just twelve separate entities (Barone, 2009).
As Figure 2-2 illustrates, the average NABERS Energy rating, which serves as a proxy for energy intensity in Australian commercial office buildings, has been steadily improving for much of the past decade. There are numerous factors driving the owners and operators of these buildings to pursue higher ratings and in doing so create more environmentally and socially responsible office accommodation. First among them is market forces. Corporate responsibility has become a significant factor in the decision making of large tenant organisations (i.e. the industry’s main customers) and with greater transparency about environmental performance, this desire to project an image of good corporate citizenship is influencing accommodation choices (Colliers International, 2010). Likewise, major property owners are competing to demonstrate ‘sustainability’ leadership credentials to their array of stakeholders, most notably investors who are increasingly concerned about environmental, social and governance (ESG) issues (including perceptions of other stakeholders, such as employees, tenants and business partners) impacting long-term financial performance (ACSI, 2013). These demand and supply side factors may explain the market evidence that environmental performance is increasingly being associated with asset ‘quality’. Energy efficiency gains (measured and communicated by NABERS Energy ratings) are therefore contributing to higher investment returns that significantly exceed the dollar value of the energy savings themselves (IPD, 2011).

2.4. Governments keen to be green

Governments too are encouraging this change. The Building Energy Efficiency Disclosure Act 2010 which introduced the CBDP in 2010, has had a profound impact on the level of transparency and accountability for environmental performance of large commercial office buildings in Australia. Phase one of the CBDP was introduced in 2010 as a voluntary measure. In November 2011, participation for qualifying buildings became
compulsory. Official statistics covering all 1,081 participating office buildings during the second year of full mandatory disclosure, from 1 November 2012 to 1 November 2013, show the average NABERS Energy rating was 3.03 Stars. This is considerably lower than the average shown in Figure 2-2 for the same period (approximately 4.1), with the difference between the average by number and by NLA explained by the fact that in recent times larger buildings have generally achieved higher NABERS ratings (Department of Industry, 2014b; IPD, 2014). Recognising the effectiveness of clear disclosure, and the leading performance of the larger end of the market, the Council of Australian Governments’ (COAG) National Strategy on Energy Efficiency includes provisions to consider expanding the program under a second phase to cover other types of commercial buildings such as hotels, shopping centres, schools and hospitals (Department of Industry, 2014b).

Complementing this policy are a variety of local-level initiatives aimed at overcoming barriers to improving the efficiency of privately owned and smaller office buildings, in particular those associated with finance for upgrades. For example, in 2010 the City of Melbourne and the state of New South Wales introduced schemes to facilitate “Environmental Upgrade Agreements” which allow building owners the opportunity to recover the cost of financing environmental retrofit works from tenants through charges linked to council rates collection (City of Melbourne, 2011; Office of Environment & Heritage (NSW), 2010). Despite active marketing, only a handful of EUAs had been taken-up some four years after their introduction (City of Melbourne, 2014).

2.5. **Design of buildings is a constraint**

Despite the spread in environmental performance observed between large and small office buildings and the significant efficiency advances being achieved in some parts of
the Australian commercial office building market, the sector as a whole is most remarkable for its homogeneity. Tenants evaluate office accommodation by referencing the Property Council of Australia’s Guide to Office Building Quality (PCA, 2011 and previous editions) which sets out guidelines for measuring the quality of mechanical services, lifts, hydraulic and electrical services, communications, security, amenities, parking, etc. Rents are determined on a ‘per square metre’ of NLA basis by taking into account these factors (which yield a ‘grade’ of Premium, A, B, C, or D) and adjusting for lease obligations, location and other aspects such as aesthetics and views. Commercial office building owners, who by definition are driven by commercial considerations, use these factors to guide the design and operation of buildings. And it is primarily for these reasons, aided by the advent of modern HVAC services, that office buildings throughout Australian CBDs and elsewhere around the world—regardless of location or time of year—must rely on advanced building services to provide indoor environmental conditions that sit within an acceptable range (Ackermann, 2002; Lucon et al., 2014; Piette, Kinney, & Haves, 2001). Except for the leakage associated with opening doors and leaky facades, these buildings are in effect ‘closed systems’ constrained by the first law of thermodynamics: work must be performed by an office building’s HVAC system to maintain any given temperature differential between its interior and the external environment. The combination of this physical law, ventilation standards and the thermal insulating properties of (mostly glazed) building envelopes can explain why a typical Australian commercial office building’s HVAC energy-use increases by six to seven percent for every degree Celsius change in the differential between internal and external temperature (Roussac, 2013; Roussac, Steinfeld, & de Dear, 2011a; Ward & White, 2007; H. Zhang et al., 2008).
Energy consumption in modern office and retail buildings typically ranges between 200–500 kWh/m².yr including all end-uses, down to less than 100 kWh/m².yr in advanced buildings, irrespective of climate. This is mostly a function of building design and the building services technologies deployed (Lucon et al., 2014; Palm & Darby, 2014). And yet, there is ample evidence from around the world that vernacular designs, which in many cases include passive features that have evolved over centuries in the absence of active energy systems, are able to meet similar comfort and health standards using as little as 1/10th the energy of commercial office buildings such as those typically found in Australian CBDs (Lucon et al., 2014; S. Zhang, et al. 2010). More flexible operating principles coupled with the adoption of relaxed dress codes that allow variable indoor thermal conditions from openable windows and part time and part space control of lighting and HVAC, can be more important for total energy use than the efficiency of the air-conditioning plant and equipment. Such cultural factors are very important for energy usage (Lucon et al., 2014). For example, studies of Asian buildings (e.g. Jiang, 2012; Murakami et al., 2009; Zhaojian & Jiang, 2007) have found that often buildings featuring high performance centralised plant (including high profile ‘green buildings’ built in international collaborations) can use up to nine times more energy than decentralised split units that operate part time and for partial space cooling: “Factors of … up to 10 times difference in office building energy use [have been found worldwide] with same climate and same building functions with similar comfort and health levels” (Lucon et al., 2014, p. 28).

2.6. Factors explaining a building’s daily energy use

This is not to say the total energy (or work) required by HVAC systems is pre-determined by building design and the flexibility of occupants rather than the actions of building
operators. Furthermore, HVAC systems typically only account for approximately 30 percent of the energy used by a ‘whole’ commercial office building – or about 60 percent of the ‘base’ building (which comprises the whole building’s services minus all the lighting, IT and any supplementary HVAC services utilised by occupants in the tenanted spaces) (Harvey, 2013).

The amount of energy used to operate a commercial office building is influenced on a day-to-day basis by a variety of factors: some constant, such as the type and quality of the façade and the energy efficiency of the plant and equipment; some dynamic, such the changing weather conditions and occupancy levels; some random, such as unexpected calls for additional after-hours air-conditioning or emergency situations; and some relating directly to the actions and decisions of management, including the way the building’s services are tuned, when plant and equipment is scheduled to start and stop, and the amount and quality of services provided to occupants (i.e. temperature setpoints and ventilation rates which, of course, are dependent on the flexibility of occupants as discussed above).

The relative importance of each of these factors varies from building to building and, as changes are made, over time as well. Furthermore, there are interrelationships between each of the factors. For example, the introduction of enhanced management practices (such as optimising equipment start and stop times) will have a larger impact on energy use in a building where inefficient plant is installed and thermal loads are exacerbated by a poorly insulated façade. Similarly, replacement of a highly inefficient chiller will bring about only a small energy saving if it is rarely required to operate.
2.7. Opportunities through technology and non-technological interventions

An adjustment to any of the four variables mentioned above will impact the amount of energy used by a commercial office building; however, the second (dynamic factors such as weather and occupancy) and third (random factors such as changes in demand that cannot be predicted reliably) tend to be outside the control of building operators. On the other hand, the first (constant factors like characteristics of physical plant and building envelopes) and fourth (management techniques and decisions) are well within the bounds of building operator/owner control.

The extent of the opportunity to reduce energy use and greenhouse gas emissions from addressing the first category (hereinafter referred to as “technology”) has been well documented (see Barker et al., 2007; Levine et al., 2007; Lucon et al., 2014). As discussed in the IPCC’s Fourth Assessment Report (AR4), essential considerations in the design of low-energy buildings are: (1) building orientation, thermal mass, and shape; (2) high-performance envelope specification; (3) maximisation of passive features (day-lighting, heating, cooling, and ventilation); (4) efficient systems meeting remaining loads; (5) highest possible efficiencies and adequate sizing of individual energy-using devices; and (6) proper commissioning of systems and devices. Numerous studies have found that retrofits incorporating combinations of the above features routinely achieve 25–70 percent savings in total energy use, with capital expense recouped over time via the savings (Harvey, 2009; Levine et al., 2007).

Management (hereinafter referred to as “non-technological” aspects), on the other hand, has not been nearly as thoroughly assessed. The 60 page chapter on residential and commercial buildings in AR4, for example, provides a clear status of the literature
regarding non-technological options at the time: “the potential reduction through non-technological options is rarely assessed and the potential leverage of policies over these is poorly understood” (Levine et al., 2007, p. 406). The authors of AR4 went on to note “substantial market barriers that need to be overcome” and an exceedingly slow uptake of “accessible and cost-effective technologies and know-how … which can abate greenhouse gas emissions in buildings to a significant extent … (high agreement, much evidence)” (p. 406). Accordingly, within AR4 “the potential [greenhouse gas emission] reduction through non-technological options is not assessed” (p. 409).

AR4 estimated the economic potential for emission reductions from different sectors (or ‘wedges’) of the global economy as a function of carbon price in 2030 and concluded that actions to enhance the building stock’s technology present by far the largest opportunity for cost effective mitigation across all sectors, primarily from retrofitting existing buildings and replacing energy using equipment (Barker et al., 2007; Levine et al., 2007). This is summarised in Working Group III’s contribution:

“Our survey of the literature (80 studies) indicates that there is a global potential to reduce approximately 29% of the projected baseline emissions by 2020 cost-effectively in the residential and commercial [buildings] sectors, the highest among all sectors studied in this report (high agreement, much evidence).”

And yet, the report goes on to say that:

“Due to the limited number of demand-side end-use efficiency options considered by the studies, the omission of non-technological options and the often significant co-benefits, as well as the exclusion of advanced integrated highly efficiency [sic] buildings, the real potential is likely to be higher (high agreement, limited evidence).”

(Levine et al., 2007, pp. 389-390)
The IPCC in 2007 also highlighted a wide range of possible emission reduction scenarios for the building stock in developed countries in 2020 (against a 2010 start year):

- Technical: 21% - 54% (i.e. with no specific reference to costs, only to ‘practical constraints’)

- Economic (<US$ 0/CO_2\text{-eq}): 12% - 25% (i.e. from the perspective of society in general)

- Market: 15% - 37% (i.e. from the perspective of utility-maximising private consumers and companies)

(Barker et al., 2007)

While AR5 goes a long way further in its attempt to highlight the importance of non-technological aspects by providing many examples of significant behavioural change opportunities (refer to those cited in Lucon et al., 2014, p. 18), they are not condensed into a figure. Rather, the report provides the following summary observation:

“In addition to technologies and architecture, behaviour, lifestyle, and culture have a major effect on buildings’ energy use, presently causing 3–5 times differences in energy use for similar levels of energy services (limited evidence, high agreement). In developed countries, evidence indicates that behaviours informed by awareness of energy and climate issues can reduce demand by up to 20% in the short term and 50% by 2050.”

(Lucon et al., 2014, p. 5 – emphasis in original)

Such estimates are derived by taking a wide variety of factors into account (Eyre, et al., 2010; Fujino et al., 2008) and summing their contributions. Those that relate most directly to the study described here, and in particular the provision of “automated daily feedback” to operators of Australian office buildings, will be discussed in the following pages and
the scale of the potential opportunity to improve the sector’s energy efficiency if these experiences are replicated is discussed at the conclusion of this thesis.

2.8. Comfort and IEQ considerations

There is an underlying pattern of growing dependence on HVAC systems to meet the comfort ‘requirements’ of commercial building occupants, especially in developed countries such as Australia where the costs of salaries and benefits typically amount to 100 times energy expenditure on a per square metre basis (Woods, 1989). The cognitive performance and productivity of occupants is increasingly being found to correlate with their indoor environmental quality (IEQ) and comfort (Allen et al., 2016). It stands to reason, therefore, that in commercial offices where individual occupants have very limited control over their surrounding environments, provision of a comfortable workplace that can satisfy a majority of occupants is often regarded as the primary measure of an FM’s performance (Kim et al., 2013).

Given that money is most often identified as the main motivational driver behind energy saving behaviours (Martiskainen, 2007), any trade-off between occupant comfort or wellbeing and actions aimed at reducing energy expenditure for environmental reasons might be regarded as financially counter-productive and therefore problematic. Actions that aim to save energy for environmental reasons should either seek to generate co-benefits including enhanced comfort and / or productivity gains, or alternatively (or in addition) engage with occupants and other stakeholders around the notion of adequacy or ‘sufficiency’ (Herring, 2006; Oikonomou, et al., 2009). Identification of co-benefits can generally help with the ‘sell’ to building owners and occupants, whereas the imposition of constraints aimed at capping or discouraging increasing energy use due to increased floor space, comfort levels, and equipment would likely require policy intervention (Lucon et
al., 2014), which is unlikely. Currently more than a half of the final energy demand from buildings globally is for space heating and cooling and yet if current policy plans and efforts in developed countries are implemented, HVAC demand in commercial buildings is expected to grow by 83% over the period 2010–2050 (Ürge-Vorsatz et al., 2013).

On a more positive note, there is growing evidence of an association between energy savings, the use of adaptive indoor temperature setpoints that take into account ambient outdoor conditions, and a reduction in the number of comfort complaints from building occupants. Where observed, this association may be a reflection of more vigilant operation and tighter building control, suggesting better managed office buildings can lead to more comfortable and productive occupants (Allen et al., 2016; Bordass, et al., 2001; Fisk, et al., 2012; Piette et al., 2001; Roussac, et al., 2011b).

The economic benefits are potentially large, both in the case of new ‘green’ construction and also from enhancing existing buildings. Fisk estimated the potential annual savings and productivity gains from direct improvements in worker performance that are unrelated to health are, for the United States alone, between USD$20 to $160 billion per annum (Fisk, 2000). That study identified numerous energy efficiency measures that also improve IEQ in a manner that often significantly increases productivity and health. Similar claims (about effects of improved IEQ on health and productivity, generally without the financial extrapolation contained in Fisk) have been made in relation to “green” (e.g. LEED rated) buildings although most have been qualitative in nature or involve quite small sample sizes (Altomonte & Schiavon, 2013; Birt & Newsham, 2009; Singh, et al., 2010). One influential study, by Singh et al., sought to longitudinally evaluate these impacts by collecting survey data from employees before and after moving from conventional to LEED-rated buildings and found improvements in perceived
productivity and health symptoms which supported previous qualitative studies about the benefits of green office buildings (Singh, et al., 2010).

Whether these benefits are more prevalent in ‘green’ (i.e. LEED, Green Star rated) buildings, rather than well-managed conventional buildings, is unclear, however. A recent study by Altomonte & Schiavon (2013) has called into question claims about the productivity benefits of LEED rated buildings, finding “occupants of LEED certified buildings have equal satisfaction with the building overall and with the workspace than occupants of non-LEED rated buildings”. That study included 144 buildings (65 LEED certified) and 21,477 individual occupant IEQ survey responses (10,129 in LEED buildings) (Altomonte & Schiavon, 2013). Their findings are significant because they indicate that operational factors within the control of the FM, as distinct from design, may be important in determining occupants’ IEQ experience. The challenge of balancing energy saving goals with occupant comfort and IEQ requirements may therefore be ‘front of mind’ for many FMs.

### 2.9. Buildings getting greener, but more complex

The ‘green building’ industry has been growing rapidly for many years (USGBC, 2009) and there are indications that the prevalence of rating systems may be contributing to an increase in the technological complexity of modern commercial buildings and a reduction in their energy use (GBCA, 2013; IPD, 2014). Even if the causal factors are not clear, there is growing evidence suggesting that new Australian office buildings are more energy efficient than buildings completed in the years prior to the advent of NABERS and the Green Building Council of Australia’s Green Star rating scheme (GBCA, 2013; Warren Centre for Advanced Engineering, 2009). As already discussed, it is clear that ‘technology’ improvements have been contributing to these results; however, the extent to
which these carefully designed higher-performance buildings are meeting their design potential is unclear. To what degree are non-technological aspects contributing, and can they contribute more?

Without vigilant operations and maintenance, many highly complex buildings will not perform as designed; a problem first highlighted by Bordass et al. after they reviewed in-operation performance of UK buildings and found that more and higher technology without a commensurate increase in management input was ‘a hole that many buildings fall into’ (Bordass et al., 2001). Building operations personnel with the background and skills required to meet greater performance expectations from increasingly complex buildings are difficult to source, and the demands being placed on them continually grow. Not surprisingly, given the focus for most commercial office building owners is on keeping tenants (their customers) happy, many FMs are appointed for customer-relationship management skills rather than technical qualifications. In conducting the research for this thesis it was found that many FMs had experienced no technical training whatsoever.

This combination of complexity and skills shortages is giving rise to a significant marketplace for building automation software and facilities resource management (FRM) products and services (MacDonald & Bray, 2011). In 2015 the global marketplace for FRM solutions was expected to reach USD$36 billion, part of the rapidly evolving (largely technology-based) market for energy efficiency products and services to buildings which is forecast to reach USD$103.5 billion by 2017, up more than 50 percent from 2011 (Pike Research, 2011). It is not yet clear whether the emergence of this market will drive further improvements in individual high-performance buildings, whether it will
supplant and contribute to further down-skilling of building operators, or whether it will instead empower them.

One thing that is clear, however, is that new commercial buildings sporting the latest technology will only ever make up a small proportion of the sector overall. At any point in time more than 80 percent (currently about 18 million square metres) of Australian commercial office accommodation is 10 years old or older (Davis Langdon, 2009). Given that the NABERS and Green Star rating schemes only came into existence in 1999 and 2003 respectively, the vast majority of existing stock has been designed and built without those drivers being taken into consideration. Accordingly, the focus of this thesis is on addressing the question of whether the energy-efficiency of existing Australian office buildings can be significantly improved without resorting to costly technology upgrades or relying on an army of highly skilled technicians.
3. RESEARCH IS REQUIRED

3.1. Previous studies focussed on residential

The majority of research relating to behaviour-based approaches for improving energy efficiency in buildings is focussed on the residential sector (CSE / ECI, 2012; Lopes et al., 2012; Lutzenhiser, 1993). Studies have addressed a wide array of promising innovations, including Home Energy Management Systems (HEMS) which have generated a particular interest recently. HEMS generally take the form of energy monitors, also known as In-Home Displays (IHDs) – “intermediary products that can visualize, manage, and/or monitor the energy use of other products or whole households” (van Dam et al., 2010, p. 458). The residential sector provides a reliable testing ground for such approaches because dwellings and occupants are relatively numerous and able to be recruited to studies at scale. Furthermore, selections and interventions can be effectively randomised and controlled, thus facilitating randomised controlled trials (RCT) and rigorous statistical analysis (Todd et al., 2012).

Non-residential buildings, on the other hand, tend more often than not to be owned and operated by commercial or government entities that are often cautious about participating in studies – particularly those involving physical actions that impact building operations or independent evaluation of performance. This recruitment barrier can make randomisation of samples difficult. Likewise, the difficulty of matching uses, sizes, designs, ages etc. between intervention and control groups limits the possibilities for inferential statistical analysis (for a discussion of requirements, see Freedman et al., 2007). These factors may help explain why a thorough review of journal articles published during the period 2000-2011 found “… despite the significant existing background knowledge on energy behaviours in the residential sector, behavioural studies
... are almost inexistent for other type of buildings such as services [i.e. any non-residential typology]” (Lopes et al., 2012, p. 4098).

Notwithstanding the obvious differences and associated limitations, residential studies do provide useful indicators of the potential for behaviour-based approaches to save energy in commercial buildings. Faruqui et al. (2010), for example, reviewed the literature going back at least four decades and found “studies have consistently demonstrated that direct feedback motivates behaviour change, resulting in energy savings ranging up to 20 percent” (p. 1599). That study involved a review of a dozen utility-managed IHD pilot programs and found savings in the order of 7-14 percent. Those results are supported by the findings of Darby (2006) who reviewed the literature on residential metering, billing and direct displays for the British government and found that direct feedback (i.e. immediate, from the meter or an associated display monitor) reduced energy usage in households by 5-15 percent on average (Darby, 2006) and Delmas et al. (2013) who reviewed 156 residential field trials published from 1975 to 2012 and found an overall saving of 7.4 percent (Delmas et al., 2013). Other recent studies have suggested savings in the range of 20–30 percent are achievable across the residential sector (see for example Gardner & Stern, 2008; Laitner et al., 2009; WBCSD, 2010), a range supported by the WBCSD’s estimate (p. 62) that “wasteful behaviour can add one-third to a building’s designed energy performance, while conservation behaviour can save a third.”

3.2. Information alone is not enough

While indications about the benefit of direct feedback approaches are quite strong, much of the research shows that “information” by itself is generally not enough to change households’ energy consuming behaviour and indirect feedback, e.g. raw data processed by the utility and sent out to customers, has in many cases been found to achieve no
noticeable effect and even in some cases have a negative effect (see for example Bittle et al., 1979-80; Brandon & Lewis, 1999; Carrico & Riemer, 2011; Darby, 2006; Henryson et al., 2000). Often people simply fail to comprehend or absorb the energy saving information, or fail to follow through with action. Information combined with other measures, such as feedback on energy use coupled with advice and help in interpreting the feedback, is therefore considered much more effective in contributing to behavioural change (see for example Abrahamse et al., 2005; Darby, 2006; Martiskainen, 2007). One of the first to identify this particular issue was Geller (1981) who obtained questionnaire data from 117 subjects that attended three-hour long energy-conservation workshops near Richmond VA (U.S.A.) and then visited 40 of the workshop attendees in their homes and also the homes of 40 non-attendees. It was found that workshop attendees showed few applications of the energy-conservation strategies emphasised at the workshops and no appreciable difference in behaviours when compared to the control group (Geller, 1981). One explanation, from a Swedish study, is that householders often feel “overloaded” by energy efficiency information and yet lack the basic knowledge to apply that information in their own homes (Henryson et al., 2000). This calls for tailoring and personalisation of information, rather than provision of advice which individuals may find difficult to adapt to their circumstances (Martiskainen, 2007). Brandon and Lewis (1999) put it succinctly: “Consumers want customised or particularised advice” (p. 84). In their study, conducted across 120 households in Bath, U.K. over a nine month period during 1995-96, the only feedback form in which the authors felt they were “able to place any confidence” (p. 84) was information supplied by computer software (in this case software that included a year on year comparison, a questionnaire on general aspects of energy saving and a directory of energy saving information and advice) (Brandon & Lewis, 1999).
3.3. *Some studies lack methodological strength*

The range of results reported in the literature and the uncertainty expressed about the effectiveness of various approaches can be explained, at least in part, by inherent limitations in the design of many of the reported studies. As Figure 3- illustrates, the number of studies involving the application of behaviour-science techniques to the problem of energy saving grew rapidly throughout the 1970s and then gradually declined from the early 1980s onwards: a pattern that reflects the concern over the OPEC oil price shocks of the mid-1970s and its subsidence with the recovery in the world economy (for a commentary on the growing interest in energy conservation in the 1970s and the social and institutional barriers it faced, and still does, see Blumstein et al., 1980). Most of these early studies which sought to analyse information-based measures were not specifically designed to allow for meaningful comparisons between alternative intervention strategies; their aim was to focus attention to particular conservation measures and approaches – many for the first time. Replicability was also limited in many cases due to
methodological weaknesses; for example by limiting focus to only one intervention (and therefore not drawing attention to alternative explanations for results) or, where multiple interventions were identified, confounding results by failing to isolate the contribution of individual measures (Dwyer et al., 1993). Many of the studies have also been criticised for not being representative of the wider population due to small sample sizes and selection bias (for critiques, see for example Bittle et al., 1979; Delmas et al., 2013; Hayes & Cone, 1981).

Limited progress has been made in this regard since. Some 25 years later Martiskainen (2007) observed that “many of the existing studies use small sample sizes, are prone to selection bias, fail to include a control group or have other methodological weaknesses … [making] it difficult to estimate the potential impact of different forms of interventions with any confidence” (p. 6). Such criticisms are not easy to avoid, of course; a point highlighted in the IPCC’s AR5 where its authors achieved “high agreement” but found “limited evidence” to support the proposition that “behaviour, lifestyle, and culture have a major effect on buildings’ energy use, presently causing 3–5 times differences in energy use for similar levels of energy services” (Lucon et al., 2014, p. 5).

Even where significant numbers of subjects can be recruited to intervention and control groups, issues of selection bias and representativeness are difficult to overcome because there are so many factors to be accounted for – a challenge being compounded now that addressing environmental problems such as climate change has become one of the primary reasons for studying energy conservation (Abrahamse et al., 2005). Take, for example, the ideological persuasion of participants. One study based on data from almost 50,000 households in Sacramento, California who received home energy reports provided by Opower, a company that works with energy utilities to provide customers normative
feedback, found that feedback conveyed to politically liberal households about their own and peers’ electricity usage is two to four times more effective than it is with politically conservative households who are much more likely to opt out or ignore such feedback (Costa & Kahn, 2013). Concerned by this finding, Opower conducted similar research based on large samples from other communities in the U.S. and found negligible, and in some cases opposite, effects (Curtis, 2013). Caution must therefore be applied when extrapolating results from studies conducted in one community to another, and even more so when considering the implications of a study in one country for another where significant cultural, technological and climatic differences exist.

3.4. Uncertainty about the drivers of energy behaviours

One thing is clear: there is not yet enough evidence to definitively say which intervention measures lead most reliably to the kinds of behavioural changes (or combination of changes) able to produce sustained energy savings in the order of the 20–30 percent cited above. Thorough and relatively recent reviews (for example those by Andrews & Johnson, 2016; Martiskainen, 2007; Stern et al., 2016; van Dam et al., 2010) consistently call for more robust research to untangle the complex web of factors that account for significant savings in some contexts and negligible savings in others. In relation to commercial buildings, the need is even more pressing. Clearly there have not been enough studies focussed on commercial buildings and those that do exist suffer from many of the same limitations as the residential studies, but compounded by the difficulties associated with recruitment and also diversity of building design and use described above.

It can be said with confidence, however, based on the many reviews already cited, that bespoke “feedback” strategies are emerging as the category of behaviour-based interventions with greatest potential to significantly influence energy efficiency. And yet
there is no general prescription for how feedback should be structured and delivered because individual behaviours are deeply embedded in social and institutional contexts and reactions, therefore, can vary widely (see for example Abrahamse et al., 2007; Gardner & Stern, 2002; Jackson, 2005; Lopes et al., 2012; Martiskainen, 2007). This was exemplified in a finding from a recent U.S. study that considered the effectiveness of energy information systems (EIS) deployed across 26 portfolios of commercial buildings. The factors most strongly associated with high energy savings were a building’s previous energy use intensity (EUI) (i.e. the higher the usage, the greater the savings) and the extent of efficiency projects already deployed prior to EIS installation, whereas “user empowerment” and “user engagement” with the EIS installation was found to have had very limited impact on energy savings (Granderson, Lin, & Piette, 2013). As Martiskainen (2007, p. 11) observed, “a general theme is emerging from the earlier research – energy efficiency and consumer behaviour have been puzzling researchers for over 35 years and still are…”

3.5. Opportunity

Throughout 2007-2008, The Warren Centre for Advanced Engineering at the University of Sydney engaged a large section of the Australian commercial property sector involved in the building management chain (owners, operators, designers, consultants, contractors) to participate in its Low Energy High Rise (LEHR) project. The objective was to evaluate non-technical barriers to the upgrading of large commercial office buildings (defined as office buildings exceeding 7,500m² NLA), including barriers to the procurement and deployment of technology solutions. It was an unusual undertaking for a centre for “advanced engineering” and came about from an acknowledgement that “buildings are significantly human rather than merely technical systems and that the key lies in balancing relatively well trodden technical issues with the far more challenging human
and organisational issues” (Warren Centre for Advanced Engineering, 2009, p. 11). The LEHR project analysed survey and utility meter data from 127 large commercial office buildings in Australia’s capital cities in an attempt to identify factors that influence building performance, as measured by NABERS. The results were significant, with data suggesting “there is potential for most buildings to achieve a 4 Star NABERS Energy base building rating with only limited recourse to major technical refurbishment, which corresponds to a performance improvement of approximately 30% for an average building” (p. 9) (emphasis added). The report found that “management” can have an impact of 1.3 stars, the difference between an average performing building and a 4 Star NABERS Energy rated building regarded at the time as “excellent energy performance”.

3.6. More data needed

The LEHR study was based on surveys and energy data collected for each building on one occasion only. There was no follow-up and no attempt made to study longitudinally the patterns of improvement / decline. This is unfortunate, because as Lucon et al. observed in the IPCC’s AR5, a more detailed understanding of the potential for efficiency improvements and co-benefits would be gained if operating dynamics and performance could be continuously monitored – thus providing “better feedback to the policymaking process, to education, to capacity building, and to training” (Lucon et al., 2014, p. 67).

The focus of this thesis therefore centres on the question of how to stimulate, develop and maintain the ‘motivation’ of non-residential building operators to improve energy efficiency and, in doing so, enhance their knowledge, skills and the performance of their buildings. The technique I’ve developed falls into the categories of ‘feedback’ and ‘social learning’. They will be discussed in chapters 5 and 6 respectively before I move on to describe the methods employed.
4. HUMAN MOTIVATION AND BEHAVIOUR

4.1. Motivation drives behaviour

A glance at the shelves of any airport bookstore will reveal myriad publications on the subjects of motivation and behaviour change generated for a public eager to imbibe their recipes for success. Most titles simplify and popularise concepts found in the scientific and peer-reviewed literature, so a review is not necessary. For a flavour, however, here is a quote from one of the genre’s more prominent and relatively rigorous recent examples: “We can say this much with confidence: When change works, it tends to follow a pattern. The people who change have clear direction, ample motivation, and a supportive environment” (Heath & Heath, 2010, p. 255).

According to Nevid (2013), the term motivation “… refers to factors that activate, direct, and sustain goal-directed behaviour... Motives are the ‘whys’ of behaviour – the needs or wants that drive behaviour and explain what we do. We don't actually observe a motive; rather, we infer that one exists based on the behaviour we observe” (p. 288). But energy consumption is not behaviour. It is a consequence of behaviours like adjusting thermostats or scheduling the operation of plant and equipment (Becker et al., 1981). Therefore, in order to change the patterns of building energy use through behaviour-based approaches, it is important to address the underlying motivations (and habits) of the operators – the “whys” of their behaviour.

Motivation is derived from emotional, social, biological, or cognitive forces and it has three core components: activation, persistence, and intensity (Cherry, 2015; Nevid, 2013). Each of the three components can be considered essential for initiating and maintaining activities that deliver higher energy efficiency in buildings. Activation is the initial phase where a decision is made to act or implement something; Persistence is the effort to
continue toward a goal; and Intensity is the level of concentration and vigour applied to pursuing it.

Human motivations are multi-faceted and highly complex, and there is no single theory that can adequately explain all their dimensions. Furthermore, they are often embedded in ordinary, routine and habitual behaviours (i.e. unconscious actions) which are themselves heavily influenced by a variety of social, institutional and cognitive constraints (Jackson, 2005). Following are a few of the most prominent theories of motivation which take on different levels of importance depending on the context being considered.

**Instinct Theory** – inspired by the theory of Evolution, it suggests people are motivated to behave in certain ways as a result of biological, genetic programming (see McDougall, 1918). This theory has been controversial, in part because of its associations with eugenics, and has been largely discarded by mainstream psychology although the influence of genetics and heredity on human behaviour remains relevant to the field of evolutionary psychology.

**Drive Theory** – people are motivated to minimise the homeostatic disturbance (internal tension) that arises when psychological needs are not satisfied. This theory derives from the work of Freud in the 1920s and was first formally proposed by Clark Hull (1943) and systematised by the Hungarian psychiatrist and psychologist Leopold Szondi (1947 [1952]). According to this theory, ‘drive’ tends to progressively build up over time and is regulated by a feedback loop, much like a building’s air conditioning system.

**Arousal Theory** – people try to maintain an optimum level of physiological and psychological arousal, i.e. their sensory alertness, mobility and readiness to respond to stimuli. The comfortable or “optimum” level varies from person to person and research in the field focusses on the neural systems involved in what is collectively known as the
arousal system (neurotransmitters, dopamine, histamine, acetylcholine, norepinephrine, and serotonin).

**Incentive theories** – people are motivated to do things because of rewards. These rewards can be either internal (intrinsic) or external (extrinsic) and may either pull people towards a goal or push them to avoid undesirable outcomes (see Bernstein, 2011). Major contributions to the field came from Skinner's theory of ‘operant conditioning’ (1948) drawing on the 'Law of Effect' developed by Thorndike (1905), both of which will be described further below.

**Content theories** – people will focus on satisfying their most basic biological needs and, once they are met, will progressively give higher priority to addressing unsatisfied needs, ranging up to fulfilment and realisation of personal potential. These theories stem from Abraham Maslow’s (1943) paper *A Theory of Human Motivation* where he described man as “a perpetually wanting animal” (p. 370), were expanded upon in his (1954) book *Motivation and Personality*, and were subsequently extended by Frederick Herzberg’s two-factor theory (1959) where he introduced the distinction between “motivators” which give satisfaction, and “hygiene factors” which are not motivational but will result in demotivation if they are not present. These and other content theories remain highly influential; however, important limitations have been identified – particularly in the ranking aspects and for their disregard of personality traits. These weaknesses may be traced to the fact that Maslow’s primary research focused on unrepresentative demographic groups (exemplary people; the healthiest one percent of college students) primarily from the United States (see for example Goebel & Brown, 1981; Hackman & Oldham, 1976).
Of all the theories of motivation mentioned above, incentive theories hold the most promise for the development of levers to influence energy use in buildings. Instinct, drive and arousal theories have helped shape our understanding of why humans are the way we are. Likewise, the various content theories help us to understand why we make the decisions we make. Incentive theories, on the other hand, not only give us deeper insights into human behaviours, but they also suggest tools for manipulating the patterns of those behaviours.

4.2. Behaviour theories – conditioning

Incentive theories are derived from the ‘behaviourist’ school of thought that emerged from the work of Russian physiologist Ivan Pavlov during the 1890s. Pavlov noticed that his dogs appeared to associate him with food and would begin to salivate whenever he entered the room they were in – even if he had no food with him. He decided to conduct an experiment where he would ring a bell and then feed the dogs. After a few repetitions he observed the dogs salivating in response to the bell. Thus emerged the field of Classical conditioning where in this case the bell becomes a conditioned (or conditional) stimulus because its effect is dependent on its association with food and the salivation is the conditioned (learned) response (Cherry, 2005). The term ‘behaviourism’ was first coined by the psychologist John B. Watson who in his (1913) paper, Psychology as the behaviorist views it, presented a view of psychology as “a purely directive experimental branch of natural science … its theoretical goal is the prediction and control of behaviour” (p. 158). Watson was a strong proponent of Pavlov’s work; however, the ethics of his famous experiment to condition fear of a white rat into an 11-month-old orphan "Little Albert" by clanging a pot each time the child saw it, makes him a controversial figure in the history of psychology. His extreme belief in the defining role
of ‘nurture’ over ‘nature’ was also controversial, and deliberately provocative, as it pitted him squarely against eugenicists whose ideas were popular at the time (Hothersall, 2004).

Whereas Watson held the firm view that psychology should focus on the "behaviour" of the individual rather than the “mind” (Hothersall, 2004), B.F. Skinner simply believed that it was more practical to focus on behaviour, stating: "It is toward the reduction of seemingly diverse processes to simple laws that a science of behaviour naturally directs itself" (Skinner, 1938, p. 425). Skinner was influenced by Thorndike’s (1905) 'law of effect', which states that a behaviour is likely to be repeated if it leads to satisfying consequences, whereas one that leads to undesired outcomes is likely to be stopped. The two men employed similar experimental methods using operant conditioning chambers.

Skinner’s work extended Thorndike’s by rejecting all references to unobservable mental states, such as ‘satisfaction’, and focussing more explicitly on observable behaviour and its consequences (Skinner, 1950). Crucially, he introduced the concept of “reinforcement”, a key element of what he termed Operant conditioning (Skinner, 1938). Unlike classical conditioning, where behaviour (e.g. salivation) is influenced by pairings of a neutral stimulus (e.g. bell) and a potent biological stimulus (e.g. tasty food), operant conditioning seeks to change behaviour through reinforcement and punishment, as illustrated in Figure 4-1.
Figure 4-1. Operant conditioning diagram

Skinner identified two types of ‘operant’ that can be used to influence behaviour: *reinforcers* that increase the likelihood of a behaviour being repeated and *punishments* that decrease the likelihood. He also recognised *neutral operands* that have neither effect.

Both positive and negative forms of reinforcement strengthen behaviour. Punishment, on the other hand, is designed to weaken behaviour. A positive punishment involves the introduction of an unpleasant stimulus (e.g. penalty) and a negative punishment involves the removal of a potentially rewarding stimulus (e.g. denial of an anticipated benefit). The distinctions can obviously become confusing and less clear-cut outside the experimental chamber. Financial bonuses, for example, which are commonly paid each year to the managers of commercial buildings, are designed to reward (reinforce) good performance. However once a building manager comes to expect a certain level of bonus, a payment of less than the anticipated amount can be perceived as punishment. This will have the effect of weakening behaviour, even if the employer’s intention in paying the bonus was to reward it.
4.3. Behaviour modification

Behaviour modification involves the systematic use of operant conditioning techniques to reinforce desired behaviours and ignore or punish undesired ones, mainly by changing environmental stimuli that relate to a person's behaviour (see Skinner, 1974). It divides positive reinforcement into two categories: primary reinforcement, which is essentially biological and does not require any learning (e.g. food) and secondary reinforcement, also known as conditioned reinforcement, which works through its association with a primary reinforcer. For example, if an FM is given a day off work for doing something impressive that saves energy, the day off would be primary reinforcement. If instead the FM receives a message from the boss saying “congratulations, well done – keep up the good work!” the message is secondary reinforcement because it is taken to indicate something desirable may happen in the future (money is also generally regarded as a secondary reinforcer – it provides the means to acquire things like food, shelter, holidays, etc.). Such use of praise and encouragement as a form of positive reinforcement is widely recognised as being effective in the workplace and elsewhere. In one prominent study, it was observed that the highest-performing workplace teams maintain an average ratio of 5.6 positive comments (reinforcement) for every negative one (punishment) they make to one another. The lowest-performing teams, on the other hand, were found to make 2.8 negative comments for every positive one (Losada & Heaphy, 2004). The 5:1 ratio has also been found to hold in other settings, for example as a predictor of marital happiness and divorce (Gottman, 1993).

4.4. Behaviourism provides an incomplete account

For all its merits and continued appeal, Skinner’s proposition that humans learn behaviour in much the same way as caged rats learn to press a lever (Skinner, 1974) is nowadays
regarded as an inadequate explanation for human behaviour and learning. Many prominent critics, including Sigmund Freud, Leon Festinger and Noam Chomsky, have argued that behavioural theories do not account for the influence of the unconscious mind's thoughts, feelings, and desires, nor do they account for other types of learning derived from thinking or mental processes such as problem-solving, decision-making and the acquisition of language (see Mills, 1998). Chomsky’s (1959) review of Skinner’s book Verbal Behaviour, published in 1957, was highly critical of the behaviourist’s “gross and crude” oversimplification of linguistic behaviour and accused it of overlooking “just about everything of interest” (Chomsky, 1959, pp. 54-58). His critique, summarised in the following sentence, is credited with giving rise to the cognitive movement in psychology and the gradual decline of behaviourism.

“One would naturally expect that prediction of the behavior of a complex organism (or machine) would require, in addition to information about external stimulation, knowledge of the internal structure of the organism, the ways in which it processes input information and organizes its own behavior.”

(Chomsky, 1959, p. 27)

Leon Festinger, a pioneer of social psychology, is most widely recognised for introducing the theory of ‘cognitive dissonance’ – people’s tendency to seek to resolve the discomfort that arises when they hold two or more competing beliefs (or ideas or values), either by reducing it or avoiding it in the first place (Festinger, 1957). Festinger’s demonstration of this effect in laboratory experiments was a major challenge to behaviourist theory and highlighted the importance of considering the ‘internal structure of the organism’, as suggested by Chomsky. Likewise, his description of the way people tend to evaluate their own abilities and opinions by comparing them against others’ introduced social comparison theory (Festinger, 1954) and drew attention to another of behaviourism’s
limitations. Social comparison theory, and advances made since, has contributed to our understanding of humans’ ability to learn by observing others’ behaviour and its consequences – sometimes referred to as vicarious reinforcement. This concept of imitative learning, which forms the basis of social learning theory developed by Albert Bandura (1977), requires cognitive processes which are not accounted for in behavioural theories. (Given the importance of social learning to the method developed for the study reported here, it will be discussed in more detail in chapter 6.)

Cognitive psychology, more broadly, focuses on many of the mental process and influences that behaviourism’s critics argue it fails to adequately address, such as thinking, decision-making and problem-solving. The rise of cognitive psychology should not be taken as an indication that behaviourist insights and methods have lost their significance, however. It has simply become a lot clearer that effective behaviour research requires the ‘balance of evidence’ from a wide variety of studies, and from different kinds of perspectives, to be considered in context and ‘weighed up’ (Jackson, 2005). As Jackson observed in his extensive review of evidence on consumer behaviour and behavioural change, “… it is virtually impossible to derive universal causal models with which to construct behaviour change policies” (p. 6).

Behaviourism can provide a ready explanation for somebody’s action in seeking to earn a reward or avoid a punishment (extrinsic motivation), but intrinsically motivated actions are often more problematic – sometimes a person’s internal machinations are not obvious (e.g. concepts from Festinger cited above). Take the action of donating blood, for example. Since the 1970’s it has been widely accepted that people are less likely to give blood if they receive a financial incentive to do so (Titmuss, 1970). This is explained by cognitive evaluation theory which asserts that intrinsic motivation is made up of
psychological needs for ‘autonomy’ and ‘competence’, an advance that blends elements of incentive and content theories and explains why sometimes people see rewards as controlling or threatening their autonomy and, therefore, reducing their intrinsic motivation (Deci et al., 1999). While that principle is generally accepted, recent studies have challenged the blood donation results presented in Titmuss – particularly as they relate to developing economies. A large-scale field study conducted by Iajya et al. (2012) in Argentina, for example, found a very different effect: only higher-valued economic incentives (i.e. extrinsic positive reinforcement) generated more donations. Intrinsic and social rewards had no effect whatsoever. The explanation for this was that unlike in more developed countries, the act of anonymously donating blood is not associated with being ‘pro-social’ in Argentina because blood supply is normally collected through donations by relatives and friends of individuals needing transfusions, or otherwise through services run by ‘for profit’ companies (Iajya et al., 2012).

The difference between whether an incentive motivates and engages people or demotivates and alienates them is largely determined by the social conditions in which they develop and function (Ryan & Deci, 2000; Stern, 1993). In their meta-analysis of 128 studies, Deci et al. (1999) concluded that “… tangible rewards tend to have a substantially negative effect on intrinsic motivation” (pp. 658-659). They found this is because reward contingencies weaken people's ability to self-regulate and transfer their responsibility for motivating themselves, particularly in relation to activities they find stimulating and interesting. Furthermore, there can also be a risk that the recipient will interpret the offering of a reward as conveying something negative about the nature of the task or their ability (Bénabou & Tirole, 2004). This issue is often compounded when organisations introduce greater surveillance, evaluation, and competition to aid in their assessment and handling of rewards because each of these things have been found to
consistently undermine intrinsic motivation (Deci et al., 1999). Exogenous rewards, therefore, need to be used sparingly and with caution (for more examples, see Steed, 2013).

4.5. Influence of values and ethics

Governments throughout the world seek to influence behaviour by applying incentives and punishments at every level, from the individual to the corporation and the state. And yet, from the motivational perspective, punishments have similar effects to rewards: people working to avoid a punishment, or the threat of punishment, are less intrinsically motivated than people who are not (Deci & Moller, 2005). The same holds in situations where people are acting out of a fear of failure, or the fear of being perceived as a failure (Elliot & Harackiewicz, 1996). Given that a person’s values and ethics (intrinsically) guide their motivations (Martiskainen, 2007), an important consideration for policy planning and implementation is the potential for these influences to be “crowded out” by exogenous factors. And also, noting that not all of us are consistently inclined to pro-environmental behaviour, there is also the potential for undesirable tendencies to be reinforced or “crowded in” (Rode et al., 2015). The crowding phenomenon was first documented by Frey (1992) who observed that in many cases pricing and other policy interventions can effectively “crowd out” environmental ethics and lead to unintended consequences (e.g. increasing pollution) which can even “spill-over” to unrelated areas of environmental concern (Frey, 1992). The consequence can be to undermine or preclude political action in favour of the environment, thus helping to explain why “… environmental agencies use a style of enforcement which is predominantly conciliatory and based on compromises rather than on compulsion and coercion” (Frey, 1992, pp. 409, citing Hawkins et al.).
Frey’s assessment is supported by research on public attitudes and behaviours toward the environment conducted by the British Government’s Department for Environment, Food and Rural Affairs (DEFRA) which found that “… for most people, being ‘green’ is seen as the socially acceptable norm” (DEFRA, 2007, p. 2). That research also found that many factors stand in the way of action and, drawing on additional information from a subsequent analysis, DEFRA produced a segmentation model that categorises the British population into distinct ‘clusters’ according to their attitudes and beliefs towards the environment, environmental issues and behaviours (DEFRA, 2008). The clusters are shown at Figure 4-2 plotted on a scale of ‘willingness’ and ‘ability’ to act on a set of 12 environmental behaviour goals.

![Figure 4-2. The seven population segments (DEFRA, 2008)]
The profiles for each of DEFRA’s clusters take into account a range of characteristics including ecological worldview, lifestyle, attitudes towards behaviours and current behaviours, “socio-geo-demographics”, motivations and barriers, and knowledge and engagement. The composition of clusters was found to vary across population groups and evidence was also found that people can tend to move from one category to another over time according to life stage and other individual circumstances (DEFRA, 2008, p. 8).

A wide variety of studies looking at consumer behaviours relating to the environment have found that people actively look to minimise the monetary or psychological costs associated with making changes to their behaviour and extrinsic motivators generally have a strong effect. In many cases monetary savings and methods for making energy use visible have been found to have a greater influence than environmental beliefs and attitudes (see Gärling et al., 2002; Martiskainen, 2007). However, there is little evidence that rapid results achieved through targeted intervention lead automatically to locked-in pro-environmental behaviours or to permanent shifts in people’s core values and beliefs (Stern, 2011). Gärling et al. (2002), in their review of household travel and private car use, make a compelling argument that it is, therefore, necessary to consider both internal and external factors when seeking to change people’s environmental behaviour. DEFRA’s work serves to highlight that, when considering interventions aimed at influencing behaviour, it is possible to categorise people generally, but it is also very important to adopt a targeted approach that can accommodate each individual’s values and ethics.

4.6. Influence of habits

Humans are creatures of habit and most energy consuming behaviours are habitual and based on routine. Only a small proportion are ‘one-offs’ that prompt people to pay careful
attention to available alternatives (Martiskainen, 2007). For behaviour to be ‘habit’ it must have an automatic element (in contrast to repetitive actions requiring carefully considered choices) with an intensity that can be gauged by reference to how frequently and recently, i.e. routinely, it occurs (Darnton, 2008). Numerous empirical studies, such as those conducted by Verplanken et al. (1998), have demonstrated that although external incentives are able to increase the follow-through from energy-saving ‘intention’ to energy-saving ‘action’, habits set “boundary conditions” that limit their effect. For example, it has been found that even when an intervention or method has demonstrated success in altering a person’s attitudes and behaviours, these changes are generally not maintained when they come into competition with an entrenched behaviour that is habitual (Gärling et al., 2002).

Verplanken et al. (1997) demonstrated that the habits of experimental subjects could be broken (in relation to their patterns of car use) by introducing techniques that increased focus on the information and decision options available to them; however, the effect was only temporary. They went on to conclude that habits may profoundly affect the way individuals process information in situations where they are dealing with choices (Verplanken et al., 1997), i.e. they are generally entrenched and can override all other considerations. This conclusion has been criticised by Gärling, et al. on the basis that many habits are “functional and convenient,” and it may therefore be possible to change them by manipulating the environment, for example by introducing a level of inconvenience to the habituated behaviour (Gärling et al., 2002). This implies a combination approach may be effective in some situations, such as one that first attempts to break the habit and is then followed up with the introduction of alternative choices. A good example of this was observed by Fujii et al. during an 8-day temporary freeway closure in Osaka, Japan. The closure forced drivers to consider public transport options
and, having done so, they were able to reflect on misconceptions they’d held previously (e.g. overestimating travel times by public transit) and make changes to their routines (Fujii, Gärling, & Kitamura, 2001). Measures aimed at discouraging an activity, in this case breaking habits, are generally referred to as “push measures”. “Pull measures” come from the other direction and are designed to encourage an activity and hopefully facilitate the development of pro-environmental habits. It has been suggested, therefore, that an effective method for overcoming entrenched habits is to commence with push measures and then introduce pull measures once the push measures have been successfully implemented (Fujii et al., 2001; Gärling et al., 2002).

Another method for breaking habits draws on the pioneering work of Lewin (1951) who demonstrated that an individual’s habits are greatly influenced by membership of a group. Activities and interactions in a group or community are governed by social ‘norms’—guidelines that indicate how members should behave—and we develop our understanding of these norms by observing how others are acting and then we choose to act in a way that we regard as consistent (McKenzie-Mohr, 2000). Norms, therefore, help to mediate between the identity of the individual and that of the group (Darnton, 2008). Lewin found that because a key factor that makes habits particularly hard to break is their close relationship to social norms, in order to break them and effect lasting change it is necessary to work on the broader group rather than just the individual.

**4.7. Triandis' Theory of Interpersonal Behaviour**

The segmentation model developed by DEFRA acknowledges the fact that some individuals are more receptive to initiatives aimed at encouraging pro-environmental behaviours than others. And it has been established that these intrinsically motivated people quickly develop new behaviours and exhibit larger savings in response to
interventions aimed at reducing their energy use, even though most will revert to their old habituated behaviour patterns over time (van Dam et al., 2010). Personalities, values and ethics clearly matter for behaviour change, and so do habits because they impact at every phase in the process. Therefore, in order to effectively influence non-residential building FM’s patterns of energy-use behaviour, it is helpful to draw on a social-psychological model that takes these factors into account.

The objective of identifying a social-psychological model, in this study, is not to establish a rigid framework for analysis but rather to assist with thinking about the intervention design and approaches, and aid in the assessment of results. As Cronbach (1975) put it:

“Though enduring systematic theories about man in society are not likely to be achieved, systematic inquiry can realistically hope to make two contributions. One reasonable aspiration is to assess local events accurately, to improve short-run control…. The other reasonable aspiration is to develop explanatory concepts that will help people use their heads.”

(p. 126; as cited in Entwistle, 2012, p. 7)

Unfortunately many of the more prominent and widely utilised social-psychological models are too limited to be used in relation to energy-use behaviours because they do not account for habits or routine (Martiskainen, 2007). The most prominent model—one which underpins much of the government policy in this area due to its place in neoclassical economics—is rational choice theory, also known as choice theory, rational action theory and the rational actor model (G. S. Becker, 1978; Blume & Easley, 2008; Sen, 2008; Stern, 1986). This very limited theory of human behaviour assumes that people make choices based solely on economic utility, i.e. as if balancing costs against benefits to achieve maximum personal advantage – what Milton Friedman called the “maximization-of-returns hypothesis” (Friedman, 1953). The argument for its use was
very simple: “… unless the behaviour of businessmen in some way or other approximated behaviour consistent with the maximisation of returns, it seems unlikely that they would remain in business for long” (p. 22). While this may be true (and noting, of course, that many businessmen do not remain in business for long), the model’s failure to account for the influence of factors such as habits, emotions, social norms, morals and intellect (Jackson, 2005) may go some way to explaining why information-only campaigns are generally unsuccessful (Gardner & Stern, 2002; Malone et al., 2011; Martiskainen, 2007) and why behaviour-based strategies still represent untapped potential.

As noted by another highly influential winner of the Nobel Prize in economics, Daniel Kahneman, “The agent of economic theory is rational, selfish and his tastes do not change” (p. 162). Real people, on the other hand, are rather more complex (Kahneman, 2003). Given everything discussed so far, something more sophisticated than rational choice theory is required in order to capture the key influencers of energy consuming behaviours. Of all the established social-psychological models that offer insights into human behaviour as it relates to environmental decision-making (for extensive reviews see Darnton, 2008; Jackson, 2005; Martiskainen, 2007), the one that is most helpful for the present study is Triandis’ (1977) Theory of Interpersonal Behaviour (TIB) which is illustrated in Figure 4-3 below.
Triandis’ model takes into account the influence of personal attitudinal variables, contextual factors and the influence of habits, and while it does not offer the most complex account of human behaviour, this is also an advantage because generally the more complex a model is, the less able it is to be used in experimental research (Martiskainen, 2007). Accordingly, the TIB has not been used as widely as some of the simpler models that fail to take account of habits (including rational choice theory) but it has been used successfully in empirical research, for example in defining the importance of morals and habits in determining the patterns of students’ car use (e.g., Bamberg & Schmidt, 2003).

The TIB can be described as a ‘dual process model’ because it shows how intention and habit contribute independently to the behavioural outcome. Facilitating conditions, the other key factor determining the probability of action (behaviour), can be thought of as the environment or context external to the individual. According to the TIB, the role of deliberate intention declines over time as experience increases and the influence of habit
takes over (Triandis, 1977): an idea that corresponds with the notion that with experience comes familiarity, strong habits and a reluctance to change. Bamberg & Schmidt found the TIB to be a better predictor of pro-environmental outcomes than Ajzen’s (1991) more recent and more widely used Theory of Planned Behaviour (TPB)3 because of this emphasis on habit (Bamberg & Schmidt, 2003). Accordingly, the TIB is of increasing interest to researchers and policy makers wanting to explore energy consuming behaviours, in particular those interested “… in exploring the influence of habitualization on everyday behaviours” (Bamberg & Schmidt, 2003, p. 269; see also Darnton, 2008; Martiskainen, 2007)

4.8. Stages of behavioural change

As Geller (2002) notes, “Before an environment-destructive habit (or routine) can be changed to an environment-protective habit, the target behaviour must become self-directed” (p. 532). This calls for a staged approach, requiring that individuals first become aware of their undesirable habits before they can be motivated to improve, and then, ultimately shift to new self-directed behaviours that might in time become automatic (Geller, 2002). Geller identifies four stages of competence4 (unconscious incompetence, conscious incompetence, conscious competence, and unconscious competence) and four intervention approaches (instructional intervention, motivational intervention, supportive intervention, and self-management), all of which are depicted in Figure 4-4, below. Under this model, the pathway from environment-destructive habits (top left of figure) to environment-protective habits (top right) requires three transitions between the four phases, each with a distinct set of challenges. Put broadly for the purposes of this study, the first involves awareness raising, the second requires the deployment of motivational

---

3 For an explanation of the TPB, see (Ajzen, 1991)
4 Derived from unpublished leadership training material originally developed by psychologist Noel Burch during the 1970’s while working for Gordon Training International.
techniques, and the third involves the provision of appropriate support. These all come together to one broad objective, as summarised by Geller: “The critical challenge is to help people get so personally committed to environmental protection that they would use self-management techniques to increase their pro-environment behaviour” (Geller, 2002, p. 535).

Figure 4-4. The flow of behavior change model (Geller, 2002)

4.9. Limitations of behaviour-based approaches

And yet there are significant questions about the effectiveness of behaviour-based approaches in practice, mostly arising from uncertainty about the ‘plasticity’, or ‘stickability’ of changes, i.e. can people be relied upon to maintain energy saving behaviours over the long-term? Because of this uncertainty, behaviours that lead to
adoption of energy-efficient equipment are regarded as having greater impact than behaviours focussed on changes in the use of equipment (Stern, 2011). This is the case even where the impact of operating decisions may be more immediate and dramatic. Consider an ornamental light fitting, for example. A person who makes the conscious decision to use the light less regularly may save more energy than someone who decides to replace the lamp with a low-wattage alternative. But if the resolution of both individuals wains, the low-wattage lamp will continue to use less energy in operation whereas the energy used by the unaltered fitting may revert to pre-intervention levels. Stern (2011) has likened the difficulty of changing environmental behaviours to the challenges faced in health-related campaigns such as those involving diet and smoking (see, for example, Snyder et al., 2004) and presents the following equation to express the likely impact of behavioural change interventions:

\[
I = t \cdot p \cdot n
\]

(1)

Where:

(I) is the impact of a behavioural change on climate change
(t) is the technical potential of the behaviour to alter emissions
(p) is the plasticity of the behaviour
(n) is the number of individuals who might change their behaviour

(Stern, 2011)

Plasticity is the opposite of elasticity. If the plasticity of a behaviour change is low, it is likely to revert (or spring back) to its earlier form. If plasticity is high, the change is likely to be more permanent. Evidence from residential contexts shows the levels of plasticity demonstrated for equipment “use” behaviours are consistently lower than those for equipment “adoption” behaviours (Dwyer et al., 1993; Stern, 2011). Unfortunately the literature relating to non-residential buildings is far more limited, however it can be
concluded that regardless of whether our energy behaviours are based on continuous or one-off efforts, the actions of individuals are influenced by a complex array of factors.
5. FEEDBACK

“The acquisition of skills requires a regular environment, an adequate opportunity to practice, and rapid and unequivocal feedback about the correctness of thoughts and actions.”

(Kahneman, 2011, p. 416)

5.1. What is feedback?

Feedback is a word and a concept that holds many different meanings and its use has evolved considerably over the past century. Given that this study explores the nexus between technology and behaviour through the use of feedback, a short review of its etymology is required.

The first recorded use of the verb phrase "to feed back" arose in various U.S. patents during the mid-1860s where it described how components of mechanical processes would return to their earlier positions (see, for example, Cole, 1866; Jay, 1865). This idea of components of a system extending and reversing, according to Mayr (1989), had started to enter economic theory in Britain as early as 1720: “Trade causes a Vibration, or continual Ebbing and Flowing; which may be called the natural Ballance [sic] of Trade” ('The System or Theory of the Trade of the World' by Isaac Gervaise cited in Mayr, 1989, p. 165). However, it was not until 1920 that the Oxford English Dictionary first included a definition of “feed-back”; a noun specifically related to electronic amplifiers: “An inductive feed-back in relation to the secondary system generates local oscillations” (OED, 2014). The earliest recorded use of the word in that context was by Karl Ferdinand Braun who used it in his lecture at the ceremony for the awarding of the 1909 Nobel Prize
in Physics (which he shared with Guglielmo Marconi in recognition of their contributions to the development of wireless telegraphy) to refer to (undesired) coupling between components of an electronic circuit (Braun, 1909).

It wasn’t until the mid-twentieth century that the term “feed back” began to emerge as the verb we use in common language today. In 1943 the OED recorded its use as an abstraction relating to “transfer” of information, citing an article from the journal Philosophy of Science: “Purposeful active behavior may be subdivided into two classes: ‘feed-back’ (or ‘teleological’) and ‘non-feed-back’ (or ‘non-teleological’)” (OED, 2014). It has since evolved further to the form used in the present study, as a noun that is given or transferred:

*The modification, adjustment, or control of a process or system (as a social situation or a biological mechanism) by a result or effect of the process, esp. by a difference between a desired and an actual result; information about the result of a process, experiment, etc.; a response.*

(Oxford English Dictionary, 2014)

The first clear attempt at a definition of “feedback” in management theory was by Ramaprasad (1983) who identified that the definition used by management theorists was quite different to those used in general systems theory, control theory (cybernetics), etc. Ramaprasad asserted that without a common understanding of the concept, knowledge from related disciplines could not be translated for use in the context of management. His article ‘On the Definition of Feedback’ proposed the following definition of feedback which is generally accepted in management theory today:
Feedback is information about the gap between the actual level and the reference level of a system parameter which is used to alter the gap in some way.

(Ramaprasad, 1983, p. 4)

If information is merely stored, or if it is passed on without including qualities necessary for interpretation, or if it is directed to people who lack the knowledge or power to use it, then the control loop cannot be closed and it becomes what Sadler has termed “dangling data” (Sadler, 1989). This view accords with Ramaprasad who emphasises three crucial points about feedback:

1. The focus of feedback may be any system parameter: input, process, or output.

2. The necessary conditions for feedback are the existence of data on the reference level of the parameter, data on the actual level of the parameter, and a mechanism for comparing the two to generate information about the gap between the two levels. There cannot be any feedback if any one of the three (data on the reference level, data on the actual level, mechanism for comparing) is absent.

3. The information on the gap between the actual level and the reference level is feedback only when it is used to alter the gap. If the information is stored in memory it is not feedback.

(Ramaprasad, 1983, p. 5)

An important aspect of this definition, which will be considered in more detail in the discussion in Part D, is the focus on the closing or narrowing of the “gap” between an actual or observed performance level and the reference level. Feedback that minimises deviations from a goal or reference level is termed ‘negative feedback’ (Gärling et al., 2002), whereas ‘positive feedback’ works to increase them. If feedback methods are to help in facilitating the built environment’s long-term progression from current levels of
energy efficiency to the levels proposed by Lucon et al. (2014), it may be necessary to define a target standard (reference level) for each individual building. Given the multiplicity of factors that contribute to a building’s energy use, both physical and behavioural, this is not likely to be achievable in many cases. Over the short-term, however, for many buildings it may be sufficient to establish baseline reference levels and inform the operators about divergences above or below them. In this scenario, if energy use is higher than baseline the operator receives feedback, but if energy use falls below baseline the FM receives no feedback because this information is not used to close the gap (otherwise performance would revert).

Being able to reduce the gap requires that FMs: (1) have a concept of the baseline standard of performance which aligns with the evaluation being presented to them; (2) reflect, self-evaluate and continuously compare the actual level of performance with the baseline; and (3) engage in appropriate actions in an effort to close the gap, which means that FMs (or their technicians) have appropriate knowledge and awareness of suitable strategies to draw upon (Sadler, 1989, 1998, 2010; Van der Schaaf et al., 2013). This obviously introduces a few methodological questions for experimentation in the field which will be addressed in Part B.

**5.2. Elements of highly effective feedback and assessment**

Much of the research on feedback has its roots in Thorndike’s (1905) 'law of effect' (Sadler, 1989), in particular the idea that, if an action is closely followed by ‘satisfaction’ it is more likely to recur under the same or similar circumstances in future (Thorndike, 1905, 1911). Exploration of this concept by Ammons (1956), Bilodeau and Bilodeau (1958), et al., led to the advance known as knowledge of results (KR): augmented (as opposed to intrinsic) feedback typically provided as a verbal evaluation at the completion
of a task. KR is additional to the sources of feedback that are naturally received through cause-effect and can help to generate positive motivational beliefs and self-esteem (Nicol & Macfarlane-Dick, 2006; Sadler, 1998), an aspect that aligns more closely with Thorndike’s assessment than it does with Skinner who eschewed reference to mental states in his behaviour experiments (Skinner, 1950).

A related concept, also involving augmented feedback, is knowledge of performance (KP), a term focused on informing the performer about the way the performance or task was carried out. The distinction between KR and KP can be quite subtle but is well illustrated in this example from Salmoni:

“In learning a springboard dive … it may be that the principles of learning are the same for KR (“Your score was 4.5”) and KP (“You tucked too late”), but each could cause the performer to learn something different about the dive, or one could be more efficient than the other in a particular circumstance.”

(Salmoni et al., 1984, pp. 356-357)

For such information to become feedback it must be specific to an individual’s situation and it must be timely enough to facilitate learning (Stern, 2011), as would be the case in the above example if KR and KP were delivered while the memory of the dive was still fresh in the mind of the diver. Delivered in an educational setting or workplace with the aim of forming or moulding behaviour, it becomes formative assessment, where “… judgements about the quality of student responses (performances, pieces or works) [are] used to shape and improve the student’s competence by short-circuiting the randomness and inefficiency of trial-and-error learning” (Sadler, 1989, p. 120). Formative assessment techniques have become widely used in education where research is increasingly focussed on methods of designing feedback to give learners a progressive appreciation of what constitutes high quality work and the techniques required to attain progressively higher
standards of achievement (Sadler, 1998). The reasons for the increasing popularity of qualitative judgements in assessment are clear from the meta analyses (see, for example, Hattie & Timperley, 2007) which show consistently that the most effective forms of feedback are those that provide “cues or reinforcement to learners; are in the form of video-, audio-, or computer-assisted instructional feedback; and/or relate to goals” (p. 84). Just as importantly, it has been found that programmed instruction, praise, punishment, and extrinsic rewards are generally ineffective (and in some cases detrimental) to learning in educational settings (Hattie & Timperley, 2007). These observations tie back to many of the principles of behaviour change discussed in the previous chapter, but also to the ultimate goal of many instructional systems: to facilitate the transition from feedback to self-monitoring and self-regulation (Sadler, 1989).

Caution must be exercised in prescribing a universal ideal that feedback should conform to because of the complex and variable relationship that exists between its form, timing and effectiveness (Sadler, 2010). Furthermore, the present study is concerned with the FMs of large commercial office buildings and there is no available literature evaluating the efficacy of energy performance feedback methods in that context. This makes it necessary to try to abstract common threads from the extensive work conducted in other fields, in particular higher education settings (because they relate to adult learning) and from studies of residential energy use (because they relate to energy behaviours).

A synthesis of the literature relating to feedback in higher education by Nicol and Macfarlane-Dick (2006) produced the following seven principles widely regarded as good practice for strengthening a student’s capacity to self-regulate performance. They found that good feedback practice:
1. helps clarify what good performance is (goals, criteria, expected standards);

2. facilitates the development of self-assessment (reflection) in learning;

3. delivers high quality information to students about their learning;

4. encourages teacher and peer dialogue around learning;

5. encourages positive motivational beliefs and self-esteem;

6. provides opportunities to close the gap between current and desired performance;

7. provides information to teachers that can be used to help shape teaching.

(Nicol & Macfarlane-Dick, 2006, p. 205)

Prescriptions for effective feedback to household energy consumers are also emerging in the literature (Darby, 2010; DECC, 2009; Fischer, 2008). Fischer, for example, suggests that:

“Successful feedback has to capture the consumer’s attention, to link specific actions to their effects and to activate various motives. If this is the case, then different characteristics of the feedback itself become relevant, among them, its frequency, content, breakdown, presentation, inclusion of comparisons, and combination with additional information and other instruments.”

(Fischer, 2008, p. 85)

Fischer found that the frequency of feedback is crucial in the residential context, finding that “… none of the ‘less than monthly’, and all but one of the ‘daily or more’ projects are among the best performing (as far as they can be compared)” (Fischer, 2008, p. 97). Furthermore, information should be ‘delivered’ in a readily accessible form rather than simply being made ‘available’. Martiskainen, whose review of the literature on household energy consuming behaviours found that feedback on energy use has the “greatest
potential” of all behaviour-based techniques to reduce consumption (Martiskainen, 2007, p. 6) suggests that highly effective feedback:

- consists of simple messages clearly presented
- contains information directly relevant to the recipient
- involves some kind of a goal or a commitment
- will be visible, consistent and frequent.

One of the most influential articles underpinning much of the recent research on the nature and function of formative assessment in the development of expertise is Sadler (1989) (see Nicol & Macfarlane-Dick, 2006). Sadler, whose work focusses on higher education, argues that if students are to improve they must come to share the same concept of ‘quality work’ as that held by the teacher, be able to continuously monitor and objectively evaluate the quality of their work with reference to this higher standard, and be able to draw upon a range of alternative approaches to regulate or modify their work if required (Sadler, 1989). To directly improve learning, and not simply motivate learning, feedback must “help the recipient to reject erroneous hypotheses” (Kluger & DeNisi, 1996, p. 265). Essentially, this requires that the student should assimilate the teacher’s evaluation capabilities. Sadler argues that these skills can be developed and that “…providing direct and authentic evaluative experience is a necessary (instrumental) condition for the development of evaluative expertise and therefore for intelligent self-monitoring” (p. 143).

The above review suggests some crucial ingredients for feedback in the context of helping FMs develop the evaluative skills necessary for continuous energy efficiency
improvement. These have been condensed into the following four criteria that will underpin the model for the action research methodology adopted in this study.

To be effective, energy performance feedback to operators of non-residential buildings must:

1. consist of simple, clear, reliable, high quality messages delivered consistently and frequently (with limited delay);

2. make reference to some kind of a goal, commitment or expected baseline standard;

3. contain information that is directly relevant to the recipient in order to promote self-assessment (reflective thinking) and investigation;

4. build positive motivational beliefs and self-esteem that promote experimentation and support constructive dialogue with experts, specialists and peers.

5.3. Important contextual factors

There are also a few preconditions and environmental factors that must be carefully addressed for FMs to effectively apply feedback. One emphasised by Sadler (2010) is the need for critical background knowledge in order for them to be able to identify the aspects of their work that need attention. Without a reasonable understanding of building systems and operations (i.e. a level of knowledge to be expected in someone employed for the role), even the clearest performance feedback will leave the recipient feeling confused. Associated with this is the issue of “validity”. A recipient must perceive the feedback as valid (regardless of whether it is or not) in order to act upon it. In this regard, computerised feedback is generally considered more neutral, is more trusted, and leads to
stronger feelings of self-efficacy (task-specific confidence) and to better performance compared with identical feedback from a supervisor (Kluger & DeNisi, 1996; McCalley & Midden, 2002). Reliability (as noted at criterion 1, above) is a necessary condition for valid feedback, however it is not sufficient. To be valid it must be “on target” and well-based from the point of view of the recipient (Sadler, 1989). This also applies to the goal or expected standard (criterion 2) which must also be regarded as valid: if the recipient does not understand the goal or does not assume some level of ‘ownership’ over it or does not believe that it can reasonably be achieved, then it is unlikely that it will be regarded as a valid basis for measuring progress (Black & Wiliam, 1998; Nicol & Macfarlane-Dick, 2006).

Whereas criteria 1 and 2 address the largely mechanistic and measurable dimensions of effective feedback design, an approach to criteria 3 and 4 will likely need to be more situational and nuanced (e.g. taking into account personalities, workplace dynamics, etc.). This presents an important co-benefit because the degree to which the feedback stimulates reflective thinking, investigation and dialogue, can also indicate its overall effectiveness (van der Schaaf et al., 2013). A clear message must be understood and internalised by the recipient before it can inform a productive investigation that draws on available resources and the assistance of peers and expert support (Nicol & Macfarlane-Dick, 2006). The supportiveness of the environment is crucial in this regard. A review of the literature by Ames (1992) found that, among the other factors mentioned above, self-perception is critical: feedback is most effective when kept private and when it encourages the view that mistakes are a part of learning (Ames, 1992). Such an environment provides an opportunity to focus on strengths and weaknesses and offer praise alongside constructive criticism – all without undermining motivation and self-esteem (Black & Wiliam, 1998; Nicol & Macfarlane-Dick, 2006).
Associated with the issue of validity is that of trust. Feedback may be perceived as reliable and well-based (e.g. generated by a computer) and yet, due to concerns over the integrity or credibility of its source, may still not be trusted by its recipient (Greenberg, 2014; Sovacool, 2014b). This can be a particular concern when the source of information is regarded as conflicted, as is often the case when retailers of electricity offer efficiency advice (Martiskainen, 2007; Stern, 1993). A specific example of this problem was identified in research conducted by Craig and McCann during the 1970s energy crises where it was found that information from utilities tended to be ignored by household consumers even though it was of high quality and would have saved the recipients money (Craig & McCann, 1978). It is highly probable that a similar issue may arise in the context of operating commercial buildings if FMs perceive ‘management’ as having priorities that compete or conflict with their own (e.g. “management is trying to save money by making FMs work harder”, or “management is looking for an excuse to lay off staff”). Regardless of whether such concerns are well-founded or otherwise, recipients will inevitably speculate on the motives of the information source, so it is important that they feel those motivations align with their own. A good model for commercial buildings would therefore credibly demonstrate the independence of the feedback source and draw upon the recipients’ peers as a source of additional information and interpretation. It would also involve a significant component of interpersonal contact rather than relying purely on technology or printed media, given that existing professional and social networks tend to be among the most highly trusted sources of advice and information (Stern & Aronson, 1984)

5.4. Insights from Residential feedback – HEMS

As technologies advance, the options for providing direct and automatic energy feedback expand. This is particularly true for the residential sector where various automatic forms
of energy usage feedback have been evolving for approximately four decades, thus allowing for rigorous analysis and evaluation (see, Darby, 2010; Delmas et al., 2013; Ehrhardt-Martinez et al., 2010; Faruqui et al., 2010; Fischer, 2008). A key of element Advanced Metering Infrastructure (AMI)—an umbrella term for the hardware, software, communications systems, consumer displays, controllers and associated systems—is the “smart meter”. Smart meters are utility meters that record and communicate electronically at intervals of less than an hour, rather than relying on a person to physically read them (Darby, 2010). The pace of smart meter rollouts has been accelerating recently. For example, 37 million smart meters were installed in the U.S. between the beginning of 2010 and July 2014 and they now cover more than 43% of all households in that country (IEI, 2014). Extensive rollouts are also taking place elsewhere throughout the world, including in Australia where in 2006 the COAG committed to “the progressive national roll out of 'smart' electricity meters from 2007 to allow the introduction of time of day pricing and to allow users to better manage their demand for peak power” (COAG, 2006).

Households (and also small businesses) are the target of these initiatives in virtually every instance globally because, as U.S. President Barrack Obama stated at the launch of the USD $3.4 billion smart meter investment program in 2009: “Smart meters will allow you to actually monitor how much energy your family is using by the month, by the week, by the day, or even by the hour, so coupled with other technologies, this is going to help you manage your electricity use and your budget at the same time” (Mooney, 2015). Evidence is emerging that this may be the case. For example, a study conducted across 437 Connecticut households (207 in the control group) in 2011 found that consumers with smart meters connected to IHDs that received time-of-use pricing signals reduced their consumption by 14 percent on average when prompted about a price increase. Consumers without the displays who received time-of-use pricing signals saved between zero and
seven percent (Jessoe & Rapson, 2014). It must be noted that this study had a very limited duration and concentrated on demand shifting rather than influencing or developing long-term energy efficiency behaviours. Nonetheless, the findings are significant from an economics point of view because they demonstrate that a key factor limiting ‘rational’ energy usage behaviour (i.e. application of rational choice theory) is imperfect information: the price elasticity of residential customers’ demand increases with the provision of feedback on how much they are using. While this finding is important and encouraging, the degree to which it can be transferred to non-residential building contexts is uncertain given it is almost invariably the case that FMs will not directly bear the costs / benefits of their energy management behaviours.

Further advances in AMI, including disaggregated feedback down to the appliance level, offer much promise for significant energy efficiency gains (Armel et al., 2013; Delmas et al., 2013; Stern, 2011), but they need to be used. Currently only one percent of Americans with a smart meter receive feedback via an IHD connected to it (Mooney, 2015) indicating that the benefits of this hi-tech approach are far from being fully realised and there is still much to learn about the potential benefits and limitations. This was emphasised by the British government in its response to a national consultation about the rollout of smart meters, citing evidence from Sweden and the USA where it has been found that only 2 to 4 percent of customers choose to view smart meter data online (DECC, 2009). According to that report:

“The Government’s position remains that a standalone display should be provided with the smart meter ... the provision of a display is important to securing the consumer benefits of smart metering, delivering real time information to consumers on their energy consumption in a readily accessible form.”

(DECC, 2009, p. 31)
This judgement has recently been confirmed by the DECC’s commitment that every single gas and electricity customer across Britain will be offered an in-home display as part of the forthcoming rollout of 53 million smart meters from 2015 (Mooney, 2015).

But will householders be motivated to look at the displays and use the information to tune their energy use behaviours to achieve greater efficiency? And what should the displays show? Contrary to the findings of Jessoe and Rapson cited above, Delmas et al. (2013), found in their near-comprehensive meta-analysis of residential energy information studies that pecuniary feedback (i.e. relating to price or cost) actually leads to a relative increase in energy usage rather than inducing conservation – a phenomenon sometimes associated with ‘rational inattentiveness’, e.g. where effort and inconvenience may lead to a 5 percent saving that adds only $5 per month to the hip pocket (Delmas et al., 2013). A key finding of Delmas et al. is that information delivered in person appears to be consistently more effective than information provided through other media. Much of the above discussion about human motivation and feedback design may help to explain this finding and, again, it is important to highlight the risks of transferring knowledge from the residential sector to non-residential buildings without carefully considering all the contextual factors. As noted by Stern (2011), the introduction of new technologies designed to influence energy consuming behaviours calls for the practical application of psychological knowledge; a point emphasised by the DECC in relation to smart meter technologies for non-residential buildings:

“In recognition of the different needs of the wide range of customers and premises in the non-domestic sector, the Government does not intend to require a real-time display device to be provided to electricity or gas consumers in this sector. As part of further preparatory work on the roll-out of smart meters, we will consider what data should be made available to
Potentially the most promising answer for commercial office buildings is a hybrid approach. Malone et al. (2011), citing interview research and not empirical data, suggest that feedback is most effective when matched to the recipient’s area of responsibility, e.g. fine-grained for building operators and less frequent and more summative for their managers and those charged with capital purchasing decisions. They also suggest feedback should provide information about dollars and units of energy because these are used for different purposes and assist analysis (Malone et al., 2011).

All of this serves to highlight the efficiency improvements that may be achieved in non-residential buildings by introducing rigorous methodologies that draw on decades’ of residential field studies focused on consumer interactions with energy monitoring and feedback systems. Despite uncertainties about the transferability of techniques, plus the obvious differences in the form and function of residential and non-residential buildings, there are important similarities. In both contexts energy is used to provide thermal comfort and indoor environment quality. Likewise, the level of operator awareness and competency (and other behaviour-based factors) can often explain significant variances in baseline energy use between otherwise comparable buildings – factors that may be influenced with the provision of effective feedback.

**5.5. Problems with feedback loops in buildings**

Sadler (1989) makes an important observation: “Few physical, intellectual or social skills can be acquired satisfactorily simply through being told about them. Most require practice in a supportive environment which incorporates feedback loops” (Sadler, 1989, p. 120).
The term ‘feedback loop’ is used widely and, like ‘feedback’ itself, it can take on very specific meanings depending on the field of application. Generally, however, drawing on the work of Ashby (1957) in the field of cybernetics (or control theory) and of Bandura (1977) in the psychology of behaviour change, it involves four distinct stages as illustrated in Figure 5-1.

Figure 5-1. A typical feedback loop.

In an effective feedback loop, evidence about the effect of an action is collected as raw data and stored. It is then communicated after being converted to a form that makes it relevant and able to be interpreted by its intended recipient. The information must then be absorbed by the recipient as knowledge about potential consequences and alternatives. And then, finally, the recipient makes a choice and takes action once more in a process of repetition and recalibration that edges towards optimum performance.

The development of expertise depends not only on the quality and speed of feedback, but also on there being sufficient opportunity to practice (Kahneman, 2011). This is where the concept of a loop is particularly important for buildings. As noted by Brown and Arens
(2012) it is commonly the case that feedback loops in buildings are “broken”, i.e. data does not get converted to information in a usable form or directed to the appropriate recipient in a timely manner. This means that FMs (and designers, technicians, occupants, etc.) have limited ability to see the effects of their decisions and are therefore not able to effectively develop and refine their skills. Building management and control systems (BMCS), for example, emphasise monitoring for automatic control rather than for the provision of feedback to operators (Brown & Arens, 2012). This creates a dependence on the technology and disempowers the operators who must rely on the controls systems because they are ‘better informed’ than they are. Furthermore, if they interfere with the BMCS’s operation and settings they run the risk of introducing consequences they may not fully understand or even be aware of.

Closing these feedback loops requires cultural change. However, as Brown and Arens note, “Research interventions to close the feedback loops sometimes enjoy transient success, but often run up against strong cultural barriers against using certain types of information or communicating with other groups in the building social network” (Brown & Arens, 2012, pp. 12-68). In some cases there are vested interests to consider (empowered FMs may be less dependent on advice, technology and support from paid service providers who control the data) and more generally there is the issue of inertia to overcome. This requires an additional stage: initiation, i.e. “the framing of the problem that is to be solved” – a point highlighted by Kahneman (2011, p. 418) and addressed in the methodology adopted for this study by the application of social learning theory introduced in chapter 6.
5.6. The boomerang effect

The term ‘boomerang effect’ is used in social psychology to describe a situation where an effort to persuade someone of something leads unintentionally to them adopting an opposing position. Various studies (see for example, Bittle et al., 1979-80; Brandon & Lewis, 1999) reveal boomerang effects arising from the provision of normative energy usage feedback to households where their consumption is compared to broader populations (social norms) or previous periods: feedback in some cases prompts high users to conserve energy but low users to use more. In the study by Brandon & Lewis, which involved the provision of various forms of normative feedback on the energy consumption of 120 British households, it was found that following the feedback high and medium users reduced their usage by an average of 3.6% and 2.4% respectively, whereas low users increased their consumption by an average 10.7% (Brandon & Lewis, 1999). Similar effects have been observed in studies targeting gambling, eating disorders, alcohol consumption, drug use, littering, and recycling (see Schultz et al., 2007, for examples and discussion).

Normative feedback (i.e. feedback that makes reference to past performance or the performance of others) intended to increase the prevalence of prosocial behaviour (e.g. energy saving) relies on the recipients forming a view that positive behaviours are common amongst the broader population. The problem arises for recipients who gain an awareness that others are not doing as well as they are – because most people prefer not to deviate from the community standard (Schultz et al., 2007). A related issue has been observed in regard to political ideology and highlights some of the challenges in predicting people’s reaction to normative information about their communities’ energy use. One of the explanations for the earlier cited finding (chapter 3) that energy feedback conveyed to liberals is more effective than it is with conservatives (Costa & Kahn, 2013).
is that normative feedback of the type provided affects people differently – liberals may regard energy saving as a virtue and desire to be ranked highly. Costa and Kahn’s study was based on data from households in Sacramento, California who received home energy reports provided by Opower, a company that works with energy utilities to provide customers with normative feedback. While confirming the results, Opower subsequently published data from other utility regions in the U.S. showing political affiliation exhibits little (or a reverse) relationship to energy saving in response to feedback (Curtis, 2013).

The Opower data, despite inconsistencies in results (possibly) associated with political affiliation, do appear to support the company’s assertion that “behavioural energy efficiency is a universal resource” (Curtis, 2013). One reason it is able to make such a claim is a feature of its feedback design that was first identified in the study by Schultz et al. (2007). Schultz et al. (a team that included Robert Cialdini, currently the world’s most cited living social psychologist and chief scientist at Opower) recognised that social norms can be divided into two types: descriptive norms and injunctive norms. Descriptive norms refer to perceptions about what is common practice or a common level of performance. Injunctive norms, on the other hand, refer to perceptions about what is appropriate or socially desirable within the community (Reno et al., 1993). Studies have shown that the type of norm that is most prominent in a person’s consciousness will exert the greater influence over behaviour (Cialdini & Goldstein, 2004). The study by Schultz et al. aimed to test whether an undesirable boomerang effect induced by descriptive normative information could be avoided by adding an injunctive message. 290 households with visible energy meters from a small Californian city were randomly selected to participate in the study. At each observation period the households were divided according to whether their energy use was below or above average for the community. Participants in each group were given one of two kinds of feedback:
descriptive normative information comparing their energy use to the average, or that plus a simple injunctive message. The injunctive message was either a positively valenced emoticon (😊) for those that used less than average or a negatively valenced emoticon (😊) for those that consumed more. The boomerang effect, observed (as predicted) with those households that received only the descriptive normative information, was eliminated by the addition of the 😎 (Schultz et al., 2007). Similar results have been observed in work environments (Dixon, et al., 2015b).

5.7. Feedback without goals ineffective?

A basic requirement that feedback it be used to ‘alter the gap’ between an actual level of performance and a reference level (Ramaprasad, 1983) is common to several prominent theories of behaviour regulation, including control theory and goal-setting theory. However, whereas control theory (as applied to behaviour) postulates that people will be motivated to reduce a discrepancy between the performance level and the reference (Carver & Scheier, 1982), goal-setting theory suggests people are more motivated by the desire to achieve a goal than they are by eliminating the discrepancy (Locke & Latham, 2002). Goal-setting can elicit a variety of behaviours, e.g. people may strive to attain the goal, seek to amend it, dismiss or reject feedback, or abandon their commitment to the goal. The typical response, though, is for people to work towards eliminating the discrepancy between feedback and the reference (i.e. the goal) by applying more effort if performance falls short, or reducing (or maintaining) effort if performance exceeds the standard, as illustrated in Figure 5-2 below (Kluger & DeNisi, 1996).
Figure 5.2. The effects of feedback intervention (FI)-induced attention on task-motivation processes and their consequences for performance (Kluger & DeNisi, 1996)

Despite most studies reporting that immediate or timely energy feedback prompts recipients to change behaviour to save energy, some researchers challenge such findings. They point to an apparent tendency for many studies to overlook “a considerable body of evidence suggesting that feedback intervention effects on performance are quite variable” (Kluger & DeNisi, 1996, p. 254), (see also, Delmas et al., 2013; McCalley & Midden, 2002). According to McCalley and Midden (2002), one assumption implicit in most residential energy feedback analyses and yet often left unquestioned is that “most residents already have a goal to save energy (or money)” (p. 591). This, they argue, is a fundamental flaw because residents may not have this goal or, even if they do, may give it a lower priority than maintaining comfort or convenience, for example.

A clear understanding of the relationship between goals and feedback is necessary in order to achieve optimal feedback effectiveness (McCalley & Midden, 2002). Goals
mobilise and direct effort towards goal attainment (Gärling et al., 2002) and they affect performance through four mechanisms (Locke & Latham, 2002):

1. **choice / direction** – they direct attention and effort from goal-irrelevant activities towards goal-relevant activities

2. **effort** – they energise, with the most challenging goals channelling the most effort

3. **persistence** – challenging goals prolong effort, although there is typically a trade-off between time and intensity, i.e. people can maintain high intensity effort for short periods, or persistent effort for long periods at lower intensity. Tight deadlines increase the pace of work

4. **strategies** – they arouse thought processes leading to the discovery of task-relevant knowledge and strategies.

A distinction should be made between a performance goal and a learning goal. If a goal to achieve a certain level of performance is beyond a person’s present capabilities it can lead to evaluative pressure and performance anxiety that manifests itself in unsystematic effort and a failure to learn and improve. This can be avoided by focussing on learning goals, for example goals focussed on the identification, evaluation and introduction of performance improvement strategies rather than the achievement of results (Locke & Latham, 2002). High performance goals work best when people have appropriate training and background information (priming) about effective strategies whereas learning goals work better where they do not (Earley & Perry, 1987).

The effects of goal setting are very reliable, so long as appropriate methods are used to set them. Locke and Latham (2002) identify a variety of factors that have been associated with failure to achieve goals, including failure to secure goal commitment, not taking into
account a person’s ‘personal’ goals, not including a graduated range of goal difficulty levels, not providing feedback and not conveying sufficient knowledge for complex tasks and situations. Each of these elements can be addressed quite effectively in a rigorous goal-setting and feedback design. Kluger and DeNisi (1996) argue that, because attention is limited, only feedback-standard gaps that receive attention are addressed. Goal-setting can, therefore, be a highly effective process for directing focus to an area of performance that may otherwise not be considered, e.g. energy saving. This was illustrated in an experiment by McCalley and Midden (2002) that used an advanced washing machine control panel to provide energy efficiency feedback to 100 participants. It was found that those who were aware of a goal produced energy savings of 21% by applying feedback information whereas no energy savings were achieved by those without a goal. Overall it made no difference whether a goal was self-set or assigned, however, individuals identified as “pro-self” saved more energy when allowed to self-set a goal and “pro-social” individuals saved more when assigned a goal (“pro-social individuals are more willing to accept costs to themselves in favour of societal gains than pro-self individuals”) (McCalley & Midden, 2002). Goal setting can help remove ambiguity and direct attention to task processes (Kluger & DeNisi, 1996), an idea supported by the findings of L. J. Becker (1978) who asked 80 randomly selected families in California to set a goal to reduce their home energy use for a few weeks over summer by either an easy (2%) or a difficult (20%) amount. The combination of feedback and goal was found to be important. Those in the 20% group who received feedback reduced their energy use by 15.1% compared to 4.5% for those who had no feedback. Within the 2% goal group, those receiving feedback achieved a 5.7% saving whereas those with no feedback showed an insignificant 0.6% saving compared to the control group.
A habit is an automatic behaviour directed towards a goal (Verplanken et al., 1997), so it follows that deliberate goal-setting can play an important role in helping to bring habitual behaviours under cognitive control (Darnton, 2008). This may help to explain the results from McCalley & Midden, L. J. Becker, et al., as they challenged people’s patterns of behaviour and drew their attention to the feedback provided to them. As people become more competent in a task their attention wains over time (Verplanken et al., 1997) and they tend to think about a task at increasingly higher levels (e.g. the task of “changing equipment start-time schedules” becomes “working on optimising building efficiency”) (Kluger & DeNisi, 1996). This, according to Kluger and DeNisi, suggests that the effectiveness of task feedback will reduce over time and implies that it is important to regularly reassess and reset goals.

5.8. Fatigue limitations

The problem of attention and performance reverting over time to pre-feedback levels is known as the ‘fatigue effect’. Fatigue effects, which arise when research subjects become mentally disengaged with a process or task, work in the opposite direction to ‘practice effects’ – improvements in performance that occur due to repeated exposure to a process or task (Cozby, 2009). Both of these ‘order effects’ can be problematic for repeated measure research designs that require randomisation because they introduce a time variable that must be accounted for in addition to the treatment variable.

In this action research study, however, the practice effect is something we wish to encourage and fatigue effects are to be avoided through design, if possible. This requires the consideration of time as a key independent variable because we are interested in both the magnitude and rate of change in energy use associated with provision of feedback and facilitated learning environments (research design will be discussed in Part B).
Fatigue effects have not been adequately addressed in the literature relating to energy use, primarily because most study durations have been too short to determine whether savings will be maintained over the long-run (Delmas et al., 2013; Kluger & DeNisi, 1996; Stern, 2011). Delmas et al. (2013) found that close to 60% of published field studies that occurred between 1975 and 2012 lasted for less than three months (based on a near-comprehensive sample). Often, it seems, information from feedback devices gradually drifts into the background as the novelty wears off and people “simply lose interest” (van Dam et al., 2010, p. 460), an observation supported by Delmas’ meta-analysis which found that, “for each additional month of treatment [e.g. feedback], there is a small, but significant increase in energy usage…” (Delmas et al., 2013, p. 734).

Fatigue may be a particularly significant problem for fashionable or novel approaches, such as those involving software ‘apps’ or digital interaction (Chen et al., 2012), and there is considerable uncertainty regarding the long-term benefits of highly engaging short-term interventions. Some studies suggest an enduring impact arising from ‘carryover effects’ associated with investment in efficiency measures and changed habits (Bittle et al., 1979; Darby, 2006; Fischer, 2008), whereas others find that withdrawal of feedback is often associated with a return to original levels of consumption, particularly if savings are mostly attributable to an increase in task motivation (De Young, 1993; Hayes & Cone, 1981; Kluger & DeNisi, 1996). Studies of longer duration are therefore needed to resolve such questions (Delmas et al., 2013) and overcome an overarching view among policy-makers that behavioural change information measures are less reliable over time than measures involving investments in equipment (Martiskainen, 2007; Stern, 2011). Furthermore, if continuous feedback intervention is required in order to avoid reversal of savings in the majority of cases, a much clearer understanding of the ongoing
maintenance investment is required in order to evaluate behaviour-based methodologies (Kluger & DeNisi, 1996).
6. SOCIAL LEARNING CONTEXT

“Most of the behaviours that people display are learned, either deliberately or inadvertently, through the influence of example.”

(Bandura, 1977, p. 5)

6.1. Complex behaviours require models

Feedback can be a highly effective tool for accelerating learning. However it is not without limitations. Dynamic and highly cross-connected systems have complex interrelationships which cannot be completely understood through feedback. As Ashby (1957) observed, in a simple system where the relationships between actions and components are easily understood, feedback provides important and useful information. Introduce a few more components and interdependencies, however, and the relationships multiply dramatically with the consequence that specific and targeted feedback can no longer provide much useful information about the system as a whole (Ashby, 1957).

This proposition has been widely influential in the fields of control theory and robotics (Di Paolo, 2010), and is also reflected in the development of social psychology. As Bandura (1977) notes:

“Some complex behaviours, of course, can be produced only through the influence of models. If children had no opportunity to hear speech, for example, it would be virtually impossible to teach them the linguistic skills that constitute language. It is doubtful that one could ever shape intricate individual words, let alone grammatical speech, by differential reinforcement of random vocalisations”

(Bandura, 1977, p. 5).
People are able to learn more efficiently if they are given pointers so that they know, or at least think they know, what actions are likely to induce rewarding and punishing consequences. This is especially the case where tasks require a degree of analysis or customisation, for example when implementing operational changes to large commercial buildings which typically have highly customised combinations of features and operating characteristics. If an FM lacks background knowledge on a process or task, behaviour modelling (e.g. introducing social cues) is an indispensable aspect of learning and much more effective than merely applying consequences to unguided actions (Bandura, 1977).

6.2. Social Learning Theory

Albert Bandura’s (1977) social learning theory recognises that people learn from observing and imitating the actions of others. Unlike behaviourist approaches and social-psychological theories that focus on the direct relationship between stimulus and response, social learning theory is concerned with the cognitive processes that come into play when people are able to observe and learn from the experiences of others (models) by modelling their behaviour. Bandura identified the following four interrelated elements that govern how modelling occurs.

Attention: people seek out models that they find interesting or relevant to themselves. They are far less likely to pay close attention to, or seek to imitate, actions and characteristics of those they are not attracted to, even if such models could provide helpful guidance.

Retention: people are far more likely to remember and independently repeat modelled patterns of behaviour if they think about them and practice them.
Reproduction: people must be able to reproduce the behaviour they are attempting to model in order to learn it. This requires that they already have the component skills.

Reinforcement and motivation: people are more likely to pay close attention and attempt to match behaviours if they are incentivised to do so. Motivators affect the speed of translation from observation to action.

Early research underpinning social learning theory focussed on preschool children (3-6 years old). A famous experiment by Bandura et al. (1961) which exposed young children to adult models behaving aggressively towards a ‘Bobo doll’ found that they tended to imitate the behaviour of models—not just their aggressiveness, but their other behaviours as well—and they were more inclined to imitate models of their own sex. Behaviour is reinforced when people similar to ourselves show approval. Children also take into account what happens as a consequence, i.e. what consequences they observe. This is known as vicarious reinforcement because knowing that a behaviour is likely to bring about desired rewards (e.g. respect from or engagement with peers) or avoiding punishment is enough to increase attentiveness and motivation to imitate (Bandura, 1977). In social learning theory, reinforcement facilitates learning, but is not necessary (Bandura, 1977).

More broadly, it has since been found in studies involving older children and adults that people are inclined to emulate actions, beliefs, values and attitudes of those they admire and respect and that have a characteristic or quality they can relate to (see, for example, Braaksma et al., 2002; Y. Kim, 2007). This has interesting implications for learning relationships. Braaksma et al. (2002) and Y. Kim (2007) both studied the effects of similarity in competence between observers and models in educational settings. In both studies it was found that weak learners learned more and showed higher self-efficacy.
beliefs from working with less-competent models and the stronger students learned more and gained more confidence from working with highly competent models. In terms of perception and respect, weak learners had a higher regard for the lesser competence models than they did for the more highly competent models. The opposite was the case for the stronger students. Furthermore, the stronger students preferred situations where the models controlled the learning environment whereas the weaker students preferred to control the environments themselves.

In tasks requiring creativity or complex problem solving, such as energy optimisation problems associated with large commercial office buildings, strong learners gain most from observing the performance of models. People with lower learning aptitude, on the other hand, are more likely to succeed with direct instruction (Groenendijk et al., 2013). In such cases it has been found that it is generally most effective to provide an opportunity to follow a worked example broken into digestible components that can be performed and evaluated ‘step-by-step’ (van Gog & Rummel, 2010). Studying worked examples (e.g. a description of steps taken to achieve a goal) is a kind of observational learning; however, by breaking problems or activities into smaller components, learners gain more clarity and experience less cognitive strain than they do from following models because they are required to call on less working memory in order to successfully carry out a task (for discussion about cognitive load theory, see Sweller et al., 1998).

**6.3. Social proof and Theory of Reasoned Action**

The formation of connections between individuals is not based purely on shared characteristics and abilities. Physical proximity is also important, as was demonstrated by Festinger et al. (1950) who found that small social groups and interpersonal bonds form naturally on the basis of “sociological variables” such as common occupation,
background, gender etc. and are most likely to emerge between people who are neighbours and come into regular contact with one another (Festinger et al., 1950). This concept of ‘propinquity’ (physical or psychological proximity) is one of the key factors that determine attitudes and social behaviour patterns.

As already discussed, most people prefer not to deviate from a recognised community standard – especially if it applies to their direct peers. Knowledge of this has been used to inform various studies and campaigns promoting pro-social behaviour (see Schultz et al., 2007) including some prominent examples such as the study of household energy conservation behaviour by Schultz et al. (2007) cited above and one by Goldstein et al. (2008) that addressed water conservation in hotels. The Goldstein study found that reuse of towels in hotel rooms increased from 37 percent to 49 percent when a standard message card about water conservation was replaced by one that included information about a reference group (people similar to the hotel guest reading the card). They experimented with a variety of messages and found that an important “…factor affecting norm adherence is the extent to which individuals identify with the reference group” (Goldstein et al., 2008, p. 475).

These factors—propinquity and the referencing of behaviour to that of peers—combine to form the principle of ‘social proof’. For any given situation people will generally look to the prior responses of others for social proof to determine the appropriateness of their behaviour, with the influence level increasing according to how similar the peers are to the individual (Cialdini et al., 1999). Applied to the context of non-residential building operations, it suggests that FMs are most likely to form interpersonal bonds with operators of other nearby buildings within the same organisation and that they will look to those FMs for cues and guidance regarding their actions in managing their buildings. This
is further supported by the theory of reasoned action (Ajzen & Fishbein, 1980)—extended to become the theory of planned behaviour (Ajzen, 1991)—which asserts that people undertake activities on the expectation that they will derive positive value from the outcomes of their behaviour, including the benefits derived from the positive judgements of others.

**6.4. Social marketing and action**

Given the tendency of people to form social groups based on propinquity and the influence these connections have on behaviour, there is a clear opportunity to utilise groups as a means to facilitate action-learning (Bedwell et al., 2014; Deline, 2015; Dixon, Deline, McComas, Chambliss, & Hoffmann, 2015a; Jackson, 2005; McKenzie-Mohr, 2000). When applied to a community, this concept is known as ‘social marketing’ and generally involves the selection of a behaviour associated with an environmental goal (e.g. raise air conditioning temperature setpoints in summer), the identification of barriers to performing the activity, and then the design, piloting and evaluation of the impact of the program / strategy to overcome the identified barriers (McKenzie-Mohr, 2000).

When asked, individuals tend to underestimate the influence that social normative factors have on their actions because they are simply not aware of them (Schultz et al., 2007). And yet there are many examples of social marketing successes from around the world (see McKenzie-Mohr et al., 2012). Common to most of them is a clear focus on gaining a thorough understanding of the perceived barriers to performing the desired behaviours and the use of social norms and community engagement in addressing them (Martiskainen, 2007).

As already noted in relation to the ‘boomerang effect’, for social marketing it is important to be able to reference a norm that is in favour of a target behaviour, i.e. to be able to
inform individuals that “most people do the right thing” rather than pointing to information that suggests many people behave badly or that no effort is required (McKenzie-Mohr et al., 2012; Schultz et al., 2007). It is generally the case that referencing the behaviour of a more focused population (e.g. people that are similar in almost every way – location, experience, world-view, gender, etc.) has a slightly stronger effect than generic references, but even they can have a strong effect (Dixon et al., 2015a; McKenzie-Mohr et al., 2012). The use of ‘out of group’ references (e.g. to competitors or rivals) should be avoided, however, as they have been shown to have only limited influence and, in some cases, to backfire when people focus on the differentiating characteristics (Abrams et al., 1990).

One widely cited example of a successful social marketing strategy is the EcoTeam programs (Staats et al., 2004) developed in the Netherlands and since delivered to tens of thousands of homes in 20 countries (McKenzie-Mohr et al., 2012). EcoTeams are groups of six to ten people who meet once per month for a period of four to eight months to share ideas, experiences and achievements related to their household environmental behaviour. Participation is voluntarily and usually members know each other beforehand through friendship circles or neighbourhood relationships. The program’s methodology takes into account much of the knowledge about successful feedback and social learning processes described throughout this introduction and applies it to more than 90 behaviours in six themes: waste, gas, electricity, water, transport, and consumer behaviour. The 3-year longitudinal study of 150 Dutch participants reported in Staats et al. (2004) found that they changed 20 of the 38 targeted behaviours in a pro-environmental direction with significant reductions in use of gas, electricity and water. No behaviours changed in an anti-environmental direction. A follow-up study two years after completion showed
improvements were maintained or increased with 7 percent water savings and a 32 percent reduction in solid waste generation.

The EcoTeam program is also interesting for the fact that it achieves success without requiring participants to make commitments or set targets. As already discussed, techniques that encourage the making of commitments are effective in promoting pro-environmental behaviours, with public commitments generally more effective than those made privately (Pallak et al., 1980; see also L. J. Becker, 1978; Dwyer et al., 1993; Kluger & DeNisi, 1996; Locke & Latham, 2002). During the 1940’s and 50’s, pioneers in the field including Lewin and Revans brought groups of peers together to focus explicitly on achieving goals. Lewin (1951), who in 1944 coined the term ‘action research’, found participation in discussion groups was much more effective than lectures in promoting behaviour change and, furthermore, that the effects were maintained over time. Revans, who in 1945 coined the term ‘action learning’, made similar observations in his work bringing pit managers in the UK coal industry together in small groups of 5–8 peers to share their experiences and ask each other questions (Revans, 1997). Revans showed that by having the managers meet periodically at each other’s pits to focus on a problem faced in practice, over a 6 month period productivity increased by over 30% (Revans, 1980). Revans regarded learning as a function of ‘programming’ (or programmed knowledge, i.e. knowledge that is taught or read) and ‘questioning’ involving the use of closed (e.g. what?), open (e.g. why?), relative (e.g. where?) and objective (e.g. how many?) questions. More recently a ‘reflection’ component has been added (see Marquardt, 2004) to account for the refinement that occurs through considering trial and error feedback. Figure 6-1, developed by Kolb (1984), shows how the approach works.
Lewin described action research as “a spiral of steps, each of which is composed of a circle of planning, action and fact-finding about the result of the action” (Lewin, 1946, p. 37). For FMs, the context that ‘learning’ and ‘reflection’ take place in is therefore important because they may be unable to use feedback systems effectively, either in the short- or long-run, unless they have the support of an accompanying social learning framework. This point has been emphasised by Darby who states that feedback is necessary for a focus on energy savings, though “…it is not always sufficient – sometimes people need help in interpreting their feedback and in deciding what courses of action to take – but without feedback it is impossible to learn effectively” (Darby, 2006, p. 17). If actions are central to the process of learning in the work setting (Frese & Zapf, 1994), then environments that foster peer-to-peer questioning and reflection may provide the necessary conditions to maximise the benefits from automatic energy performance feedback (Janda, 2014; Moezzi & Janda, 2014).
PART B: RESEARCH DESIGN AND METHODOLOGY

7. QUESTIONS AND ORGANISING PRINCIPLES

7.1. Approach

The objective of this study is to establish whether facilities managers (i.e. operators) of Australian commercial office buildings will reduce their buildings’ energy use when provided automated feedback in a social learning environment. Energy use—specifically electricity—is the dependent variable we aim to measure. However, the amount of energy used to operate a non-residential building is influenced by a variety of factors aside from the competence and motivation of its FM. Weather conditions, such as temperature and humidity, occupancy patterns and user demands all have a significant influence on energy use in commercial office buildings. Given that these variables cannot be controlled by the researcher or building operator, the research requires a normalised or ‘steady state’ measure of energy use to be established so the effects of operator feedback can be clearly identified.

In a quantitative study of feedback-effects such as this, it would be too difficult (and subjective) to treat FM motivation and competence as independent and quantifiable variables because they are, in fact, unobservable intervening variables. Intervening variables assist our understanding of the relationship between independent and dependent variables but they are ‘states of mind’ that cannot be measured directly (Deci & Ryan, 1985). Significant and prolonged changes in energy usage patterns (dependent variable) will therefore point to changes in FM motivation and competence (intervening variables).
arising directly from feedback interventions (independent variable) if we select buildings with steady operating conditions (i.e. stable occupancy, consistent plant and equipment) and normalise for factors outside the operators’ control (e.g. weather). Our approach is illustrated in Figure 7-1, below.

Figure 7-1. Sequence of and interaction between variables in the research design

For the purpose of this study we regard operator motivation and competence as being closely associated and indistinguishable in terms of their effect. For example, we are not concerned whether energy savings arise through additional effort or more informed effort. Nor are we concerned with the degree to which operators deliberately act, reflect or learn (although we will consider all of these in the discussion of results). Our primary interest is to establish causality – to understand whether changes in energy use occur as a consequence of the feedback provided, and whether such changes can be observed generally across the portfolio of buildings included in the study, or not.
The intent of using the variables quantitatively is to compare outcomes for individual buildings within a portfolio and assess the effect on the portfolio overall (Mayer, 1980). To achieve this we will need to select a sample of buildings that is representative of the Australian commercial office stock and which is large enough to generate statistically significant results.

### 7.2. Model

A model can be a useful device for simplifying and ordering the reality of a complex situation and also for suggesting possible explanatory relationships among variables which can lead to the formation of hypotheses (Forcese & Richer, 1973). In chapter 4 we explained why Triandis’ (1977) Theory of Interpersonal Behaviour (TIB) is increasingly regarded as a useful social-psychological model for explaining diverse influences on energy consuming behaviours (e.g. Bamberg & Schmidt, 2003; Darnton, 2008; Jackson, 2005; Martiskainen, 2007). Its key advantage as a framework for empirical analysis lies in its treatment of behaviour as a function of intentions, habitual responses and the influence of ‘facilitating conditions’ – the conditions and constraints presented by a given situation.

The TIB considers a range of interacting factors that determine ‘intention’ and it is beyond the scope of this study to consider these in any detail. In this study we have assumed that intention and the attitudes, affect and social factors that determine it—and also the elements that influence them—remain stable throughout the baseline and intervention periods (see Figure 4-3). The justification for this is presented in the following chapter that describes the building sample and its characteristics. Similarly, due to the nature of the sample, the habits of FMs can be regarded as relatively entrenched.
Combined, these characteristics help to explain the stable energy usage patterns observed in the sample portfolio.

At the commencement of this study we can therefore consider the intention of FMs as ‘latent potential’ and their habits as ‘inertia’ and our efforts to avoid directly influencing either factor will be discussed in chapter 9. Rather, our objective is to introduce new facilitating conditions that are conducive to the adoption of energy saving behaviours and draw conclusions about the role of habits and deliberate actions (intention) by analysing the results over time. For this study we will therefore adopt a high-level application of the TIB model which is presented in Figure 7-2:

![Simplified version of Triandis' Theory of Interpersonal Behaviour (Triandis, 1977) limited to key variables considered in this study. The complete model is presented in Figure 4-3.](image)

The TIB model suggests behaviour (B) is determined by the influence of facilitating conditions over the intention (I) and habits (H) of subjects:

\[
B = F(H + I)
\]

(2)

Under the TIB, habits mediate behaviour and deliberate action becomes habituated as experience increases over time (Triandis, 1977). So if the results of behaviour are constant for an extended period (i.e. if energy use is stable), it is likely that many of the deliberate actions taken initially to achieve that level of performance will have become
habits. Changes to facilitating conditions may potentially make existing habits more effective or enable subjects participating in the study to achieve better results for any given level of effort regardless of habits. Facilitating conditions are likely to be unchanged if behaviours are unchanged.

Therefore, if conditions are adjusted to be more conducive to energy saving behaviours, i.e.

if \( F_{\text{new}} \neq F_{\text{old}} \), and if we are able to hold

\[
\sum(H + I)_{\text{old}} \approx \sum(H + I)_{\text{new}}
\]

then equation (2) can be further reduced to

\[
\Delta B \propto \Delta F
\]

The net effect on behaviour of altering facilitating conditions will, therefore, be directly proportional and independent of any reweighting of the roles of habit and intention. Specifically in relation to this study, this means that FMs’ adoption of energy saving behaviours will be proportional to the effectiveness of changes to the availability of operational feedback and the provision of learning opportunities and it will not matter whether those behaviours arise consciously or subconsciously.

7.3. Research question and sub-problems

As Creswell (2003) notes, “The most rigorous form of quantitative research flows from a test of a theory and the specification of research questions or hypotheses that are included in the theory” (p. 109). The TIB provides a strong theoretical framework to explore the relationship between the energy saving behaviours of commercial office building operators and the conditions facilitating those behaviours; however, it does not address
the relationship between the behaviours and the changes they bring about. Here, as discussed earlier and presented in Figure 7-1, we are concerned with measuring the changes in building energy use rather than the behaviours themselves. Therefore, the problem addressed by this study, and set out in the introduction, may be expressed as a research question:

*Will operators of Australian commercial office buildings reduce their buildings’ energy use when provided with automated daily energy performance feedback?*

We know from the literature presented in chapter 5 that feedback by itself is not effective unless coupled with access to critical background knowledge. In this study, the mechanism for providing this is the peer group. Therefore, to fully capture the intervention being trialled we must broaden the research question so that it becomes:

*Will operators of Australian commercial office buildings reduce their buildings’ energy use when provided with automated daily energy performance feedback within a social learning environment?*

This study is focussed on a single portfolio of Australian commercial office buildings. However, our objective is to develop an understanding of the scale of energy saving and greenhouse gas emission reduction opportunities arising from behaviour-based approaches applied to Australian commercial office buildings in general, and to consider the implications for other building typologies and for buildings in other regions. This introduces a series sub-problems tied to the interpretation of the data which, together, add up to the totality of the research problem (Leedy & Ormond, 2005).
The first sub-problem

The first sub-problem is to identify a representative sample of Australian commercial office buildings and to establish a basis for comparing performance between buildings. This will require a definition of the “Australian commercial office building” typology and an assessment of the sample portfolio with reference to its peers prior to commencement of the trial.

The second sub-problem

The second sub-problem is how to develop a baseline to evaluate changes in a building’s energy performance and, having done so, determine a method for assessing its validity. This sub-problem relates to the first insofar as a steady-state baseline will be required in order to compare energy usage changes in a building and between buildings. A solution to this sub-problem will also meet the requirement presented in equation (3) and thus allow simplification of the TIB’s relationship between facilitating conditions and behaviour presented in equation (4).

The third sub-problem

The third sub-problem is how to design a mechanism for automating energy performance feedback to FMs. The system will need to generate detailed, accurate and timely insights. This will require the development of software that is reliable and consistent so that recipients perceive the feedback as valid.

The fourth sub-problem

The fourth sub-problem is to isolate the effects of feedback and the changed environmental factors from other variables that might influence energy use in the trial.
buildings and relate these to changes in baseline energy use. This will require careful selection of the sample for trialling (and exclusion of any buildings where other energy-related factors change independently during the trial) and also rigorous normalisation to remove the influence of otherwise confounding variables such as occupancy and weather.

7.4. Causal hypothesis

Notwithstanding the limitations of previous studies, the literature gives grounds to anticipate a causal relationship between operator feedback and decreased energy use. A causal hypothesis will therefore help to translate the research question into a prediction that can be formally tested in this study:

 Operators of a representative sample of Australian commercial office buildings, when exposed to automated daily energy performance feedback in a social-learning environment, reduce energy usage relative to their building’s normalised baseline.

In the following chapters I will set out an action learning research design to test this hypothesis.
8. PORTFOLIO SELECTION AND PRELIMINARY REVIEW

8.1. Portfolio selection

There are a variety of explanations for why most research focused on behaviour-based approaches for saving energy in buildings addresses residential, rather than non-residential, buildings (Strachan, Janda, & McKeown, 2015). One reason is the difficulty of recruiting subjects. Owners of commercial office buildings are, by definition, motivated to generate commercial returns from their assets and the motivation for almost all decisions can be traced back to a commercial objective: maximising present and future income (i.e. revenue minus costs). This particular behaviour conforms to the rational choice model. For example, activities aimed at increasing a building’s appeal to occupants will normally be supported if they are expected to increase revenue by more than any associated increase in costs. Likewise, activities designed to reduce operating costs (e.g. through energy savings) without causing detriment to occupants will normally be supported if the owners are confident that net income will rise. However, because energy expenditure (in Australia) is typically 1/20th of the rent generated from tenants (Property Council of Australia, 2009), caution prevails. Typically, owners are reluctant to “experiment” with their buildings’ energy use for fear that it may adversely impact occupants or take up a disproportionate amount of management time and attention. And for the same reasons owners tend not to participate in experiments conducted by others. Given that commercial office building FMs are service providers engaged to represent the building owners’ interests – they take the same view.

It is, therefore, quite challenging to persuade an owner or operator to participate in an experimental study such as this. To be persuaded the owner / operator must form a view that the research is unlikely to harm their commercial returns and may in fact deliver a
benefit that compensates them for the inconvenience and expense they anticipate arising. On the flip side, having been satisfied that an intervention is likely to yield a commercial benefit, an owner / operator is unlikely to favour a randomised control trial (RCT) because assigning half their portfolio to a control group would simply halve the potential benefit accruing from the decision.

These were some of the realities faced when seeking to recruit the sample portfolio for the present study. The pragmatic solution was therefore to adopt an action research methodology using statistical models to provide normalised baselines for evaluation rather than referencing a randomised control group. All suitable buildings within a participating portfolio would join in the study and all would be exposed to the conditions that the causal hypothesis suggested were likely to bring about energy savings, and therefore commercial benefit to the owner of the portfolio.

**8.2. Investa Property Group portfolio**

As an employee of Investa Property Group, a large Australian Real Estate Investment Trust (REIT), I was fortunate to have the opportunity to pitch a research proposal directly to executive management and have it thoroughly evaluated (see Appendix A). The proposal was simply to test the causal hypothesis, so approval was granted on the expectation that the portfolio would benefit if the research findings supported the hypothesis.

Investa’s portfolio was ideally suited to this study. It is large and had a mature energy efficiency investment program that had yielded significant savings initially, but energy performance had plateaued in the years immediately prior to the commencement of this research project in 2009. Descriptive statistics for the portfolio are summarised in Figure 8-1 below.
Investa at the time of this research was regarded as a leader in responsible property investment globally and had received almost 100 awards in corporate responsibility fields prior to the commencement of this study. Prior to its delisting from the Australian Securities Exchange (ASX) in 2007 (upon being acquired by Morgan Stanley Real Estate Investment funds), it was the leading company on the Dow Jones Sustainability Index (DJSI World), both for the Real Estate sector and also the Financial Services ‘super-sector’. In the same year it was also included in the Global 100 list of the world’s most sustainable public corporations announced at the meeting of the World Economic Forum in Davos. It was the first property fund manager to be certified by the Responsible Investment Association of Australasia (RIAA) and its principal wholesale investment fund, ICPF, was rated 2nd out of the 688 listed property companies and private property funds included in the Global Real Estate Sustainability Benchmark (GRESB), with its “environmental management and policy” rated best in the world.

<table>
<thead>
<tr>
<th>Electricity</th>
<th>2003/04</th>
<th>2004/05</th>
<th>2005/06</th>
<th>2006/07</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buildings in program with continuous statistics</td>
<td>23</td>
<td>23</td>
<td>27</td>
<td>28</td>
<td>28</td>
<td>33</td>
<td>33</td>
</tr>
<tr>
<td>Floor area (NLA) of buildings with continuous statistics (m²)</td>
<td>463,150</td>
<td>460,911</td>
<td>469,329</td>
<td>703,110</td>
<td>703,110</td>
<td>760,403</td>
<td>772,679</td>
</tr>
<tr>
<td>Metered electricity consumption (MWh)</td>
<td>62,748</td>
<td>56,515</td>
<td>78,533</td>
<td>75,820</td>
<td>74,400</td>
<td>81,267</td>
<td>81,261</td>
</tr>
<tr>
<td>Consumption Intensity (kWh/m²)</td>
<td>135</td>
<td>123</td>
<td>116</td>
<td>108</td>
<td>107</td>
<td>107</td>
<td>106</td>
</tr>
<tr>
<td>22% Reduction since 2003/04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 8.1. Investa Property Group portfolio electricity intensity (Investa Property Group, 2010)**
These ‘sustainability’ credentials clearly demonstrated the portfolio was ahead of trends emerging at the time (see Figure 2-2). Furthermore, its plateauing energy intensity results and demonstrated commitment to environmental management reinforced the perception (both internally and from outside) that there were relatively few ‘easy wins’ still available for FMs to improve energy performance. This gave the study credibility with Investa’s executive who elected to support it, and was expected to be helpful in achieving a stable and consistent environment for evaluating the intervention.

Throughout the pre-trial period the portfolio’s management team maintained a detailed and accurate measurement, monitoring and reporting platform that included independent assurance of all published environmental performance data by a top-tier international accounting firm (Investa Property Group, 2010). It also invested considerable resources in developing the ability of its facilities management teams to drive continuous improvement as well as technologies, procedures and practices aimed at reducing impacts or enhancing the performance of its buildings for their occupants (Roussac, 2009).

The combination of robust monitoring and independent assurance suggested the published statistics were highly credible. The net lettable floor area of buildings (NLA) was calculated for the buildings’ leased and unleased spaces by qualified surveyors in accordance with the Property Council of Australia measurement standard. Occupancy statistics were maintained by Investa as part of the financial accounting function. Combined, these data showed that most buildings had stable occupancy throughout the analysis period and that occupancy was stable for the portfolio overall (Roussac, 2009). Other data, such as the utilities’ meter data and weather observations were sourced by Investa from independent and highly reputable authorities such as the Australian Bureau of Meteorology (BoM) and AEMO-accredited meter data providers.
Selecting and focusing on a single large portfolio presented other advantages. It meant that we could be reasonably sure that organisational factors—such as policies, systems and procedures, culture and technology platforms—which can greatly influence the effectiveness of behaviour-based interventions (Malone et al., 2011) were relatively uniform across all participating buildings. This would not have been the case if we had sought to recruit an equivalent number of buildings from multiple owners and operators. A lot of variables which may otherwise have influenced the results did not vary between buildings and their possible influence on the dependent variable could, therefore, be excluded from the analysis. Another, less beneficial, consequence of removing these variables is that we have no measure of their relevance. As explained in the introduction, quantifying the potential impact of organisational factors is outside the scope of this study; however it is important to reflect on them—in particular, whether they could impact the generalisability of findings—and they will therefore be discussed in Part D.

8.3. Pre-trial review of a subsample

As outlined above, at the commencement of this study Investa was a major player in the Australian commercial office sector (and it remains so). The NLA of its portfolio exceeded 1 million square metres—equivalent to almost 5 percent of the Australian commercial office market—and in some key locations such as the Sydney and Melbourne CBDs it had an interest in roughly 10 percent of the commercial office stock.

Aside from the improvement statistics and sustainability accolades, in most other respects the portfolio appears representative of the broader market (noting that many statistics are not available from the Property Council of Australia or other credible sources). The average age of buildings was 20-25 years and the average NLA was ~25,000m². The weighted-average base-building NABERS Energy rating during the 2003/04 fiscal year
was 2.6 stars, almost identical to the market average at the time of 2.5 stars (out of 5 at that time). Over the period 2003/04 – 2009 the portfolio’s base-building (i.e. excluding tenancy) electricity and natural gas usage intensity reduced by 22 and 50 percent respectively, corresponding to 28.5 percent reduction in total energy use. This was confirmed by the weighted-average NABERS Energy rating which by 2009 had climbed 1.13 stars to 3.73, indicating that by 2009 the portfolio was using between 25 – 30 percent less energy than the median Australian CBD office base-building rated using NABERS (Investa Property Group, 2010).

Consistent with the approach suggested by Mayer (1980), an initial evaluation of a subset of the portfolio was conducted in an attempt to identify patterns, inconsistencies or clues regarding the performance of the entire portfolio. It was found that energy intensity ranged widely within the portfolio as illustrated in Figure 8-2 below.

![Figure 8-2. Base-building electricity consumption during the pre-trial period 2009-10 (Roussac, de Dear, & Hyde, 2011)](chart)
A sample of eleven buildings comprising 298,000 m$^2$ of NLA (~1.5 percent of the Australian commercial office market) and representative of the portfolio’s mix of size, age, location and operating platform was selected from the portfolio for a detailed preliminary analysis (see Roussac, 2009; Roussac, de Dear, et al., 2011). Buildings younger than ten years of age were excluded. A summary of these buildings is presented in Table 8-1 below.

Table 8-1. Sample of eleven of Investa’s commercial office buildings (Roussac, 2009)

<table>
<thead>
<tr>
<th>Building ID</th>
<th>CBD location</th>
<th>Approx. age (years)</th>
<th>Approx. size (m$^2$ NLA)</th>
<th>2004 energy use (MJ/m$^2$.yr)</th>
<th>2008 energy use (MJ/m$^2$.yr)</th>
<th>Change (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Building Y</td>
<td>Nth Sydney</td>
<td>18</td>
<td>15,000</td>
<td>465</td>
<td>462</td>
<td>-1%</td>
</tr>
<tr>
<td>Building AA</td>
<td>Brisbane</td>
<td>20</td>
<td>20,000</td>
<td>469</td>
<td>346</td>
<td>-26%</td>
</tr>
<tr>
<td>Building –</td>
<td>Brisbane</td>
<td>33</td>
<td>25,000</td>
<td>516</td>
<td>317</td>
<td>-39%</td>
</tr>
<tr>
<td>Building G</td>
<td>Adelaide</td>
<td>20</td>
<td>25,000</td>
<td>535</td>
<td>604</td>
<td>13%</td>
</tr>
<tr>
<td>Building Q</td>
<td>Sydney</td>
<td>37</td>
<td>27,000</td>
<td>567</td>
<td>377</td>
<td>-34%</td>
</tr>
<tr>
<td>Building K</td>
<td>Sydney</td>
<td>18</td>
<td>23,000</td>
<td>586</td>
<td>413</td>
<td>-30%</td>
</tr>
<tr>
<td>Building AC</td>
<td>Perth</td>
<td>25</td>
<td>18,000</td>
<td>597</td>
<td>404</td>
<td>-32%</td>
</tr>
<tr>
<td>Building AB</td>
<td>Sydney</td>
<td>18</td>
<td>28,000</td>
<td>661</td>
<td>342</td>
<td>-48%</td>
</tr>
<tr>
<td>Building O</td>
<td>Melbourne</td>
<td>15</td>
<td>65,000</td>
<td>700</td>
<td>393</td>
<td>-44%</td>
</tr>
<tr>
<td>Building X</td>
<td>Sydney</td>
<td>20</td>
<td>29,000</td>
<td>791</td>
<td>388</td>
<td>-51%</td>
</tr>
<tr>
<td>Building B</td>
<td>Sydney</td>
<td>19</td>
<td>23,000</td>
<td>942</td>
<td>827</td>
<td>-12%</td>
</tr>
<tr>
<td>Mean:</td>
<td></td>
<td>22</td>
<td>27,100</td>
<td>621</td>
<td>443</td>
<td>-29%</td>
</tr>
</tbody>
</table>

The sample’s average energy use intensity in 2004 was identical to the Investa portfolio (621 MJ/m2.yr) and the 29 percent reduction in energy use per square metre measured between 2004 and 2008 was very near the 25 percent average for the portfolio, and would have been even closer if Investa had not acquired a number of ‘below average’ buildings during this period (the average reduction for the portfolio reached 29 percent the following year). Table 8-1 shows the eleven buildings sorted from lowest energy use in 2004 to highest and illustrates that there was no clear relationship between this metric and

---

5 Building IDs in the table (Roussac, 2009) have been amended so they correspond with Building IDs used in Figure 8-2 (Roussac, de Dear, et al., 2011). The third building in the table (Building –) was sold during this phase of the research.
building age, size or percentage improvement. The fact that energy intensity still varied widely in 2009 even after a significant portfolio-wide effort at achieving energy savings, as illustrated in Figure 8-2, suggests that the differences could not be easily resolved and may have come down to the influence of physical characteristics (such as façades, control systems, plant and technologies) and varying levels of management motivation and capability (Roussac, de Dear, et al., 2011).

8.3.1. Some cursory observations

Research by Granderson et al. (2013) on the costs and benefits of deploying energy information systems (EIS) found the greatest influences on the level of savings are: (1) a high pre-existing level of energy use intensity (EUI), and (2) the extent of efficiency projects undertaken before EIS installation (i.e. pre-existing projects increase effects of EIS). Interestingly, they also found that “User engagement and user empowerment were not strongly associated with greater energy savings” (Granderson et al., 2013, p. 21). They did stress, however, that these human factors may still be important because a variety of potential explanations were outside the scope of their research. All of the buildings in Investa’s portfolio had sub-metering and monitoring (hence, EIS) systems installed circa 2004 and it is interesting to note that the bottom five in the pre-trial sample (Buildings AC, AB, O, X & B) produced an average saving of 37 percent over the four year period compared to only 17 percent for the top five (Y, AA, –, G & Q). Observations from this sample of the Investa portfolio would appear to support the findings of Granderson et al. and point to the potential benefit of exploring the user engagement / empowerment aspect. The two buildings that produced the largest pre-trial savings were Buildings X and AB. These were isolated for more detailed review because as Flyvbjerg, amongst others, has observed “atypical or extreme cases often reveal more information
because they activate more actors and more basic mechanisms in the situation studied” (Flyvbjerg, 2001, p. 78).

8.3.2. Building X, Sydney

Building X is a 32 storey A-grade office building in Sydney CBD that provides 29 levels of office accommodation, one level of retail and a ground floor entry foyer. It was fully leased to a single tenant and does not have any car parking.

Figure 8-3 shows the building’s pattern of energy use throughout the period 2004-2008 inclusive. The faint grey energy baseline is a simple projection of the 2004 usage pattern over the subsequent four years and provides an approximation for the expected seasonal profile. For the first year and a half the building did not show any consistent pattern of energy savings, with measured energy consumption closely following the baseline. From June 2006 there was a dramatic shift, with energy use during the late winter and spring of 2006 lower than the preceding winter whereas previously spring was a period when energy use would normally have increased. Operations personnel attributed these savings to a combination of the completion of tenancy refurbishment works, a new senior facilities manager and upgrades to building plant and equipment (flags 2-4) (Roussac, 2009).
Each of the flags in Figure 8-3 marks the introduction of a significant intervention aimed (in part or exclusively) at reducing energy use:

1. Outside air dampers for economy cycle replaced ($72k)

2. BMCS field controllers upgraded ($93k)

3. New highly competent Senior Facilities Manager appointed to manage the property

4. Centrifugal chiller replaced with screw type ($221k)

5. BMCS controller program fine-tuned ($22k)

6. Chiller oil additive introduced to enhance thermal conductivity ($35k)

7. Installed individual floor dampers for perimeter Air Handling Units (AHU) ($109k)
8. Installed lighting controls to base-building ($25k)

- Capital invested: $627,000
- Total saving (2005–2008): $622,000
- Projected annual saving (at 25% ROI): $156,750
- Annual saving (in 2008 cf baseline): $327,500

(Roussac, 2009)

The extent to which the technology improvements contributed to the savings is unclear and the case highlights the difficulties of disentangling the contributions of technological and behavioural factors. It was a full 12 months after the actions cited above (flags 2-4 plus tenancy refurbishment) had been implemented that the performance started to improve. It is highly probable that the equipment improvements made a contribution not just to the underlying building efficiency but also the motivation and effectiveness of the building’s operations and facilities management team which suggests there may have been undocumented interaction effects.

Another important point noted in the presentation of these results was that the improving energy performance appeared to be correlated to the reduction in the number of complaints recorded by Investa’s tenant helpdesk. In other words, the occupants’ dissatisfaction with their office comfort conditions appears to have reduced as the building’s energy use reduced (Roussac, 2009). Again, the degree to which this was brought about by equipment improvements or management competency is unknown.
8.3.3. Building AB, Sydney

Building AB is an A-grade office building, also located in Sydney CBD, that comprises 15 levels of commercial office accommodation, two levels of retail, six levels of basement car parking (149 spaces) and a rooftop plant room. It was fully leased to a small number of tenants.

Figure 8-4 (below) is arranged identically to Figure 8-3. In early 2004 a few modest changes were introduced to improve the manageability of the building and for almost two years afterwards there were very few additional changes made. Initially it was thought that savings were arising due to a significant vacancy in the building, rather than the minor upgrades performed, however energy use dropped further once the building became fully occupied again. The introduction of variable speed drives (VSD) to many of the motors on the building’s pumps and fans (flag 4) will have contributed significantly to savings; however this intervention is unlikely to have accounted for the full extent of the improvement from March 2006 onwards. Likewise, the introduction of the chiller oil additive is unlikely to have had a large impact as evidenced from the trend following this intervention in Figure 8-3. As was the case with Building X, the improvement in energy performance corresponded with a very significant reduction in the number of hot/cold complaints registered with the tenant help portal.
Each flag marks a significant intervention aimed (in part or exclusively) at reducing energy use:

1. Outside air dampers for economy cycle replaced ($14k)
2. On-floor air balance ($23k)
3. BMCS ‘head end’ (user interface) upgraded ($15k)
4. VSDs to fans and pumps ($86k)
5. Chiller oil additive introduced to enhance thermal conductivity (35k)

- Capital invested: $223,000
- Total saving to-date (2005–2008): $736,000
- Projected annual saving (at 25% ROI): $55,750
• Current annual saving (2008 cf baseline): $251,000

(Roussac, 2009)

It is highly probable that the energy efficiency of Building X, like Building AB, improved on account of the technology upgrades, some behaviour-based interventions, and interactions between the two. Unfortunately, no reliable records of non-technological interventions were maintained and it has not been possible to go back and evaluate what other interventions may have been occurring due to changes in staffing, contractor engagements, etc. What this does appear to point to, however, is a marked improvement in operating practices and a general effort to achieve energy savings that went beyond what could have been anticipated from the capital investments alone. This would help to explain why measured savings in 2008 were almost five times greater than the savings anticipated from the capital investments themselves.

8.3.4. Relationships between energy intensity, occupant satisfaction, technology and behaviour

As the two cases illustrate, there appears to have been a relationship between the energy performance of each building and the number of complaints for “too hot” and “too cold” conditions registered with Investa’s tenant helpdesk. As energy use decreased, so too did the number of complaints. This relationship held remarkably well across all buildings included in the sample, with eight of the nine buildings for which reliable helpdesk data was available showing the same pattern (Roussac, 2009). One of several plausible explanations for this was put forward by Bordass, Leaman and Ruyssevelt in their ground-breaking “Probe” studies:

“There is little or no direct relationship between comfort and energy efficiency, but an important indirect one, in that good management of the
The other particularly interesting observation to come out of the pre-trial sample review was the relationship between expected and actual results. As noted in the two examples discussed above, energy savings were considerably larger than expected: for Building X they were double and for Building AB they were almost five times the expected level. While this was the case for buildings at the upper end of the spectrum, at the lower end it was found that savings were in some instances lower than anticipated. This might suggest that in those instances management may have contributed very little or even had a negative impact.

Overall, it was estimated that investments in technology and operator influences each contributed approximately half of the sample group’s total energy savings. The composition of the behaviour-based contribution was unclear, however. For example, it was found that the benefits derived from the installation of sub-meters (a capital investment made in all sample buildings immediately prior to commencement of the analysis) were heavily impacted by the FM’s technical competency and willingness to monitor them. At the same time, their motivation levels would have been influenced by factors such as the enhanced training, reporting and incentives introduced throughout the period (see Roussac, 2009).

**8.4. Initial prototype and its limitations**

Drawing on the idea that FM engagement levels were quite variable and also potentially influenced by organisational factors, an attempt was made to raise awareness (and, therefore, motivation) by making the detailed performance data for all Investa buildings
publicly accessible via a web-based data visualisation platform embedded in Investa’s 2009 Sustainability Report (Investa Property Group, 2010). A screenshot of the interactive applet is shown in Figure 8-5 below.

Figure 8-5. Publically accessible web-based energy performance data for Investa buildings. Snapshot of the Investa 2009 Sustainability Report

The idea for this initial exploration of ‘transparency’ as a first step in a series of action research interventions drew on Zeisel’s (1981) framework which suggests that “making visible the implications of the data leads to improved hypotheses, further data gathering, and so on until the problem is sufficiently redefined and a tenable solution is found” (p. 18). It was thought that the public disclosure of data would invite questions about
building performance which would in turn support, encourage and motivate the FMs (or their superiors who control the finances) to take effective action.

The interactive data visualisation tool was a highly sophisticated technical advance on the typical paper-based environmental performance reporting of the time, but as a motivation-building intervention it was quite crude. Would the savings expected to arise from a sharper focus on under-performing buildings be enough to offset a possible boomerang effect such as that identified in residential studies (e.g., Bittle et al., 1979-80; Brandon & Lewis, 1999)? And, indeed, where was the empirical evidence supporting the assumption that disclosure would directly motivate operators of underperforming buildings and help them to identify and implement successful interventions? If FMs were motivated, would it be by fear or the prospect of recognition? As noted by Block (1999), the relationship between ‘fear’ and ‘persuasion’ is complex and individuals are likely to react to the prospect of negative consequences (i.e. punishment) in a variety of ways depending on various personality and message variables. This was highlighted in a survey of FMs and their immediate superiors (off-site senior facilities managers) at the time of the 2009 Sustainability Report’s publication. While the initiative was generally regarded as beneficial for the company (reputational benefits), senior FMs were on average more supportive and less wary than the building level FMs (Roussac & Bright, 2012).

More telling was its influence on behaviour. After the initial buzz associated with the report’s release, it quickly drifted into the background. It became clear that FMs were not changing behaviours on account of the publication of their buildings’ data. It was not providing reinforcement because the interactions were only with static data. And even if it introduced the expectation of punishment for under-performance, it would most likely be ineffective at encouraging positive behaviours given that the objective of punishment is to
weaken behaviour (as discussed in chapter 4). Finally, it did not help FMs to identify actions that might deliver energy savings. This experience, therefore, confirmed the need for a more sophisticated intervention design that took into account the extensive literature presented in Part A.

Data must be refined in order for it to become information. Investa’s attempt to harness data visualisation as a motivational tool failed in part because the interface was designed to facilitate ‘exploration’ of the data rather than provide a means to refine it and make it more actionable or meaningful. The audience (which included both internal and external ‘stakeholders’, including the FMs) engaged with the data, but they were unable to confidently form judgements about whether an individual building’s performance was actually ‘good’ or ‘bad’, because crucial background information about operational constraints and requirements was omitted. Even if users of the tool had been able to access the requisite metadata to inform judgements, it is unlikely that anything more would have come of it because no clear link was provided between ‘actions and their effects’ (see Fischer, 2008), i.e. there was no feedback.
9. INTERVENTION DESIGN

9.1. Automatic feedback

Highly effective energy performance feedback for building operators requires a number of key ingredients which were all discussed in chapter 5. First among them is the requirement that it:

- consist of simple, clear, reliable, high quality messages delivered consistently and frequently (with limited delay).

The literature also suggests that effective feedback will:

- reference some kind of a goal, commitment or expected baseline standard;

- contain information that is directly relevant to the recipient in order to promote self-assessment (reflective thinking) and investigation;

- build positive motivational beliefs and self-esteem that promote experimentation and support constructive dialogue with experts, specialists and peers.

In the context of operating buildings there is also an attention aspect that needs to be considered: FMs are generally time-poor and are constantly having their attention drawn to time critical ‘issues’ raised by occupants, owners, contractors, the public, etc. It is not uncommon for a site-based FM’s mobile phone to ring up to a dozen times during the course of a half-hour meeting. I experienced this first hand when undertaking the preliminary review. It soon became apparent to me, and this was subsequently confirmed by the FMs themselves, that it would be unrealistic to expect anyone in such a role to set aside a regular amount of time each day to logon to an online energy performance feedback service.
An automatic feedback system was designed specifically to address all of these overarching objectives and constraints. The components are illustrated in Figure 9-1 and described below.

![Figure 9-1. Message system architecture.](image)

### 9.1.1. Data sources

The feedback intervention focussed exclusively on electricity use. For Investa, electricity use accounted for more than ninety-five percent of the portfolio’s greenhouse gas emissions and it also represented a large financial expense. Natural gas use, by comparison, accounted for less than three percent (Investa Property Group, 2010, 2011, 2012).

Another reason for focusing on electricity use was the quality and availability of data. The Australian Energy Market Commission (the body that sets the national electricity, gas and energy retail rules) requires accredited Metering Data Agents (MDA) to provide daily access to electricity meter data and up to two years of historical data if requested to
do so by a customer or its authorised representative (AEMC, 2014). Data must be in a 15-minute interval format that conforms to the official Meter Data File Format Specification (either NEM12 or NEM13). Two years of historical data and daily NEM12 data was sourced from Infomet and TCA Ausgrid, Investa’s MDAs. Gas data was not available for most buildings in electronic format.

Weather data was sourced from the Bureau of Meteorology (BoM) – the Australian Government’s official weather agency. Two years of historical 15-minute interval data was sourced from official weather stations nearest to the buildings in question and a daily data service was established from each.

Access to submeter electricity data (historical and ongoing) was also arranged for many of the buildings’ via Investa’s submeter service provider. This data was delivered in NEM12 format, but its reliability was found to be much lower than the MDA standard, so it was not used as a core component of the daily feedback.

All electricity data were sent from the 3rd party source to a single Secure Shell File Transfer Protocol (S/FTP) server where it was collected by the software application for processing and then stored. The BoM data were handled in a similar way, however they were transmitted via FTP rather than S/FTP.

Other data, such as public holidays and day and date identifiers, were input to the database manually.

**9.1.2. Statistical model**

A statistical model to predict daily electricity use was constructed for each building by taking into account weather conditions and also weekly operating patterns. This model provided the benchmark for advising building operators whether measured electricity use
was higher or lower than “expected” and for identifying the day in the previous year when conditions were most similar and, therefore, electricity usage profiles could be expected to be comparable.

Following is a description of the statistical model and formulas used in the trial.

As a multiple regression method was used, the first step was to calculate the required variables from the 15-minute interval data series. A matrix $DXO$ (daily external observations) of dimension $N_{columns} \times 96$ was created for each day from 00:15 hrs (first measurement, No. 1, is for the period 00:00 – 00:15) to 24:00 hrs (measurement No. 96):

$$DXO = \begin{pmatrix}
T_1 & RH_1 & P_1 & \ldots & \ldots \\
T_2 & RH_2 & P_2 & \ldots & \ldots \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
T_{96} & RH_{96} & P_{96} & \ldots & \ldots
\end{pmatrix}.$$  \quad (5)

Rows (1-96) correspond to observations and columns to variables ($N_{columns}$ in total). As depicted in equation (5), $T$ is dry bulb temperature in degrees C, $RH$ is relative humidity expressed as a percentage, $P$ is atmospheric pressure in hPa. Various other columns were also included, but not used, such as wind speed in m/s and precipitation in millimetres.

Similarly a matrix (vector) for electricity consumption data, $DEU$ (daily electricity usage), was constructed for each day:

$$DEU = \begin{pmatrix}
E_1 & \ldots & \ldots \\
E_2 & \ldots & \ldots \\
\vdots & \vdots & \vdots \\
E_{96} & \ldots & \ldots
\end{pmatrix}.$$  \quad (6)

where $E_i$ is electrical energy usage in kWh measured at 15-minute increments. As with $DXO$, other columns were included because the data supplied by the MDAs included
additional parameters including power, current and voltage, but they were not used in any calculations during the trial.

\( DXO \) matrices and \( DEU \) vectors were recorded and stored in the database.

The next step was to pre-process and quality assure the data, including (a) removal of days with fewer than 96 measurements (missing values), (b) removal of \( NaNs \) (undefined or non-representable values), (c) removal of weekends and public holidays, and (d) removal of outlier days with unusual operations such as shutdowns for maintenance of major plant and equipment.

Intra-daily data presented by \( DXO \) and \( DEU \) matrices were then consolidated and combined with day-of-week identifiers (binary – if Friday, no day would be identified) to create each building’s model calibration database of external observations, \( XO \)

\[
XO = \begin{pmatrix}
T_{1av} & RH_{1av} & T_{1max} & TP_{1av} & RHP_{1av} & M_1 & T_1 & W_1 & Th_1 \\
T_{2av} & RH_{2av} & T_{2max} & TP_{2av} & RHP_{2av} & M_2 & T_2 & W_2 & Th_2 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
T_{Nav} & RH_{Nav} & T_{Nmax} & TP_{Nav} & RHP_{Nav} & M_N & T_N & W_N & Th_N
\end{pmatrix}
\]

This database has dimensions \( 9 \times N_{days} \), where \( N_{days} \) is the number of rows (number of days recorded). The mean of daily \( DXO \) measurements from 34 (08:30) to 70 (17:30) was used to calculate values for average dry bulb temperature (\( T_{av} \)) and average relative humidity (\( RH_{av} \))\(^6\). The maximum dry bulb temperature (\( T_{max} \)) was the highest value recorded between measurements 34 and 70 (inclusive). The mean of daily \( DXO \)

\(^6\) The variable \( RH \) takes into account the moisture content of the outside air and the energy lost to condensation by HVAC systems when cooling it. As \( RH \) varies proportionally with temperature (for any given moisture quantity, \( RH \) goes down as air temperature goes up), it is acknowledged that it would be better to decompose \( RH \) and use air moisture (grams per kg of dry air) instead. This was done for an alternative evaluation model described in the next chapter, but not during the trial.
measurements from 1 to 33 was used to calculate values for average overnight dry bulb temperature ($T_{Pav}$) and average overnight relative humidity ($RH_{Pav}$).

Similarly, electricity usage was summed for the core operating period of each day from 04:30 (first measurement No. 19 at 04:45 for the period 04:30 – 04:45) to 18:00 (measurement No. 72) and presented as a vector for each building’s total daily electricity usage, $EU$

$$EU = \left( \begin{array}{c} E_{1t} \\ E_{2t} \\ \vdots \\ E_{Nt} \end{array} \right)$$

(8)

where $E_{it} = \sum_{j=19}^{72} E_j$.

Measurements 1 to 18 and 73 to 96 were excluded because of the unpredictability of after-hours electricity usage and its tendency to be driven by factors outside the control of the operators, such as tenant air conditioning requests. In some cases early-morning startup procedures commence well before occupancy hours and it was considered important for feedback to capture this period even though office buildings are normally unoccupied before 07:00.

During initial testing it was found that linear regression was able to provide a reasonably accurate prediction for electricity use during typical operating conditions (>90% of days), but relationships became nonlinear when whether conditions were anomalous (either unusually hot or unusually cold). A formula was developed to account for the residuals found in each building’s models at the extremes, i.e. the difference between the observed electricity use ($y$) and the predicted use ($\hat{y}$). This equation took the form

$$R = aT_{av}^2 + bT_{av} + c$$

(9)
where the residual, $R$, was calculated by performing a “first pass” linear regression on matrix (7). Figure 9-2 shows the residual distribution for one of the more significantly affected buildings.

![Residual Distribution Plot](image)

**Figure 9-2. Plot showing distribution of residuals for one of the Melbourne buildings.**

Following this step, an additional column $\begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_N \end{bmatrix}$ was added to matrix (7), thus forming

$$XO' = \begin{bmatrix} T_{1av} & RH_{1av} & T_{1max} & TP_{1av} & RHP_{1av} & M_1 & T_1 & W_1 & Th_1 & R_1 \\ T_{2av} & RH_{2av} & T_{2max} & TP_{2av} & RHP_{2av} & M_2 & T_2 & W_2 & Th_2 & R_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ T_{Nav} & RH_{Nav} & T_{Nmax} & TP_{Nav} & RHP_{Nav} & M_N & T_N & W_N & Th_N & R_N \end{bmatrix}. \quad (10)$$

The method described thus far was applied to a minimum of 12 months’ data leading up to the commencement of the trial at every building, thus forming a full seasonal cycle calibration dataset. Each building’s calibration database was closed immediately prior to it entering the study. Subsequently recorded $XO'$ and $EU$ trial data was stored in a ‘validation’ set using identical methods to those described above.
Given \( \{EU_i, T_{iav}, RH_{iav}, T_{imax}, TP_{iav}, RHP_{iav}, M_i, T_i, W_i, Th_i, R_i\}_{i=1}^N \) for \( N \) days in the calibration set, each building’s model for predicting daily electricity usage under any given combination of operating conditions was able to take the form

\[
EU_i = \beta_1 T_{iav} + \beta_2 RH_{iav} + \beta_3 T_{imax} + \beta_4 TP_{iav} + \beta_5 RHP_{iav} + \beta_6 M_i + \beta_7 T_i + \beta_8 W_i + \\
\beta_9 Th_i + \beta_{10} R_i + \epsilon_i.
\]  

(11)

Stacked together, these equations yielded regression coefficients for each of the independent variables impacting each building

\[
EU_{PRED} = X'O'\beta + \epsilon
\]  

(12)

where \( \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_{10} \end{pmatrix} \) and \( \epsilon \) is the error term.

To generate a prediction for any given post-calibration day (i.e. any day during the study) the relevant day’s external data from the validation dataset \( XO' \) was processed using equation (12). This gives an expected daily electricity usage (kWh) for the period 04:30 – 18:00 based on patterns and relationships observed prior to the commencement of the feedback trial.

The difference between a trial day’s \( EU \) and \( EU_{PRED} \) is then expressed as a percentage

\[
\text{DIFFERENCE}(\%) = \left(1 - \frac{EU}{EU_{PRED}}\right) \times 100
\]  

(13)

and this difference between observed and predicted performance forms the basis of the feedback signal dispatched to the building’s FM.
9.1.3. **Processing and storage**

The automatic data capture, storage, processing and messaging functions were handled by a customised version of a commercial software application procured from an Australian company, Event Zero\(^7\). The application built for this study (which was called ‘*Pulse*’) was based on Event Zero’s Event Center platform which uses a Linux operating system customised specifically by Event Zero for virtual appliances. The base Linux distribution was Ubuntu 8.04 LTS, which in turn was based on the Linux 2.6 kernel.

The container for the *Pulse* application was Java 2 (Standard Edition) and the majority of the external application logic was implemented in Java Server Pages (JSP) with the following two additional languages:

- For message templating, the Apache Velocity language was used
  (http://velocity.apache.org)
- For mathematical formulae, the Drools language was used
  (http://www.drools.org).

*Pulse* ran as a web application and required network connectivity to a range of other servers and services in order to operate. The main components are summarised in Figure 9-3 below.

---

\(^7\) Event Zero provides real time distributed data aggregation and analysis systems for commercial clients.
Figure 9-3. Pulse system components.

Listed below is the sequence of key functions performed by the *Pulse* system:

- Source NEM12 electricity consumption data from 3rd party metering providers
- Source weather observation data from the Bureau of Meteorology
- Associate electricity consumption data and weather observation data with individual buildings using common date- and time-stamps
- Compare each building’s observed electricity consumption and predictions based on statistical models
- For any given day, identify comparable (‘like’) days from database (see below)
• Compose different types of messages based on performance variations (e.g. using more or less than predicted), including text and graphical content

• Dispatch bespoke feedback to FMs (or groups) using email

• Provide data export facilities for raw, analysed and composit data to facilitate post hoc analysis.

9.1.4. Email messages

Given the nature of their job, FMs are constantly on their ‘smart’ phones checking emails, text messages, and taking calls. At the time of the trials, email offered a range of advantages for communicating with building operators compared with web- and app-based media (and even with the subsequent explosion of communication technologies, this still remains the case)\(^8\). Because it is ‘pushed’ from the software application there is no action required from the recipient in order to get the message. The subject line gives recipients an indication of the message content, allowing them to determine its significance or importance with just a glance and without requiring any deliberate action. And content can be structured to provide the crucial information up-front and more detail embedded deeper within the message. It was therefore decided that email would be the most effective medium to partly satisfy the first requirement for effective energy performance feedback: that it be in the form of ‘simple … messages delivered consistently and frequently’. Furthermore, by structuring the subject line and content so that significant results attracted more attention than average, building operators were prompted to focus more careful scrutiny (and potentially take action) on days when it was most warranted.

\(^8\) At the time the study was being conducted (2011-12) all participants were carrying pre-2010 versions of BlackBerry\(^\text{TM}\) handsets that had good email features and reasonable graphics but very limited internet functionality compared with other smart phone options.
Following is a description of the “facilities manager” email template used in the trial.

**Subject**

<table>
<thead>
<tr>
<th>If within threshold:</th>
</tr>
</thead>
<tbody>
<tr>
<td>From: Investa <em>Pulse</em> (<a href="mailto:PULSE@investa.com.au">PULSE@investa.com.au</a>)</td>
</tr>
<tr>
<td>To: &lt;Facilities_Manager’s_Name&gt;@investa.com.au</td>
</tr>
<tr>
<td>Subject: &lt;BUILDING_ADDRESS&gt;’s electricity performance: average day yesterday</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>If outside threshold:</th>
</tr>
</thead>
<tbody>
<tr>
<td>From: Investa Pulse (<a href="mailto:PULSE@investa.com.au">PULSE@investa.com.au</a>)</td>
</tr>
<tr>
<td>To: &lt;Facilities_Manager’s_Name&gt;@investa.com.au</td>
</tr>
<tr>
<td>Subject: &lt;BUILDING_ADDRESS&gt;’s electricity: &lt;good/bad&gt; day yesterday</td>
</tr>
</tbody>
</table>

**Figure 9-4. Design of the feedback email header and subject line.**

This blunt ‘good / bad news’ headline was dispatched every morning at 07:30, a time when it was known the facilities managers would still have yesterday fresh in their minds and be prioritising activities for the day ahead. It was hoped a simple injunctive message would also help to address potential boomerang effects, much like the emoticons deployed successfully by Schultz et al. (2007). Messages were sent from a trusted domain and recipients were advised that they could ‘hit reply’ and their response would be received and any comments noted. The statistician’s threshold of significance, ±5%, was used to determine whether a recipient would receive a message saying yesterday’s performance was good/bad or average, i.e. if the difference calculated at equation (13) was less than five percent the recipient was advised that nothing significant had been observed the previous day. This attempted to filter out some of the noise that would otherwise be associated with modelling error (standard deviations varied between models but on average they were approximately six percent. The *Pulse* application did not have the capacity to support sophisticated threshold setting or calculation).
Message body

Other requirements for effective feedback design that were accommodated in the email message template included: references to goals / baseline standards; direct relevance to the recipients so as to promote investigation; and, building of positive motivational beliefs and self-esteem to support dialogue and promote experimentation. Goals were not explicitly set, or self-set by the participants, but a ‘continuous improvement’ objective is espoused within the culture of Investa (and most businesses), so any clear reference to baseline performance automatically provides a signal of success or failure. This baseline was referenced in the opening sentence and the improvement trend was communicated in the third.

If within threshold:

Yesterday $\text{<BUILDING_ADDRESS>’s}$ electricity consumption was as expected. Nearby buildings used $\text{<INSERT_A>}%$ $\text{<less/more>}$ than expected. Over the past month $\text{<BUILDING_ADDRESS>}$ has significantly beaten the prediction on $\text{<INSERT_B>}%$ of occasions.

If outside threshold:

$\text{<Great/Bad> news, yesterday <BUILDING_ADDRESS> used <DIFFERENCE(%)>}%$ $\text{<less/more>}$ energy than expected. Nearby buildings used $\text{<INSERT_A>}%$ $\text{<less/more>}$ than expected. Over the past month $\text{<BUILDING_ADDRESS>}$ has significantly beaten the prediction on $\text{<INSERT_B>}%$ of occasions.

Where:

$\text{INSERT_A} = \text{absolute value of } 1 - \sum \text{of consumption predicted from all buildings assigned to the same BoM weather station divided by } \sum \text{of actual consumption from all buildings assigned to the same BoM weather station} - 1$

$\text{INSERT_B} = \text{rolling 30 day count of when savings of greater than } 5\% \text{ were achieved divided by the number of workdays.}$

Figure 9-5. Design of the feedback email message text.
A graphical illustration of a comparable or ‘like’ day profile (upper panel in Figure 9-6) and statistical and graphical references to the performance of nearby buildings (second sentence, Figure 9-5, and lower panel in Figure 9-6) enabled visualisation of the effect of energy-saving actions. Both graphs have overlayed temperature data (for visual validation) and were intended to help build self-esteem and positive motivational beliefs by showing progress and results, as well as promoting investigation, dialogue and experimentation when outperformance of nearby buildings was noted.
NMI Data for yesterday and its closest Like Weather Day (<INSERT_DATE>)

*Yesterdays Load Profile and Comparison*

Utility meter comparison

![Utility meter comparison chart](chart.png)

Normalised electricity consumption*

![Normalised electricity consumption chart](chart.png)

* NMI data for nearby buildings. Legend removed to obscure building identities.

**Figure 9-6. Design of the feedback email message graphics.**
The ‘Like Weather Day’ is the day in the most recent rolling 12 month period on which $EU_{PRED}$ (equation 12) is nearest to yesterday’s $EU_{PRED}$. The plot (upper panel Figure 9-6) shows the actual electricity usage $EU$ and dry bulb temperature $T$ data for both days. There were no other constraints imposed on the selection of like days and this led to problems on a few occasions, particularly for buildings in Melbourne (noted for its capricious weather), when on some occasions cold weather ‘like days’ were identified for warm conditions and vice versa. The problem did not exist in Brisbane and Sydney trials because conditions were never cold enough to produce a U-shaped curved such as the one illustrated at Figure 9-2. Additional temperature-based criteria for ‘like day’ selection were introduced to eliminate this issue after the conclusion of the trial.

9.2. Social learning environment

As noted by Kahneman (2011) et al., feedback loops require a spark in order for them to be set in motion. But the risk of introducing an ‘ignition’ element to a study such as this—offering of incentives or rewards or assisting with ideas, for example—is that it might potentially confound the analysis of the feedback intervention itself. It was therefore considered important to introduce as few additional variables as possible and do as little as possible to alter the environment that the participants were exposed to, and yet still initiate action somehow. The solution was to leverage existing relationships between the participants. Various scholars recommend social network influences be integrated as part of a “feedback package” (see, Brown & Arens, 2012; Malone et al., 2011; Moezzi, Hammer, Goins, & Meier, 2014) and so the introduction of a framework for ‘social’ stimulus acted as a complementary measure to stimulate engagement with the feedback rather than a competing measure whose effects would need to be separately quantified.
Trials commenced in batches of between three and five buildings from within the same local weather station zone. Each of these batches became an ‘action learning group’ typically consisting of a small team (5-8 people focussed on 3-5 buildings) that met fortnightly for an hour to discuss recent performance and initiatives undertaken. Participants included the FM s with direct operational responsibility for the buildings, a senior facilities manager (typically the team supervisor), me (the investigator) and a graduate engineer from Investa’s sustainability and environment team. The FMs were simply asked to come prepared to ‘talk about something that happened’ during the previous fortnight. The graduate engineer typically brought along excerpts from the daily Pulse messages to encourage the discussion. There were no rules and the agenda was fluid. Any discussion topic, so long as it related somehow to the Pulse feedback, was welcomed.

Organisational ‘direction’ and ‘support’ were held constant throughout the entire project. Other than the fortnightly meetings and the Pulse messages themselves, there were no significant changes to the way Investa approached or managed energy matters. Furthermore, hierarchy was excluded from the process. While it is acknowledged that I personally held a senior position in the Investa management hierarchy, I did not have any line management authority over any of the participants (other than the graduate engineer). I had been actively engaged in energy management at Investa since 2004 and my role and responsibilities had been more-or-less static since 2008 – three years prior to the commencement of the study. Likewise, the presence of the senior facilities managers (managers of the FMs) was not anything out of the ordinary. People in those roles spend more than half of their working hours in the 3-5 buildings that are assigned to them, so their presence was familiar and unthreatening – in fact it would have seemed odd for them not to be part of the dialogue. The Head of Facilities Management (the person to
whom the senior facilities managers reported) did not attend meetings, did not receive *Pulse* messages and was deliberately excluded (with consent) from the entire process because of concerns that the presence of this more senior figure may have altered the meeting dynamics.

The overarching objective was to construct a trusting environment for learning and information exchange that drew on the concepts of social proof and the Theory of Reasoned Action discussed in chapter 6. While the FMs shared bonds and a common position within the organisational structure, they typically brought very different skill-sets and career experiences to the meetings. It was not uncommon to find backgrounds ranging from refrigeration mechanic to electrician to plumber to carpenter and ages ranging from late-twenties to mid-sixties. Likewise, participants came with a wide variety of personality types, values and perspectives. This supportive environment encouraged participants to challenge each other in a non-threatening way so that they could function as models for each other and subconsciously develop evaluation capabilities (“intelligent self-monitoring”) through a method of applied Social Learning Theory similar to that suggested by Sadler (1989) and discussed in chapter 5.
10. APPROACH TO EVALUATION

10.1. *Hypothesis is refutable*

The causal hypothesis presented in chapter 7 was structured so that it can satisfy the commonly accepted conditions for validation, i.e. that:

(1) variations between variables must be associated;

(2) the influence of the independent variable (presumed cause) must occur before the presumed effect (dependent variable); and

(3) the effect on the dependent variable cannot be produced by something other than the presumed cause (e.g. a third variable).

(Mayer, 1980)

To satisfy the first and third conditions of validation there must be an observable connection between a variation in operator feedback (the intervention) and a building’s electricity use (dependent variable). This required the introduction of measures to limit variations in the trial portfolio’s operating environment and exclusion from the analysis of any buildings that subsequently experienced significant changes to occupancy or equipment. No assumptions were made about interactions with the feedback (e.g., that FMs pay attention to messages; that they act in response to the messages; that their responses are observable, etc.) because we are not able to isolate (through quantitative means) which elements of the feedback have greatest influence on results even though their collective effects can be observed. This accords with the approach required by Popper in *The Logic of Scientific Discovery* that “Systems of theories are tested by deducing from them statements of a lesser level of universality” (Popper, 1972, p. 47).
Likewise, the second condition stated above can also be satisfied in the analysis: if the presumed cause leads to the presumed effect we should be able to observe a distinct change in energy use following the intervention.

Popper requires that a system’s “... logical form shall be such that it can be singled out, by means of empirical tests, in a negative sense: it must be possible for an empirical scientific system to be refuted by experience” (emphasis in original) (Popper, p. 42). In this study we can say that automated daily energy performance feedback delivered in a social-learning environment has no energy-saving effect if energy use in the trial buildings is not reduced relative to the trial buildings’ aggregated baseline.

10.2. Additional model

A second and independent statistical model was developed for predicting daily electricity use in addition to the regression model described in chapter 9. This was done for two reasons. First, there was a concern that some of the limitations with the regression model might impact confidence in the analysis. The use of RH (%) rather than an absolute measure of atmospheric moisture (such as mixing ratio in grams per kg of dry air), for example, was considered a potential weakness that could undermine confidence in the analysis of results due to its negative correlation with dry bulb temperature. Second, it was considered useful to have a means for calculating results that was independent of the model used for generating the daily feedback.

There are a variety of different techniques that are suitable for modelling the energy use of buildings (see Jump et al., 2013). The novel method described here uses Principal Component Analysis (PCA) to determine expected daily profiles based on a selection of Nearest Neighbours (k-NN). A major advantage of this method (subsequently deployed post-trial) is that it can generate an “expected profile” in addition to an expected
aggregate level of consumption. During the trial only the expected total daily electricity use was determined, together with identification of the most “similar day” in the recent past. This had the limitation of pointing the recipient to anomalies without giving any indication as to their cause (e.g. morning start-up, use of outside air, equipment scheduling, etc.). The PCA–k-NN approach was developed post-trial and only adopted in this study for analysis of results, not to enhance the quality / usefulness of the daily feedback – something it was subsequently employed to do post-trial.

Following is a brief summary of the PCA–k-NN method to explain how it differs from the regression approach described previously.

First, as with the regression method (eq. 5), a matrix $DXO$ (daily external observations) of dimension $N_{columns} \times 96$ was created for each day from 00:15 hrs (measurement, No. 1) to 24:00 hrs (measurement No. 96). Under the PCA–k-NN approach additional variables were introduced to account for seasonal harmonics $C = \cos \left( \frac{2\pi}{365} D_N \right)$ and $S = \sin \left( \frac{2\pi}{365} D_N \right)$

$$DXO = \begin{pmatrix} T_1 & RH_1 & P_1 & C_1 & S_1 & \cdots \\ T_2 & RH_2 & P_2 & C_2 & S_2 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ T_{96} & RH_{96} & P_{96} & C_{96} & S_{96} & \cdots \end{pmatrix}.$$

Columns $S$ and $C$ reflect the seasonal dependence of electricity usage over the course of a year. It is normal in Australian cities for energy use to be highest during the summer period (December – February) and lowest during the winter period (June – August). However, in some situations due to climate and equipment configurations, overall energy usage can rise during winter to provide heating. For this reason a second harmonic is
needed. Further, the amplitude of annual variations in $DXO$ varies with location and building typology.

The same matrix for electricity consumption data, $DEU$ (daily electricity usage), was used for both methods (eq. 6). Pre-processing procedures were identical for both methods with one important exception. Under the PCA–$k$-NN method $RH$ data was converted to actual moisture content ($H$) measured in grams of moisture per kilogram of dry air and added to the matrix $DXO$ (eq. 14) for subsequent calculations and in order to calculate enthalpy in kilojoules (kJ) per kg of wet air. Enthalpy ($E$) was also then added to the matrix $DXO$ (eq. 14).

Intra-daily data in the expanded $DXO$ matrix was then combined with the $DEU$ measurements to create each building’s model calibration database of external observations, $XO$. This was very similar to the $XO$ presented in equation (7); however, it was found that day-of-the-week dummy variables did not contribute significantly to the model’s predictive power so they were excluded. Likewise, it was found that separate variables for overnight temperature and humidity were not necessary. The simplified XO matrix took the form

$$XO = \begin{pmatrix}
T_{1av} & H_{1av} & E_{1av} & C_{1av} & S_{1av} & \cdots \\
T_{2av} & H_{2av} & E_{2av} & C_{2av} & S_{2av} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
T_{Nav} & H_{Nav} & E_{Nav} & C_{Nav} & S_{Nav} & \cdots
\end{pmatrix}$$

(15)

And in exactly the same process as that used for the regression method, electricity usage was summed for the core operating period in each day to create a vector for each building’s total daily electricity usage, $EU$ (eq. 8).
Having established $XO$ and $EU$, the next step was to introduce the PCA. Figure 10-1 shows a typical set of daily profiles for electrical energy demand recorded over two years at one of the typical buildings in the portfolio. As can be seen, there is a general pattern, but the level of use varies significantly between days.

![Figure 10-1. Daily electricity demand profiles recorded over a two year period for one of the typical office buildings in a temperate climate.](image)

The profiles of daily ambient temperature ($T$), moisture content ($H$) and enthalpy ($E$) also form curves that take on consistent diurnal patterns that reflect the weather in a particular location. Therefore, we can form linear combinations from columns in $XO$ to account for the relationships. The first three columns $T_{av}$, $H_{av}$ and $E_{av}$ are shown here to illustrate:

$$\mathbb{R} = a * T_{av} + b * H_{av} + c * E_{av} = EU_{av} * \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$  \hspace{1cm} (16)
The above combinations are determined by the constant coefficients a, b and c, and accordingly, there are just 3 coefficients a, b and c, which determine the matrix

\[ EU_{av} = [T_{av} H_{av} E_{av}] \]  

(17)

PCA is a mathematical procedure that converts a set of observations of possibly correlated variables into a reduced set of linearly uncorrelated (orthogonal) variables called principal components. PCA allows us to identify the most “informative” combination, i.e. that containing the most variance exhibited by the parameters. The number of principal components is less than or equal to the number of original variables with the first principal component having the largest possible variance (that is, accounts for as much of the variability in the data as possible), and each succeeding component in turn having the highest variance possible under the constraint that it be orthogonal to the preceding components. PCA can thus be applied to the observed data in order to construct a set of uncorrelated principal components which embody the largest variance (i.e. significance) of the observations.

In the MATLAB programming language and environment, the command for applying PCA processing to the vector \( EU_{av} \) (eq. 17) is written as \([\text{coeff}, \text{score}, \text{latent}] = \text{pca}(EU_{av})\) (for a detailed explanation see Mathworks, 2015). Each column of the matrix 'coeff' \((N_{columns} \times N_{columns})\) contains principal component coefficients formed from \(XO\) (eq. 15). The vector (i.e. column) 'latent' contains variances associated with each combination depicted in equation (16) in descending order and 'score' returns the principal component scores. Note that PCA combinations are orthogonal and diagonalised, i.e.:

\[ \sum_{j=1}^{3} \text{coeff}(i,j) \times \text{coeff}(j,k) = \delta_{ik} = \begin{cases} 1, & \text{if } i = k \\ 0, & \text{if } i \neq k \end{cases} \]  

(18)
Accordingly, all the external data (matrix $XO$) can be represented to a high degree of accuracy by just one or two daily values (principal component values), with the first principal component in these analyses typically containing 95-98% of all variance and the second one typically containing 2-4%. Accordingly, just one or two principal components are required and the other components can be regarded as noise.

It was found with the regression method described in chapter 9 that the relationship between a building’s energy use and environmental conditions is normally non-linear and when a non-linear element is introduced to achieve a better fit (e.g. a residual), overfitting, or “tuning” can occur, potentially leading to false predictions, particularly at the tails of the distribution. This was one of the primary reasons for exploring PCA as a more reliable basis for generating daily feedback. Under this alternative PCA method, daily energy use is predicted by comparing principal component values with the building response (vector $EU_{av}$).

The following sum of additive terms was formed, each comprising a principal component coefficient and a time-dependent value of an associated external parameter:

$$PC_{cal1}(t) = coeff(1,1) \cdot T_{cal}(t) + coeff(2,1) \cdot H_{cal}(t) + coeff(3,1) \cdot E_{cal}(t) +$$
$$+ coeff(4,1) \cdot C_{cal} + coeff(5,1) \cdot S_{cal} + \cdots$$

(19)

where $PC_{cal1}$ means the first principal component value and $t$ is the time or number of the measurement (1 to 96).

Note that equation (19) presents the first principal component values only. Other principal component values can be easily calculated by replacing the second index 1 in matrix 'coeff' by 2, 3, etc.
A $k$-NN algorithm was developed which overcomes problems with linear and non-linear regression. This method takes $XO$ values measured in 15 minute intervals for any given day during the study period (or any day not used to generate the calibration set) and identifies a number of comparable or "like days" from the database (calibration set) where $XO$ values are closely matched. Matches are determined by comparing principal component curves and selecting for minimum difference and it has been found that five days ($k = 5$) is generally an appropriate number, as illustrated in Figure 10-2 below.

![Figure 10-2. Illustration of the k-NN method with electricity consumption data for 5 historic calibration 'like days' and the predicted and actual data for the specified validation day.](image)

Having identified ‘like days’, a prediction for the validation day (e.g. a day in the study period, or any other day not used for model calibration) may be calculated by performing a simple averaging, also illustrated in Figure 10-2 above. The area underneath the average curve represents the predicted daily electricity consumption which can then be compared with measured consumption to determine a saving or increase relative to the baseline.
This method was applied to exactly the same calibration data sets as the models used to
generate the feedback messages use in the study and, therefore, provides an independent
means of verifying results.

10.3. **Statistical tests and reproducibility**

All analyses of results were performed using the open-source R language for statistical
computing version 3.3.1 (R Core Team, 2016) and all files necessary to reproduce the test
results are available in .txt, .R and .csv formats.
PART C: RESULTS

11. SUMMARY OF RESULTS

11.1. Summary of trial

Thirty commercial office buildings from Melbourne, Vic (4), Sydney, NSW (19) and Brisbane, Qld (7) were initially included in the study. Each city is regarded as climatically dissimilar from a building services (heating and cooling) point of view, despite Sydney and Brisbane both being classified as ‘humid sub-tropical’ (Cfa) under the Köppen–Geiger climate classification system and Melbourne being in the slightly cooler ‘temperate oceanic’ (Cfb) classification (see Peel et al., 2007). The spread in mean monthly temperatures for the three cities is illustrated in Figure 11-1, below.

![Figure 11-1. Mean monthly maximum and minimum temperatures from official weather stations nearest to the participating buildings (Bureau of Meteorology, 2017).]
All thirty buildings are commercial office buildings with conventional HVAC systems (for large office buildings in Australian capital cities) comprised of centralised mechanical plant, electric chillers, water-cooled condenser systems (i.e. cooling towers), Variable Air Volume (VAV) distribution systems, electric duct heaters and, in the Sydney and Melbourne buildings, gas-fired boilers for primary space heating. All buildings had relatively modern BMCS and had at least one Investa-employed building manager onsite most of the time. At the time of the study the average building age was approximately 25 years and the average net lettable area (NLA) was 27,295m².

Seven of the initial cohort had to be removed from the analysis because significant occupancy or technology changes occurred throughout the study period, disrupting the research design. All buildings in Investa’s portfolio which had sufficient historical consumption data for the generation of models and stable operation during the model calibration period were included. Since these were fully-operational office buildings it was not possible to prevent technology or occupancy changes, so the loss of only seven from thirty over the course of a fifteen month study was considered a good result. Another building (BID0026) was removed from the consolidated analysis because it joined the study many months after the other twenty-two commenced and its model did not share any common calibration days with the others. Despite it having stable occupancy and no technology changes, there was a concern that this building’s results may have been influenced differently on account of the significant commencement lag. The removal of the eight buildings was expected to have no net effect on results overall as impacts on energy use were in both directions. The average NLA of the twenty-two
buildings included in the final set used for consolidated analysis was 28,148m² and the average age was approximately 25 years old.

Models were calibrated (or “trained”) using data from the period 1 June, 2010 – 22 December, 2011 and the trial period ranged from 15 August, 2011 – 30 November, 2012\(^9\), with each calibration window selected to be the nearest possible reflection of each building’s operation at the commencement of the study. The portfolio was divided into six groups of buildings based on spatial proximity. Trials commenced progressively in the following order, with a pre-commencement meeting convened for each, followed by the regular fortnightly meeting.

Table 11-1. Summary of buildings included in the trial.

<table>
<thead>
<tr>
<th>Group</th>
<th>Building Id.</th>
<th>NLA (^10) (m²)</th>
<th>Location (State)</th>
<th>Model training period</th>
<th>Trial period</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Start</td>
<td>Stop</td>
</tr>
<tr>
<td>A</td>
<td>BID0001</td>
<td>30,000</td>
<td>NSW</td>
<td>10-Dec-10</td>
<td>10-Aug-11</td>
</tr>
<tr>
<td></td>
<td>BID0022</td>
<td>25,000</td>
<td>NSW</td>
<td>29-Nov-10</td>
<td>10-Aug-11</td>
</tr>
<tr>
<td></td>
<td>BID0033</td>
<td>23,000</td>
<td>NSW</td>
<td>11-Aug-10</td>
<td>10-Aug-11</td>
</tr>
<tr>
<td>B</td>
<td>BID0017</td>
<td>28,000</td>
<td>NSW</td>
<td>01-Jun-10</td>
<td>05-Sep-11</td>
</tr>
<tr>
<td></td>
<td>BID0019</td>
<td>18,000</td>
<td>NSW</td>
<td>06-Oct-10</td>
<td>05-Sep-11</td>
</tr>
<tr>
<td></td>
<td>BID0021</td>
<td>22,000</td>
<td>NSW</td>
<td>06-Sep-10</td>
<td>05-Sep-11</td>
</tr>
<tr>
<td>C</td>
<td>BID0007</td>
<td>8,000</td>
<td>NSW</td>
<td>06-Oct-10</td>
<td>05-Oct-11</td>
</tr>
<tr>
<td></td>
<td>BID0038</td>
<td>15,000</td>
<td>NSW</td>
<td>06-Oct-10</td>
<td>05-Oct-11</td>
</tr>
<tr>
<td></td>
<td>BID0039</td>
<td>14,000</td>
<td>NSW</td>
<td>10-Nov-10</td>
<td>05-Oct-11</td>
</tr>
<tr>
<td></td>
<td>BID0041</td>
<td>19,000</td>
<td>NSW</td>
<td>29-Oct-10</td>
<td>28-Oct-11</td>
</tr>
<tr>
<td></td>
<td>BID0047</td>
<td>27,000</td>
<td>NSW</td>
<td>28-Oct-10</td>
<td>28-Oct-11</td>
</tr>
<tr>
<td></td>
<td>BID0006</td>
<td>10,000</td>
<td>NSW</td>
<td>09-May-11</td>
<td>15-May-12</td>
</tr>
<tr>
<td></td>
<td>BID0048</td>
<td>17,000</td>
<td>NSW</td>
<td>03-Nov-10</td>
<td>02-Nov-11</td>
</tr>
<tr>
<td></td>
<td>BID0043</td>
<td>26,000</td>
<td>NSW</td>
<td>04-Jan-11</td>
<td>28-Oct-11</td>
</tr>
<tr>
<td>E</td>
<td>BID0004</td>
<td>65,000</td>
<td>VIC</td>
<td>29-Nov-10</td>
<td>28-Nov-11</td>
</tr>
<tr>
<td></td>
<td>BID0015</td>
<td>65,000</td>
<td>VIC</td>
<td>13-Jan-11</td>
<td>28-Nov-11</td>
</tr>
<tr>
<td></td>
<td>BID0034</td>
<td>32,000</td>
<td>VIC</td>
<td>07-Jan-11</td>
<td>29-Nov-11</td>
</tr>
<tr>
<td></td>
<td>BID0042</td>
<td>25,000</td>
<td>VIC</td>
<td>01-Feb-11</td>
<td>25-Nov-11</td>
</tr>
</tbody>
</table>

\(^9\) With the exception of BID0026 which was a late inclusion with calibration period 4/01/2012 to 27/07/2012 and trial limited to 1/08/2012 to 30/11/2012.

\(^10\) Building size has been rounded to the nearest 1,000m² to mask identities.
For the reasons discussed in chapter 10, the models developed using the PCA–$k$-NN method are regarded as more suitable than the regression models for evaluating the treatment effect (the effect of the feedback and ‘social learning’ environment). The standard deviations for the two sets of models based on the calibration data are presented in Table 11-2 below.

### Table 11-2. Standard deviations of the two models.

<table>
<thead>
<tr>
<th>BID</th>
<th>State</th>
<th>Std Dev Regression</th>
<th>Std Dev PCA–$k$-NN</th>
</tr>
</thead>
<tbody>
<tr>
<td>BID0001</td>
<td>NSW</td>
<td>0.085078</td>
<td>0.071756</td>
</tr>
<tr>
<td>BID0047</td>
<td>NSW</td>
<td>0.093737</td>
<td>0.098454</td>
</tr>
<tr>
<td>BID0041</td>
<td>NSW</td>
<td>0.101296</td>
<td>0.093828</td>
</tr>
<tr>
<td>BID0005</td>
<td>NSW</td>
<td>0.065010</td>
<td>0.073483</td>
</tr>
<tr>
<td>BID0048</td>
<td>NSW</td>
<td>0.050314</td>
<td>0.051961</td>
</tr>
<tr>
<td>BID0017</td>
<td>NSW</td>
<td>0.098388</td>
<td>0.087033</td>
</tr>
<tr>
<td>BID0019</td>
<td>NSW</td>
<td>0.104537</td>
<td>0.101316</td>
</tr>
<tr>
<td>BID0021</td>
<td>NSW</td>
<td>0.096054</td>
<td>0.089795</td>
</tr>
<tr>
<td>BID0022</td>
<td>NSW</td>
<td>0.085832</td>
<td>0.081279</td>
</tr>
<tr>
<td>BID0043</td>
<td>NSW</td>
<td>0.098365</td>
<td>0.091284</td>
</tr>
<tr>
<td>BID0026</td>
<td>NSW</td>
<td>0.067925</td>
<td>0.074869</td>
</tr>
<tr>
<td>BID0033</td>
<td>NSW</td>
<td>0.069600</td>
<td>0.067125</td>
</tr>
<tr>
<td>BID0035</td>
<td>NSW</td>
<td>0.114064</td>
<td>0.106819</td>
</tr>
<tr>
<td>BID0038</td>
<td>NSW</td>
<td>0.119625</td>
<td>0.095217</td>
</tr>
<tr>
<td>BID0039</td>
<td>NSW</td>
<td>0.090887</td>
<td>0.084656</td>
</tr>
<tr>
<td>BID0044</td>
<td>QLD</td>
<td>0.043023</td>
<td>0.043817</td>
</tr>
<tr>
<td>BID0046</td>
<td>QLD</td>
<td>0.057666</td>
<td>0.058995</td>
</tr>
<tr>
<td>BID0050</td>
<td>QLD</td>
<td>0.045699</td>
<td>0.049750</td>
</tr>
<tr>
<td>Building Code</td>
<td>State</td>
<td>Mean Standard Deviation</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td>---------------</td>
<td>-------</td>
<td>-------------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>BID0051</td>
<td>QLD</td>
<td>0.039850</td>
<td>0.045600</td>
</tr>
<tr>
<td>BID0004</td>
<td>VIC</td>
<td>0.073032</td>
<td>0.060582</td>
</tr>
<tr>
<td>BID0015</td>
<td>VIC</td>
<td>0.065650</td>
<td>0.074126</td>
</tr>
<tr>
<td>BID0034</td>
<td>VIC</td>
<td>0.066551</td>
<td>0.059110</td>
</tr>
<tr>
<td>BID0042</td>
<td>VIC</td>
<td>0.098828</td>
<td>0.089083</td>
</tr>
<tr>
<td><strong>Mean:</strong></td>
<td></td>
<td><strong>0.079609</strong></td>
<td><strong>0.076084</strong></td>
</tr>
</tbody>
</table>

In some cases multiple regression produced a lower standard deviation than the PCA–$k$-NN method. This indicates that in some cases the regression approach produced a better “fit” for the training period; however, it does not necessarily indicate it has better predictive power or validity for evaluating the trial at any of the buildings.

The ability of both sets of models to explain the variability of electricity consumption during the pre-trial calibration period is illustrated in Figure 11-2, below.

![Figure 11-2](image_url)

**Figure 11-2.** Mean ‘work day’ electricity consumption and model predictions for all buildings in the analysis set during calibration.
Figure 11-2 shows the mean daily electricity consumption over 160 calibration days for all twenty-two buildings in the analysis set (the number of ‘work days’ used to calibrate models ranged from 161 to 310) together with the mean predictions from the two modelling methods. It is clear that both methods produced a good “fit”: the standard deviation for the regression models shown in Figure 11-2 is 2.02% whereas for the PCA–*k*-NN models it is 1.85%.

Figure 11-3 shows the mean ‘workday’ electricity consumption throughout the first 220 days of each building’s trial (length of trial ranged between 228 and 310 days) together with the mean prediction from the PCA–*k*-NN models. The grey shaded area indicates the 95% confidence band (2 standard deviations).

Figure 11-3. Mean ‘work day’ electricity consumption and predictions for all buildings in the analysis set during the first 220 days of the trial.
A significant divergence between actual and ‘predicted’ daily electricity consumption began to appear after about 50 days of the treatment. This can be seen more clearly on the scale in Figure 11-4, below.

![Figure 11-4](image)

**Figure 11-4.** Mean ‘work day’ electricity consumption and predictions for all buildings in the analysis set during the first 220 days of the trial.

This divergence is also clearly illustrated in Figure 11-5 which shows daily variance between actual and predicted electricity consumption expressed as a percentage for all twenty-two buildings in the analysis set during their first 220 days of the intervention. Note the x-axis is identical for Figure 11-3, Figure 11-4 and Figure 11-5 with “Day 0” the trial start for each individual building in the sample.
Figure 11-5. Variance between ‘predicted’ and actual ‘work day’ electricity consumption.

As each of the above figures illustrates, energy savings appear to have become significant from about day ~50 and were sustained throughout the remainder of the treatment period.

11.2. Regression discontinuity design

Regression discontinuity design (RDD) has become a popular method for testing causal hypotheses, particularly in natural experiments where potential confounders are numerous and where subjects cannot be separated into control and treatment groups (Imbens & Lemieux, 2008). RDD was first introduced to the literature by Thistlethwaite and Campbell in a study of the effect of scholarship awards on students’ attitudes and career plans (Thistlethwaite & Campbell, 1960). They recognised that there was generally very little difference in the merit and abilities of students on either side of the award cut-off threshold and developed RDD to estimate the treatment effect by examining data for subjects immediately on either side of it.
Under the RDD approach, observation data are assigned to treatment or pre-treatment groups by introducing a binary predictor – in this case “0” for “no intervention” (i.e. for days belonging to the calibration set) and “1” for days after the commencement of the trial. Variance between actual and predicted energy use is assumed to be unaffected by the binary predictor and not a function of it (i.e. the null hypothesis predicts the treatment will have no effect), and so any discontinuity at the cut-off is interpreted as evidence of a causal effect of the treatment (Imbens & Lemieux, 2008).

It can be observed from Figure 11-6, below, that energy use was stable prior to the intervention, as indicated by the (-0.0013) slope of the black dashed line of best fit. This is in stark contrast to the improvement trend observed after the feedback and social-learning environment was introduced (gradient = -0.0332). The vertical dashed line marks the commencement of the intervention.

![Figure 11-6. Regression discontinuity design (RDD) showing improvement trend prior and post intervention (vertical dashed line).]
The homoscedasticity of the two groups’ sample variances was confirmed using a Fisher’s $F$-test as the calculated $F = 0.2248$ is less than the tabulated $F = 1.2716$ with numerator df = 159 and denominator df = 219.

The mean variance for the calibration set was -0.95%. For the trial it was -7.32% noting, however, that the variance at the conclusion of the trial was greater than the mean, as illustrated in Figure 11-6, with a rate of reduction in energy use of approximately one percent every 30 days.

Because the trial variance is obviously skewed, it is not appropriate to compare the means of the two groups using a Student’s $t$-test. Instead, a non-parametrical equivalent of the Student’s $t$-test for independent samples such as the Wilcoxon-Mann-Whitney rank sum test is required. A Wilcoxon-Mann-Whitney rank sum test was performed in R (R Core Team, 2016) and a significant effect was found ($U = 33044, Z = 14.6089, p\text{-}value < 2.2e-16, d = 1.79$). An effect size indicated by a Cohen’s $d$ of greater than 0.8 is considered “large” (Cohen, 1988) and as the p-value is less than 0.05, we can reject the null hypothesis that the means of the two groups are statistically equal.

11.3. Individual building results

The results varied between individual buildings, of course. Figure 11-7 presents a histogram of the improvement trends observed in the twenty-two trial buildings prior to commencement. The distribution was approximately normal ($M = -0.0013, s = 0.0185$).
Figure 11-7. Distribution of calibration set gradients.

Figure 11-8 presents a histogram of the average improvement trends observed during the trial. This distribution was also approximately normal ($M = -0.0332, s = 0.0356$).

Figure 11-8. Distribution of gradients for the trial period.
Table 11-3 and Figure 11-9 show for each individual building the magnitude and trend of variances between actual and predicted energy use during the calibration and trial periods derived from the RDD. The excluded building (BID0026) is presented in bold font.

**Table 11-3. Trends for individual buildings, prior to and during the trial.**

<table>
<thead>
<tr>
<th>BID</th>
<th>State</th>
<th>Calibration gradient</th>
<th>Trial gradient</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>BID0001</td>
<td>NSW</td>
<td>0.01644</td>
<td>-0.01813</td>
</tr>
<tr>
<td>2</td>
<td>BID0047</td>
<td>NSW</td>
<td>-0.00396</td>
<td>-0.00179</td>
</tr>
<tr>
<td>3</td>
<td>BID0041</td>
<td>NSW</td>
<td>-0.00686</td>
<td>0.00641</td>
</tr>
<tr>
<td>4</td>
<td>BID0004</td>
<td>VIC</td>
<td>0.01301</td>
<td>-0.07137</td>
</tr>
<tr>
<td>5</td>
<td>BID0005</td>
<td>NSW</td>
<td>-0.00555</td>
<td>-0.03683</td>
</tr>
<tr>
<td>6</td>
<td>BID0048</td>
<td>NSW</td>
<td>0.00210</td>
<td>-0.13315</td>
</tr>
<tr>
<td>7</td>
<td>BID0044</td>
<td>QLD</td>
<td>-0.01926</td>
<td>0.00794</td>
</tr>
<tr>
<td>8</td>
<td>BID0015</td>
<td>VIC</td>
<td>0.03492</td>
<td>-0.03792</td>
</tr>
<tr>
<td>9</td>
<td>BID0017</td>
<td>NSW</td>
<td>-0.01351</td>
<td>-0.03231</td>
</tr>
<tr>
<td>10</td>
<td>BID0019</td>
<td>NSW</td>
<td>0.02833</td>
<td>-0.07782</td>
</tr>
<tr>
<td>11</td>
<td>BID0046</td>
<td>QLD</td>
<td>0.00937</td>
<td>-0.02621</td>
</tr>
<tr>
<td>12</td>
<td>BID0021</td>
<td>NSW</td>
<td>0.03835</td>
<td>-0.05741</td>
</tr>
<tr>
<td>13</td>
<td>BID0022</td>
<td>NSW</td>
<td>-0.00869</td>
<td>-0.05821</td>
</tr>
<tr>
<td>14</td>
<td>BID0043</td>
<td>NSW</td>
<td>-0.05353</td>
<td>-0.10541</td>
</tr>
<tr>
<td>15</td>
<td>BID0050</td>
<td>QLD</td>
<td>0.01755</td>
<td>-0.04203</td>
</tr>
<tr>
<td>16</td>
<td>BID0051</td>
<td>QLD</td>
<td>0.00544</td>
<td>-0.00952</td>
</tr>
<tr>
<td>17</td>
<td>BID0026</td>
<td>NSW</td>
<td>-0.03567</td>
<td>-0.07322</td>
</tr>
<tr>
<td>18</td>
<td>BID0033</td>
<td>NSW</td>
<td>0.01384</td>
<td>-0.01231</td>
</tr>
<tr>
<td>19</td>
<td>BID0034</td>
<td>VIC</td>
<td>0.03062</td>
<td>-0.03070</td>
</tr>
<tr>
<td>20</td>
<td>BID0042</td>
<td>VIC</td>
<td>-0.04232</td>
<td>-0.04238</td>
</tr>
<tr>
<td>21</td>
<td>BID0035</td>
<td>NSW</td>
<td>-0.01449</td>
<td>-0.00963</td>
</tr>
<tr>
<td>22</td>
<td>BID0038</td>
<td>NSW</td>
<td>-0.01287</td>
<td>0.07928</td>
</tr>
<tr>
<td>23</td>
<td>BID0039</td>
<td>NSW</td>
<td>-0.02387</td>
<td>0.02008</td>
</tr>
</tbody>
</table>
Figure 11-9. Regression discontinuity design for each building in the analysis (and also BID0026).
Table 11-4 shows the standard deviation for each building’s model in calibration compared to the mean improvement observed during the trial. The twelve buildings where improvement exceeded one standard deviation are italicised. The excluded building (BID0026) is presented in bold font.

<table>
<thead>
<tr>
<th>BID</th>
<th>State</th>
<th>Std_Dev</th>
<th>Mean_Imp</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>BID0001</td>
<td>NSW</td>
<td>7.18%</td>
<td>-7.85%</td>
<td>-0.68%</td>
</tr>
<tr>
<td>BID0047</td>
<td>NSW</td>
<td>9.85%</td>
<td>-5.42%</td>
<td>4.43%</td>
</tr>
<tr>
<td>BID0041</td>
<td>NSW</td>
<td>9.38%</td>
<td>-10.77%</td>
<td>-1.39%</td>
</tr>
<tr>
<td>BID0005</td>
<td>NSW</td>
<td>7.35%</td>
<td>-2.86%</td>
<td>4.49%</td>
</tr>
<tr>
<td>BID0048</td>
<td>NSW</td>
<td>5.20%</td>
<td>-20.44%</td>
<td>-15.25%</td>
</tr>
<tr>
<td>BID0017</td>
<td>NSW</td>
<td>8.70%</td>
<td>-10.42%</td>
<td>-1.72%</td>
</tr>
<tr>
<td>BID0019</td>
<td>NSW</td>
<td>10.13%</td>
<td>-12.61%</td>
<td>-2.48%</td>
</tr>
<tr>
<td>BID0021</td>
<td>NSW</td>
<td>8.98%</td>
<td>-5.92%</td>
<td>3.06%</td>
</tr>
<tr>
<td>BID0022</td>
<td>NSW</td>
<td>8.13%</td>
<td>-5.36%</td>
<td>2.76%</td>
</tr>
<tr>
<td>BID0043</td>
<td>NSW</td>
<td>9.13%</td>
<td>-26.56%</td>
<td>-17.43%</td>
</tr>
<tr>
<td>BID0026</td>
<td>NSW</td>
<td>7.49%</td>
<td>-0.60%</td>
<td>6.89%</td>
</tr>
<tr>
<td>BID0033</td>
<td>NSW</td>
<td>6.71%</td>
<td>-8.43%</td>
<td>-1.71%</td>
</tr>
<tr>
<td>BID0035</td>
<td>NSW</td>
<td>10.68%</td>
<td>-11.44%</td>
<td>-0.75%</td>
</tr>
<tr>
<td>BID0038</td>
<td>NSW</td>
<td>9.52%</td>
<td>4.01%</td>
<td>13.53%</td>
</tr>
<tr>
<td>BID0039</td>
<td>NSW</td>
<td>8.47%</td>
<td>-4.96%</td>
<td>3.51%</td>
</tr>
<tr>
<td>BID0044</td>
<td>QLD</td>
<td>4.38%</td>
<td>-6.47%</td>
<td>-2.09%</td>
</tr>
<tr>
<td>BID0046</td>
<td>QLD</td>
<td>5.90%</td>
<td>-3.17%</td>
<td>2.73%</td>
</tr>
<tr>
<td>BID0050</td>
<td>QLD</td>
<td>4.98%</td>
<td>-6.22%</td>
<td>-1.24%</td>
</tr>
<tr>
<td>BID0051</td>
<td>QLD</td>
<td>4.56%</td>
<td>-5.21%</td>
<td>-0.65%</td>
</tr>
<tr>
<td>BID0004</td>
<td>VIC</td>
<td>6.06%</td>
<td>-7.14%</td>
<td>-1.08%</td>
</tr>
<tr>
<td>BID0015</td>
<td>VIC</td>
<td>7.41%</td>
<td>2.85%</td>
<td>10.26%</td>
</tr>
<tr>
<td>BID0034</td>
<td>VIC</td>
<td>5.91%</td>
<td>0.44%</td>
<td>6.35%</td>
</tr>
<tr>
<td>BID0042</td>
<td>VIC</td>
<td>8.91%</td>
<td>1.31%</td>
<td>10.22%</td>
</tr>
</tbody>
</table>

The pattern of improvement observed at each of the trial buildings is illustrated in Figure 11-10, below.
Figure 11-10. Pattern of improvement at each of the trial buildings in the analysis (and also BID0026).
PART D: DISCUSSION & CONCLUSIONS

12. COMMENTS ON THE INTERVENTIONS AND SIGNIFICANCE OF THE STUDY

12.1. Significance of the study

This study, which has involved the operators of twenty-two commercial office buildings that were unchanged in terms of their building services technology and equipment for a period exceeding 220 business days, is longer in duration and larger in number of participants than any reported in the literature previously (see Lopes et al., 2012, for a thorough review to the end of 2011). No other studies reported in the literature have used a feedback generation and delivery methodology of the type developed for this study and the social learning context which was adopted to support the participants, while based on established research and experience from other contexts, is a unique application.

Given Australia has over 22 million square metres of commercial office accommodation (PCA, 2009), it is conceivable that there might be some thousands of office buildings in Australia that directly compare to those included in this analysis and for which the results may be directly relevant. And noting that Australia’s GDP represents just 2.35 percent of the global economy (World Bank, 2016), the number and floor area of comparable buildings around the world may be up to 40 times larger.

The results reported in the previous chapter are, therefore, highly significant. The observed reduction in operating energy use, which by the end of the analysis period had reached approximately 10 percent on average, is consistent with those reported from
studies of feedback interventions in the residential sector (Darby, 2006; Delmas et al., 2013; Faruqui et al., 2010), noting that the extensive literature from residential studies shows a wide range.

It is typical in the literature for building typologies to be categorised as being either ‘residential’ in nature or ‘non-residential’ or ‘commercial’ (Levine et al., 2007; Lucon et al., 2014). The similarity in results may suggest that a distinction is not necessary when it comes to quantifying the energy saving possibilities from buildings: similar results may reasonably be expected from all building typologies. More research will be required to establish whether this is in fact the case and whether there is consistency across other types of non-residential buildings in addition to offices. The policy implications will be significant if there is. As noted earlier; buildings account for approximately 51 percent of global electricity consumption, 32 percent of total global final energy use, and 19 percent of energy-related greenhouse gas emissions (Lucon et al., 2014). And, in 2010, non-residential buildings (all types, not just offices) accounted for one quarter of the 32.7 PWh of energy used by buildings globally (IEA, 2013, cited in Lucon et al., 2014).

12.2. Significant results

The results presented in the previous chapter give a clear indication that the intervention produced a significant reduction in energy use across the sample building portfolio. The Wilcoxon-Mann-Whitney rank sum test showed the intervention generated a large effect with Cohen’s $d$ of 1.79 (greater than 0.8 is considered a “large” effect size) and significant result with $p$-value $< 2.2e-16$ suggesting the null hypothesis can be rejected.
There are various other indicators of improvement that could be adopted. For example, of the twenty-three buildings\textsuperscript{11}:

- nineteen had an improvement trend over the course of the trial (i.e. energy use decreased as duration of exposure increased)
- seventeen had a steeper improvement trend during the trial than they had pre-trial
- based on a visual inspection of the variances presented in Figure 11-9, twenty-two showed a pattern of energy use during the trial that may be regarded as ‘better’ than that observed during the calibration period.

The only building that showed a clear increase in energy use was BID0038 (see Figure 11-9 and Figure 11-10). There is no obvious explanation for why BID0038’s results differed so markedly from the sample averages. One possible explanation is a change in the building’s FM that occurred mid-way through the trial period. Very few personnel changes occurred during the trial and the effect of changes was expected to be small based on the fact that senior facilities managers (FMs’ direct superiors) were unchanged, the transfer of knowledge was seamless and feedback was continuous. In the case of BID0038, the incoming facilities manager was given responsibility for that building on top of another nearby building which he had been responsible for since the beginning of the trial (BID0041). The intervention was initially associated with a reduction in energy use at BID0041, however performance appeared to slip approximately 100 days into the trial at both buildings when the facilities manager’s responsibilities expanded. Another factor that may have also intervened is the behaviour of the tenant. BID0038 is fully occupied by a single government department that exercised considerable influence over the operations of the building and was known to be dissatisfied with thermal

\textsuperscript{11} Includes BID0026 which was not part of the consolidated analysis.
environmental conditions. Furthermore, much of the tenancy’s supplementary HVAC equipment had been configured to conflict with the base building services (the landlord was powerless to intervene). A plausible explanation for the increase in energy use, therefore, is that the incoming FM may have placed more emphasis on addressing tenant concerns than his predecessor did, or the tenant may have changed its requirements in some way. Unfortunately, it is impossible to dissect the exact cause from the available data\textsuperscript{12}. Given the FM responded positively to the feedback and social learning environment when it was introduced at their first building, BID0041, the inconsistency of results observed at BID0038 can probably be put down to a confounding variable such as one (or a combination) of those mentioned above. And it is worth noting that due to the normal distribution of results, similar influences may have acted in the reverse direction for buildings at the other end of the spectrum.

\textit{12.3. Impact of removing eight buildings from the sample}

It was noted in chapter 11 that seven of the buildings in the initial trial cohort had to be removed from the analysis due to significant occupancy or technology changes that occurred throughout the study period. Another (BID0026) was removed because of its late inclusion and problems with its model not having any calibration days in common with the others. The removal of the eight from a sample of thirty buildings was not expected to significantly affect the overall results.

Buildings were removed for a variety of reasons, with the principal ones being significant changes to occupancy (both increases and decreases associated with leasing activity) and major equipment changes including the life-cycle replacement of chiller plant. The

\textsuperscript{12}Electricity consumed by the tenant’s supplementary HVAC equipment was separately metered and not included in the data monitored for this study; however, it was suggested by the FM that system conflicts imposed higher demand for base-building HVAC services.
excluded buildings were BID0006, BID0049, BID0007, BID0018, BID0024, BID0028 and BID0036 and their average NLA was 21,195m$^2$. The RDD scatter plots for all 30 buildings are presented in Appendix B where it can be observed that one of the excluded buildings (BID0049) trended towards significantly lower electricity usage over the course of the trial and the other six showed no clear or consistent changes (noting that the dominant influences were all attributed to factors other than the intervention itself).

Figure 12-1, below, presents the combined results for all twenty-eight buildings in which sufficient data were generated (BID0018 and BID0026 had to be excluded because of insufficient calibration and trial data respectively). As expected, Figure 12-1 closely resembles the RDD scatter of the study’s results presented in Figure 11-6. Unlike Figure 11-6, this figure suggests a slightly increasing usage trend (gradient = 0.0001) prior to the intervention. But like Figure 11-6, Figure 12-1 also shows a rapid improvement after the feedback and social-learning environment was introduced (gradient = -0.0273).

Figure 12-1. RDD showing trend for all 28 buildings in the study with sufficient data (none excluded). Vertical dashed line marks beginning of intervention.
A comparison of pre- and post-intervention gradients suggests a difference in variance over the course of the 220 day analysis period of slightly less than one percent and summary statistics indicate the result would be as reported in chapter 11 even without the exclusions. While it is desirable to have results based on a large sample size for a variety of reasons discussed earlier, this needs to be balanced with the need to achieve a high level of confidence in the causality of the intervention. The removal of eight buildings to permit a full analysis of a consistently high quality sample is therefore considered appropriate given the research objective and it would appear justified by the “large” effect size ($d > 0.8$) and statistically significant result ($p < 0.01$).

12.4. **Commonalities and differences between buildings and operators**

As discussed earlier, the buildings shared many common characteristics. They were all large CBD office buildings with multiple tenants, but the same owner, professionally managed and on a common management platform. Their owner was motivated to save energy and they had collectively achieved significant savings prior to the study. The ways in which these factors and the sample’s homogeneity may potentially impact the external validity of findings are discussed in the next chapter. In this section we will consider the consistencies and inconsistencies in results.

The average NABERS Energy Rating for all buildings in the study for the two years prior to commencement was 3.8 (Investa’s NABERS ratings cover the period April – March).

The average NABERS Energy rating$^{13}$ for the period April, 2011 – March, 2012 (spanning the pre-trial and initial stages) was 4.2. Table 12-1 shows the NABERS Energy

---

$^{13}$ It is common practice for real estate companies to report their portfolios’ *weighted average* NABERS ratings (by NLA) rather than simple averages. This combined with the differences in sample composition explains the inconsistency between average ratings shown here and those reported by Investa Property Group (2012).
ratings and trends for the periods leading up to and during the study. The stability of the sample’s energy use baseline is demonstrated by the stable ratings and calibration gradient (see also Figure 12-1, above); however there is clearly some within-sample variability.

As can be observed from Table 12-1, in one case a building (BID0015) showed a net negative energy usage trend (i.e. the gradient during the trial was more downward than during the calibration period) and yet its NABERS rating decreased by half a star (suggestive of higher energy use). In thee of the six buildings (BID0047, BID0044, BID0038) where the difference between calibration and trial gradients was positive (suggestive of increasing energy use), NABERS ratings improved. There are five potential explanations for this. The first is the misalignment of the NABERS rating period and the trial. The trend may have been upwards and then switched to downwards, but not sufficiently to offset the increase before the end of the rating period (and vice versa). The second explanation is related in that a building’s performance may have been marginally one side of a NABERS rating band and switched over to the other side by a very small increment (this advantage of the RDD approach was discussed in the previous chapter). The third is the NABERS rating methodology itself which takes into account record keeping for after-hours service requests etc. An increase in calls for after-hours HVAC services will improve a rating by virtue of an increase in operating hours but its effect has been excluded from this analysis by design because of their random / unpredictable nature. The fourth is the possible impact of variability in weather-related factors. NABERS does not normalise for the effect of weather whereas the methodology employed here does. The fifth is a possible fatigue effect in some buildings where performance may have improved initially and then drifted back over the course of the
trial. Refer to individual building RDD plots in Figure 11-9 for evidence of these possible explanations.

### Table 12-1. NABERS Energy ratings (0-6 star scale) and performance statistics prior to and during trial

<table>
<thead>
<tr>
<th>Group</th>
<th>Building Id.</th>
<th>NLA</th>
<th>State</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>Calibration gradient</th>
<th>Trial gradient</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>BID0048</td>
<td>17,000</td>
<td>NSW</td>
<td>–</td>
<td>0.0</td>
<td>1.5</td>
<td>0.002</td>
<td>-0.133</td>
<td>-0.135</td>
</tr>
<tr>
<td>B</td>
<td>BID0019</td>
<td>18,000</td>
<td>NSW</td>
<td>4.5</td>
<td>4.5</td>
<td>4.5</td>
<td>0.028</td>
<td>-0.078</td>
<td>-0.106</td>
</tr>
<tr>
<td>B</td>
<td>BID0021</td>
<td>22,000</td>
<td>NSW</td>
<td>4.5</td>
<td>4.5</td>
<td>4.5</td>
<td>0.038</td>
<td>-0.057</td>
<td>-0.096</td>
</tr>
<tr>
<td>E</td>
<td>BID0004</td>
<td>65,000</td>
<td>VIC</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>0.013</td>
<td>-0.071</td>
<td>-0.084</td>
</tr>
<tr>
<td>E</td>
<td>BID0015</td>
<td>65,000</td>
<td>VIC</td>
<td>3.5</td>
<td>4.0</td>
<td>3.5</td>
<td>0.035</td>
<td>-0.038</td>
<td>-0.073</td>
</tr>
<tr>
<td>E</td>
<td>BID0034</td>
<td>32,000</td>
<td>VIC</td>
<td>4.5</td>
<td>4.5</td>
<td>4.5</td>
<td>0.031</td>
<td>-0.031</td>
<td>-0.061</td>
</tr>
<tr>
<td>F</td>
<td>BID0050</td>
<td>10,000</td>
<td>QLD</td>
<td>3.0</td>
<td>3.0</td>
<td>4.0</td>
<td>0.018</td>
<td>-0.042</td>
<td>-0.060</td>
</tr>
<tr>
<td>D</td>
<td>BID0043</td>
<td>26,000</td>
<td>NSW</td>
<td>–</td>
<td>3.0</td>
<td>4.0</td>
<td>-0.054</td>
<td>-0.105</td>
<td>-0.052</td>
</tr>
<tr>
<td>A</td>
<td>BID0022</td>
<td>25,000</td>
<td>NSW</td>
<td>4.5</td>
<td>4.5</td>
<td>4.5</td>
<td>-0.009</td>
<td>-0.058</td>
<td>-0.050</td>
</tr>
<tr>
<td>F</td>
<td>BID0046</td>
<td>53,000</td>
<td>QLD</td>
<td>3.5</td>
<td>4.5</td>
<td>4.5</td>
<td>0.009</td>
<td>-0.026</td>
<td>-0.036</td>
</tr>
<tr>
<td>A</td>
<td>BID0001</td>
<td>30,000</td>
<td>NSW</td>
<td>4.0</td>
<td>4.5</td>
<td>4.5</td>
<td>0.016</td>
<td>-0.018</td>
<td>-0.035</td>
</tr>
<tr>
<td>D</td>
<td>BID0005</td>
<td>42,000</td>
<td>NSW</td>
<td>4.5</td>
<td>4.5</td>
<td>4.5</td>
<td>-0.006</td>
<td>-0.037</td>
<td>-0.031</td>
</tr>
<tr>
<td>A</td>
<td>BID0033</td>
<td>23,000</td>
<td>NSW</td>
<td>3.5</td>
<td>3.5</td>
<td>4.0</td>
<td>0.014</td>
<td>-0.012</td>
<td>-0.026</td>
</tr>
<tr>
<td>B</td>
<td>BID0017</td>
<td>28,000</td>
<td>NSW</td>
<td>4.5</td>
<td>5.0</td>
<td>5.0</td>
<td>-0.014</td>
<td>-0.032</td>
<td>-0.019</td>
</tr>
<tr>
<td>F</td>
<td>BID0051</td>
<td>11,000</td>
<td>QLD</td>
<td>3.0</td>
<td>3.0</td>
<td>4.0</td>
<td>0.005</td>
<td>-0.010</td>
<td>-0.015</td>
</tr>
<tr>
<td>E</td>
<td>BID0042</td>
<td>25,000</td>
<td>VIC</td>
<td>–</td>
<td>3.5</td>
<td>4.0</td>
<td>-0.042</td>
<td>-0.042</td>
<td>0.000</td>
</tr>
<tr>
<td>C</td>
<td>BID0047</td>
<td>27,000</td>
<td>NSW</td>
<td>–</td>
<td>4.0</td>
<td>4.5</td>
<td>-0.004</td>
<td>-0.002</td>
<td>0.002</td>
</tr>
<tr>
<td>A*</td>
<td>BID0035</td>
<td>18,000</td>
<td>NSW</td>
<td>4.5</td>
<td>5.0</td>
<td>5.0</td>
<td>-0.014</td>
<td>-0.010</td>
<td>0.005</td>
</tr>
<tr>
<td>C</td>
<td>BID0041</td>
<td>19,000</td>
<td>NSW</td>
<td>–</td>
<td>3.5</td>
<td>3.5</td>
<td>-0.007</td>
<td>0.006</td>
<td>0.013</td>
</tr>
<tr>
<td>F</td>
<td>BID0044</td>
<td>36,000</td>
<td>QLD</td>
<td>–</td>
<td>3.0</td>
<td>4.5</td>
<td>-0.019</td>
<td>0.008</td>
<td>0.027</td>
</tr>
<tr>
<td>C</td>
<td>BID0039</td>
<td>14,000</td>
<td>QLD</td>
<td>4.5</td>
<td>4.5</td>
<td>4.5</td>
<td>-0.024</td>
<td>0.020</td>
<td>0.044</td>
</tr>
<tr>
<td>C</td>
<td>BID0038</td>
<td>15,000</td>
<td>NSW</td>
<td>4.5</td>
<td>4.5</td>
<td>5.0</td>
<td>-0.013</td>
<td>0.079</td>
<td>0.092</td>
</tr>
<tr>
<td>B*</td>
<td>BID0026</td>
<td>51,000</td>
<td>NSW</td>
<td>3.0</td>
<td>3.0</td>
<td>4.0</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>D</td>
<td>BID0006</td>
<td>10,000</td>
<td>NSW</td>
<td>4.5</td>
<td>4.5</td>
<td>4.0</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>F</td>
<td>BID0049</td>
<td>53,000</td>
<td>QLD</td>
<td>–</td>
<td>4.0</td>
<td>5.0</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>C</td>
<td>BID0007</td>
<td>8,000</td>
<td>NSW</td>
<td>1.5</td>
<td>2.5</td>
<td>3.5</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>B*</td>
<td>BID0018</td>
<td>14,000</td>
<td>NSW</td>
<td>–</td>
<td>–</td>
<td>4.5</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>F</td>
<td>BID0024</td>
<td>15,000</td>
<td>QLD</td>
<td>1.5</td>
<td>4.0</td>
<td>4.0</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>F</td>
<td>BID0028</td>
<td>20,000</td>
<td>QLD</td>
<td>4.5</td>
<td>5.0</td>
<td>4.5</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>B*</td>
<td>BID0036</td>
<td>28,000</td>
<td>NSW</td>
<td>4.5</td>
<td>4.5</td>
<td>5.0</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

Mean: 3.8 3.8 4.2 -0.001 -0.033 -0.032

* Those buildings marked with an asterisk joined their group after the majority had commenced.
When considered as a group, the relationship between NABERS ratings and daily variance data is consistent and indicates a stable sample that experienced a significant change arising as a consequence of the intervention. There is evidence of a minor within-sample skew relating to social learning groups, however, which is highlighted in Figure 12-2, below. The results for each social learning group were reasonably well spread with the exception of Group C, which significantly underperformed in Figure 12-2, and Groups D and E which outperformed. Personal characteristics of FMs were not expected to matter for the purposes of this study and it can be said that their general patterns of behaviour, backgrounds, incentives, etc. were broadly similar (though not necessarily representative of the wider community). But when we consider a much smaller sample of just six social learning groups—each presided over by a single senior facilities manager—the influence over the results of each group’s particular dynamics and each senior facilities manager’s personality and priority differences will be greater. This may have been the case for groups C, D and E. Group C (which includes BID0038 and BID0041 discussed above) experienced some disruption and difficulties with movement of facilities managers and was led by a senior facilities manager who had recently been promoted from the FM role at a building which was included in another group. It is unclear what factors, or combination of factors, may have contributed to the underperformance of Group C, though these two are likely to have played a role. Savings at the two highest-performing buildings in Group D (BID0043 and BID0048) arose after significant problems with the operation of their HVAC services were identified and corrected. The two FMs (and their senior FM) appeared to be among the most engaged of all involved in the study, so it is considered likely that the intervention led directly to the identification of these issues which had persisted undetected for many years. Group E was comprised of large buildings in the Melbourne CBD. More so than any other group, there was a rivalry
between members of Group E, and in particular, between two of the buildings (the two largest involved in the study) which had hierarchies that included a member part-way between facilities manager and senior facilities manager level. It is likely that this complex dynamic contributed to the outperformance, a result that was unexpected given technical challenges with the feedback (to be discussed in chapter 13).

Building size had no discernible influence on changes to energy use. As illustrated in Figure 12-3, below, buildings with a floor area smaller than the mean (28,200m²) saved more and also less energy than larger buildings which were concentrated towards the middle of the sample. This narrower though still normal distribution may suggest that larger buildings are less dynamic in terms of their energy use, or it may just have been a coincidence brought to attention by the small sample of eleven buildings in each category.

![Figure 12-2. Distribution of sample variances by social learning group](image_url)
As with building size, pre-trial intensity of energy use (indicated by 2011 NABERS ratings) appears to have had no noticeable bearing on the level of savings achieved. Figure 12-4 shows the distribution of buildings scoring 4 stars was almost identical to those which scored 3.5 stars and below, with both groups split evenly either side of the trial variance midpoint (-0.033%/day). There may possibly have been a very slight skew towards underperformance for the higher NABERS-rated buildings although this skew is most likely explained simply by the small sample of thirteen buildings in the higher-rating category and nine in the lower.
Observations about the irrelevance of pre-trial NABERS ratings are reinforced in a recent analysis by Steinfeld (2016) and presented at Figure 12-5, below, which included data from 123 Australian commercial office buildings spanning 2,300,000m² (NLA) – approximately 10 percent of the Australian market (PCA, 2009). Steinfeld’s large sample included 49 buildings rated at 3.5 stars or less prior to an equivalent intervention and 74 rated at 4.0 stars or more\textsuperscript{14}. The twenty-two buildings presented in Figure 12-4 were all included in the Steinfeld analysis.

![Graph showing energy improvement](image)

**Figure 12-5. Comparison of improvement for 123 Australian office buildings categorised by pre-intervention rating on NABERS Energy (Steinfeld, 2016)**

It should be noted that Figure 12-4 and Figure 12-5 both present energy savings as a proportion of a pre-intervention baseline. Therefore, buildings that commence a behaviour-based intervention with a high EUI will need to save more than those with a

\textsuperscript{14} Data was derived from the Australian Government’s Commercial Building Disclosure database (http://www.cbd.gov.au) and Buildings Alive Pty Ltd (http://www.buildingsalive.com), a provider of building efficiency information services similar to those developed through the course of this study. Buildings Alive was co-founded by the author at the conclusion of the trial phase of this study to further explore opportunities for energy savings using feedback methodologies similar to those discussed here.
low EUI to achieve an equivalent percentage improvement. This may help explain the apparent inconsistency between the findings presented here and those of Granderson et al. who found that a high pre-existing level of EUI was the most important influence on the level of energy savings achieved from deploying energy information systems (Granderson et al., 2013). Either way, these results suggests that significant opportunities for achieving behaviour-based energy savings are consistently overlooked in commercial office buildings. It is likely given there is no clear indication that improvements are proportionally any different for ‘low’ or ‘high’ EUI buildings, that the primary difference between the upper and lower buildings is their technology, not their operators’ underlying levels of motivation or technical / managerial competence.
13. LIMITATIONS AND STRENGTHS OF THE METHODOLOGY

13.1. Multidisciplinary focus and design

A review of 4,444 energy studies published in leading energy journals between 1999 and 2013 found that only 0.3 percent of authors reported affiliations with social science related disciplines such as psychology, history, anthropology, and communication studies and that fewer than 4.3 percent of citations were to social science and humanities journals. Quantitative methodologies were by far the most common (57.9 percent of articles) followed by articles which collected no original data and relied entirely on secondary literature. Only 12.6 percent of the studies utilised qualitative methods (Sovacool, 2014a, 2014b). While the methodology adopted here is quantitative and the methods of intervention and analysis were very much technology-based, the primary focus of enquiry has been the role of social and psychological factors in influencing energy use in buildings, and in particular, the influence of human motivation, learning and choice. As Flyvbjerg (2001) has noted, a blending of the two approaches—one centred on science and technology, the other on people—can lead to much deeper practical insights than either by itself. This is not a new idea. Aristotle discusses the three intellectual virtues of *episteme* (scientific knowledge), *techne* (craftsmanship) and *phronesis* (practical wisdom) in *The Nicomachean Ethics*, arguing that natural and social science are and should be different ventures (Flyvbjerg, 2001). Drawing from this, Flyvbjerg suggests:

“In their role as phronesis, the social sciences are strongest where the natural sciences are weakest: just as the social sciences have not contributed much to explanatory and predictive theory, neither have the natural sciences contributed to the reflexive analysis and discussion of values and interests.”
which is the prerequisite for an enlightened political, economic, and cultural development in any society, and which is at the core of phronesis.”

(Flyvbjerg, p. 3)

There is a growing acceptance that a deeper understanding of the role played by social and psychological factors will be required in order to build a thorough understanding of energy efficiency improvement opportunities generally (Andrews & Johnson, 2016; Sovacool, 2014a; Sovacool et al., 2015; Stern et al., 2016), and especially in relation to buildings (Lucon et al., 2014). This deeper understanding is not only required at the policymaker level, but also at the practical level: too often technology-based options are chosen ahead of lower cost, faster, and potentially more effective human-centred alternatives. This can be observed at the government policy level (see Levine et al., 2007; Lucon et al., 2014), but for evidence of its implications for energy use in large commercial office buildings look no further than the results presented here.

As noted by Martiskainen (2007), measures for promotion of energy saving can be divided into three broad categories: Antecedent measures, through the use of information materials and modelling for example; consequence measures such as through the use of feedback, including use of rewards and incentives; and social influences, such as through the use of groups and various commitment techniques (see also Abrahamse et al., 2005; De Young, 1993; Katzev & Johnson, 1983). In this study I drew from all three categories to develop a prescription for effective behaviour-based interventions in large complex buildings. This breadth of applied measures is another factor that makes the study somewhat unusual; however the objective was to build operator engagement and measure the effect through changes to building energy use, not to isolate mechanisms or advance the underlying theory. Problem-focused, diverse and interdisciplinary methodologies are
beginning to find support for a variety of reasons, including their deeper impact, as Sovacool opined recently in Nature:

“Problem-focused research activities that centre on both physical and social processes, include diverse actors and mix qualitative and quantitative methods, have a better chance of achieving analytic excellence and social impact.”

(Sovacool, 2014a, p. 530)

There is still much about the interactions between human behaviour and energy use that has yet to be resolved and not every relevant and broadly accepted insight from the literature can be applied in the one study, of course. Such constraints are not considered important. For example, we explicitly provided knowledge of results, but formative assessment was left up to the social learning environment, i.e. there was no provision of “high quality information to students about their learning” (as suggested by Nicol & Macfarlane-Dick, 2006), just clear and timely feedback with the opportunity for discussion. Likewise, the feedback provided did adhere to the definition proposed by Ramaprasad (1983) (“Feedback is information about the gap between the actual level and the reference level of a system parameter which is used to alter the gap in some way”), however our objective was not to “close” the gap between actual performance and a goal (as suggested by Nicol & Macfarlane-Dick, 2006; Sadler, 1998) but rather expand the gap between actual performance and the normalised baseline. Implications of this are discussed in the next chapter.

Another important feature of the study design is its longitudinal nature. Other research focussed on similar questions, such as the University of Sydney’s Low Energy High Rise project, has looked at a sample of buildings at a point in time and attempted to draw conclusions from statistical analysis of the within sample variances (Warren Centre for
That approach can be effective for pointing to anomalies and opportunities, but it offers little insight into the potential to change the performance or behaviour of a subject through direct intervention. As can be observed from the results in chapter 11, the effect of the interventions described here was not immediately evident and only built up over time. An action research methodology such as this one is able to reveal the time-dependence of a response to an intervention and, when combined with observations about the sample, it can also potentially say more about the likelihood of replicability.\footnote{For a discussion about study designs and settings and their likelihood of yielding false / true claims, see Ioannidis (2005)}

### 13.2. Stable conditions

Action research, by definition, cannot be conducted in a controlled environment. But the stability of the contextual factors present in this study made it possible to attribute the energy usage effects to building operators’ deliberate actions. The publicly listed Investa Property Group was acquired by Morgan Stanley Real Estate in 2007, immediately prior to the Global Financial Crisis (Morgan Stanley, 2007) and three years before the commencement of this study. Operating conditions remained abnormally stable throughout this period because the economic climate limited supply of investment capital and drove an overarching focus on “securing tenants to minimise risk.” This stability persisted until after the trial’s conclusion (Investa Property Group, 2016b).

According to Triandis’ TIB model, it is the influence of facilitating conditions over subjects’ underlying intentions and habits that determine their behaviours. Habits, as discussed in chapter 7, mediate behaviour, and it takes time for actions influenced by changes in conditions and involving conscious effort to become habitual. We can assume that the intention of FMs—and the attitudes, affect and social factors that determine it—
remained constant given the population was unchanged throughout the baseline and intervention periods. Likewise the direction and ‘support’ provided by the owner were held constant (aside, of course, from the intervention itself) and the buildings were unchanged. Contextual stability is necessary to infer a causal relationship between changes in facilitating conditions and behaviour change. And because this was achieved, the net effect of altering facilitating conditions will be directly proportional to observed changes in behaviour (from equation [4]). It follows, therefore, that FMs’ adoption of energy saving behaviours arising from changes to the availability of operational feedback and the provision of learning opportunities would lead to a proportional reduction in the use of energy \(E\), i.e.:

\[\Delta F \propto \Delta B \rightarrow \Delta B \propto \Delta E\]

(20)

### 13.3. Quantitative study

It is a feature of this study’s design that it offers only limited insights into the characteristics, motivations and decision making processes of facilities managers. As discussed in chapter 8, the buildings and their operators were selected on the basis that they can be regarded as ‘typical’ of the Australian commercial office sector. It follows, therefore, that the FMs engaged in this study could be expected to react to interventions in the same way as their peers would. It doesn’t matter what specific actions they took (for example, did they turn lights and equipment on later or off earlier, identify heating/cooling conflicts, find opportunities for free cooling from outside air, etc.?). What’s important, in the context of this study, is that we are able to measure the effect of operator actions through a consequential change in energy use.

A quantitative method was adopted in order to focus on the significance of the intervention and its broader application. This gave an assessment of the effectiveness of
the intervention overall, but no detail on individual buildings or FMs. A statistical approach could not have yielded further insights from a categorisation of FM personalities, building systems / services design characteristics, interactions between participants, emotional reactions to feedback etc. because the size of each subsample would have been too small. And while a qualitative methodology was initially favoured and may have been quite effective for drawing insights into those aspects (Roussac, de Dear, et al., 2011), it was found that without a clear understanding of the prevalence of each of the FM archetypes there would be limited grounds for claiming any general applicability of findings (Swim et al., 2011).

Notwithstanding the above comments, a few examples of actions taken by facilities managers directly in response to the feedback interventions are presented in the following chapter. These actions are indicative of changes in operator motivation and competence (intervening variables) that led to significant and prolonged changes in energy usage patterns (dependent variable), and therefore suggest the sequence and interaction between variables proposed in Figure 7-1 was valid. Taking the statement by Nevid (2013) that “We don’t actually observe a motive; rather, we infer that one exists based on the behaviour we observe”, it follows that we should be able to infer changes in motivation and behaviour have occurred if we measure changes in energy use and can be confident there are no other explanations for those changes.

13.4. Trust and credibility was high

People will not respond positively to feedback unless they trust and respect its source. The actual quality and accuracy of the information is of secondary importance because, unless it is trusted, it will tend to be ignored (Craig & McCann, 1978; Sovacool, 2014b; Stern, 1993). There were three aspects of the methodology that engendered high levels of
trust and helped to maintain engagement and confidence in the project when challenges emerged. The first was the source of raw data. The information contained in daily feedback messages was calculated from weather and electricity data derived from the Australian Government’s Bureau of Meteorology and AEMO-accredited Metering Data Providers respectively. Both are official data sources used to provide critical services to the nation and the quality of their data is beyond question. The second was the source and structure of the daily messages. As noted by Sadler (1989), for feedback to be regarded as valid it must be “on target” and well-based from the point of view of the recipient. Each FM was introduced to the project via a face-to-face discussion with the researcher where his/her building’s statistical model was explained and any difficulties with the modelling were presented. This built trust and a reasonable level of understanding about the mechanisms behind the feedback. It was also through this process that the idea of showing profiles of nearby peer buildings emerged (a feature introduced approximately two weeks after the commencement of the first trial with Group A, see Figure 9-6). Engagement was further enhanced by the project being conceived and managed “in house” which helped dampen FMs’ ever-present suspicions about commercial motivations and conflicts of interest. Finally, noting, as it was in chapter 5, that existing professional and social networks tend to provide the most trusted sources of advice and information (Stern & Aronson, 1984), the recipients’ peers were drawn upon as the primary source of additional information, interpretation and guidance.

13.5. **Limited interaction with tenants**

The fact that operational control for most buildings was separated into “base building” and “tenancy” meant that FMs did not need to seek the cooperation or support of tenants for their energy efficiency ideas. Interactions between owners/operators and tenants are largely governed by the terms of their leases and while lease requirements can constrain
the actions of operators (see Roussac & Bright, 2012), they also codify the nature of the relationships. One implication for this study was that FMs eschewed energy saving opportunities they believed could potentially impact occupant comfort negatively. Greater savings would probably be achieved if operators were prepared to challenge lease specifications, for example if they adjusted temperature setpoints upwards in hot weather or reduced the air-conditioning service hours (Roussac, Steinfeld, et al., 2011a, 2011b).

On the other side of the coin, separation of responsibilities meant that FMs were free to adjust plant operating parameters without the need for consultation, so long as they did not breach the lease.

13.6. **Factors impacting replicability of results**

In relation to studies such as this one which seek to contribute to a better understanding and capacity for mitigating climate change, Swim et al. (2011) cautions against claiming “…that findings from any specific group have general applicability without evidence or strong theory to support such claims” (p. 248). They point out that psychological phenomena, such as the response of building operators to energy performance feedback stimulus identified in this study, are often context dependent. We must consider a study’s external validity before attempting to generalise the findings or claiming broader significance, even if the theoretical basis for such claims is sound.

Twenty-two buildings, approximately 620,000m² or 2.8 percent of the Australian investment-grade office stock, is a small sample out of a global population that could exceed 900 million square metres (based on the proportion that Australia represents of the global economy). However, as noted in preceding chapters, the intervention produced a large effect size (indicated by a Cohen’s $d$ of 1.79 where greater than 0.8 is considered “large”) and generated sufficient evidence to reject the null hypothesis. It is therefore
reasonable to suggest, based on the findings of this study, that similar interventions applied at similar buildings in similar circumstances may generate similar results. But what should be regarded as ‘similar’? And what are the implications for buildings in dissimilar settings?

Here are some features of the sample that could potentially limit the findings’ external validity:

- Large, modern, ‘institutional grade’ Australian commercial office buildings in CBD locations

- Market-leading energy performance prior to trial commencement

- Professionally managed

- Highly motivated owner-operator

- Focus of feedback limited to ‘base building’ only (tenancies ignored)

- Researcher’s association with the sample

Other factors such as the sources and reliability of data were discussed earlier in this chapter as they are not regarded as significant from a psychological or operational perspective, even though they matter greatly for replicability of the methods and results.

Each of the six potential limitations is addressed below.

13.6.1. Large modern Australian office buildings in CBD locations

The majority of the twenty-two office buildings included in the analysis were high-rise (15 storeys and higher) with centralised HVAC plant controlled by a relatively sophisticated BMCS and located in the CBDs of Melbourne, Sydney and Brisbane. The
average floor area was approximately 28,000m² and ages ranged between ten and fifty years. Other characteristics were summarised in chapter 11.

Clearly the cultures and climates in the Australian cities of Sydney, Melbourne and Brisbane have more in common with each other than they have with many other cities around the world. The trial buildings may have characteristics that are uncommon elsewhere or influence energy usage and performance differently in different settings. For example, numerous studies have found that western-style buildings in Asia with centralised HVAC plant can use up to nine times more energy than conventional (for Asian cities) buildings with simple decentralised systems that operate on an ‘as needs’ basis (Jiang, 2012; Murakami et al., 2009; Zhaojian & Jiang, 2007). While low-tech Asian high-rise buildings likely present an extreme end of the spectrum, this obviously challenges the validity of transferring conclusions between climatic regions and cultures.

Commercial office buildings exist to provide comfortable and productive work environments for their occupants, wherever they happen to be located, so we can assume office buildings in most global markets have many similarities to those included in this study – even if we have insufficient research upon which to base an extrapolation of the results. This is far less likely to be the case for non-office commercial buildings outside of Australia where the differences are more pronounced.

Other aspects of the sample selection are less likely to negatively impact the external validity of the study. The results give no indication that managers of the larger buildings achieved proportionally larger savings than those achieved at the smaller buildings and there is no reason to expect that a building’s location within a metropolitan area would have any bearing on results. Likewise, it is unlikely that the presence or otherwise of a BMCS would matter greatly. The purpose of a BMCS is to control the operation of a
building’s plant and equipment: above a certain level of size and complexity it becomes unfeasible for a building manager to control building services without the aid of a BMCS. Clearly the buildings included in the trial were of a size that required a BMCS, however smaller less complex buildings providing similar services to occupants are more likely to be able to be operated by a building manager manually or with rudimentary control systems. It seems reasonable to assume the factors influencing a building manager’s decision to alter BMCS controls would likely be similar to those influencing a decision to adjust plant and equipment settings directly.

In chapter 2 it was noted that only five percent of commercial facilities in major global office markets are larger than 5,000 square metres, and yet they represent more than 50 percent of total NLA (MacDonald & Bray, 2011). Large buildings, despite being small in number, do represent a significant proportion of the commercial property sector’s energy usage; however to extrapolate this study’s findings to a global population it will be necessary to investigate the effects of feedback to operators of buildings of varying size in diverse cultures and climates.16

13.6.2. Market-leading energy performance prior to trial commencement

It is beyond question that both Investa and its portfolio of buildings were highly unusual for their outstanding energy performance and focus on environmental concerns prior to the commencement of the trial. However, the fact that Investa was an industry leader does not make it an inappropriate host for a study such as this one. The stable energy performance arising from its mature energy efficiency investment program during the years immediately prior to commencing the feedback intervention (see Figure 11-6 and

16 In relation to the question of size, the researcher was involved in an investigation commissioned by the City of Sydney after the conclusion of this study which looked at smaller “mid-tier” office buildings and the results of that investigation were consistent with (and are included in) Figure 12-5 – however the sample size was too small to provide any clear insights.
Table 12-1) was helpful in isolating the effects of changes and it also suggested it was just a few years ahead of a general trend (illustrated in Figure 2-2). It also gave a level of assurance that ‘easy wins’ would not be available, and that the results would represent a conservative evaluation of the potential savings from similar interventions applied to the broader population of buildings.

13.6.3. Professionally managed

The office buildings selected for this study all generated sufficient rental income from tenants to justify the employment of at least one fulltime professional facilities manager. Smaller and less prestigious buildings usually generate less rental income and cannot, therefore, afford the same level of resourcing. Despite having more management time and expertise available, the buildings included in this study did not necessarily have access to proportionally more management resources for focusing on energy performance than others might have. This is because all building owners, as a rule (Investa included), would prefer not to pay for any more resources than are necessary to operate their buildings effectively. Previously discussed criticisms of the rational choice model notwithstanding (see Blume & Easley, 2008; Jackson, 2005; Sen, 2008), management time in a large institutionally-owned building is likely to be just as stretched as in any other commercial building. Indeed, the expertise and experience of the Investa senior facilities managers varied widely, as did the size of their site teams, and yet results were reasonably consistent. The pre-intervention performance level of buildings and their managers was excluded from this study by design (it was normalised out). However as illustrated in the previous chapter, the benefits of feedback are reasonably consistent irrespective of base level performance or operator profiles.
13.6.4. Highly motivated owner-operator

Investa’s portfolio is internally managed (i.e. building management teams are directly employed by Investa and not by third-parties) and this is often not the case for commercial office building operations. Often FMs are employees of large independent companies contracted to manage the buildings on the owners’ behalf. The roles and functions of FMs are fairly consistent irrespective of who their direct employer is; however the culture and priorities of the employer organisation may differ from the owner organisation’s and this can impact an FM’s capacity to focus on energy saving opportunities. For example, a specific energy-saving objective (or “goal”) may not have been set out in the contract between owner and contracted operator and, accordingly, the operator may insist its staff focus attention on other priorities (e.g. those that are stated in the contract).

In chapter 5 we discussed evidence presented in the literature for why people are motivated by the desire to achieve goals and the propensity for people to reduce their level of effort if they believe performance is exceeding the target standard (Gärling et al., 2002; Kluger & DeNisi, 1996; Locke & Latham, 2002). Evidence from residential studies shows that it is not appropriate to assume people have a pre-existing energy saving goal (McCalley & Midden, 2002). For the sample considered in this study, however, there was clearly a broad understanding of Investa’s overarching energy-saving goal among FMs because energy efficiency had for a long time been an important element of Investa’s brand identity. That is not to say that savings might not have been greater if individual participants had been invited to commit to their own goals. There is compelling evidence to suggest they would have been (L. J. Becker, 1978; Locke & Latham, 2002; McCalley & Midden, 2002).
The fact that the trial sample was drawn from a portfolio owned by a “highly motivated owner-operator” need not be a consideration in assessing external validity of this research. What is required is that FMs be working towards a clear set of goals – both performance goals and learning goals. Such goals are likely to generate sufficient focus on the performance feedback regardless of whether they are elicited from the individual (desirable) or expressed either directly or indirectly by the owner / employer (less desirable), as was the case here.

13.6.5. Focus on ‘base building’ only (tenancies ignored)

In most countries, electricity utilities measure consumption at the “whole” building level. In Australia it is more common for “base building” and “tenancy” supplies to be metered separately. Base building services include HVAC systems, lighting for external and common areas, vertical transportation and water heating. Tenancy services include the tenancy lighting, plug loads (such as IT and office equipment) and any supplementary HVAC services utilised by occupants in the tenanted spaces. Tenancy loads tend to be quite stable from day-to-day when compared to base building loads, which vary greatly on account of changes in external environmental conditions. This separation helps FMs focus on services that they are responsible for, and limits the influence of tenant operations which can overwhelm their conservation efforts (Bannister, 2012). As such, a concern could be raised that this study was more effective because the feedback was more tuned and targeted towards aspects of building operations that were directly within the focus of the FMs.

The Brisbane market is unusual for Australia in that it adopts the “whole building” approach to metering. This means that we can have an indicative assessment (given only 4 Brisbane buildings were included in the final analysis) of whether the concentration of
feedback on ‘base’ or ‘whole’ buildings made a difference. We can see that the four Brisbane buildings showed an average 5.27% improvement over the full evaluation period, slightly below the 7.31% observed for the other 18 and that the distribution was narrower than the sample overall (see Table 12-1 and Figure 12-2). When we consider that energy use within an Australian office building is typically split 60/40 between base building and tenancy (Bannister, 2012), the reductions in Brisbane buildings actually exceeded those achieved in the base building only cases. Interventions were only made on base building services and tenants were not informed or involved in the study, so this indicates that this characteristic of the Australian market is unlikely to matter when considering the broader significance of the study – although percentage reductions should be adjusted to account for the particular metering arrangements.

13.6.6. Researcher’s association with the sample

One aspect that needs careful consideration in a study such as this one is the potential for a so-called Hawthorne Effect to emerge. The term Hawthorne Effect was first coined by Henry A. Landsberger after he analysed data from a series of experiments conducted with the aim of increasing worker productivity at the Hawthorne Works (a factory outside Chicago) from 1924–32. Contrary to initial analyses by Mayo et al., which suggested that productivity increased as a consequence of the interventions (see Mayo, 1933), Landsberger established that it was, in fact, the novelty and attention from being research subjects that led workers to temporarily increase their productivity levels (Landsberger, 1958). Thus the Hawthorne Effect is now commonly defined as “an increase in worker productivity produced by the psychological stimulus of being singled out and made to feel important” (Franke & Kaul, 1978).
In conducting this study I was very careful to avoid introducing any noticeable changes to the FMss’ environments, support or their perceptions of what matters. Subjects were not aware they were participating in a study – for them it felt like yet another innovation to contend with as part of their job. Their consent to participate was not sought on the authority of their employer who wanted the study to proceed in a ‘business as usual’ manner and did not want staff distracted from their work or to feel like they were participating in research (see Appendix A). As previously discussed, I had been working on energy management at Investa since 2004 and my role and responsibilities in the organisation (where I was employed as General Manager, Sustainability, Safety and Environment) had been more-or-less static since 2008 – three years prior to the commencement of the study. From the participants’ point of view, I was not a “researcher” and there was no sign of “research” being conducted. My role and relationships were fundamentally unchanged. Likewise, other than the fortnightly meetings and the Pulse messages themselves, there were no significant changes to the way Investa approached or managed energy. Therefore, while it cannot be discounted completely, the likelihood that a Hawthorne Effect may have influenced the behaviour of subjects or the energy use of buildings is small.

* * *

Having considered the six factors discussed above, there seems to be a reasonable basis for claiming the results of this study have broader significance beyond the Australian commercial office sector. However, while the study may point researchers, practitioners and policy makers towards potentially effective feedback methodologies for different building types and in other cultures and regions outside of Australia, without more
evidence from similar studies conducted in other environments it would be appropriate to exercise a degree of caution when attempting to generalise the findings.
14. KEY FINDINGS

The FMs who participated in this study achieved significant energy savings after being exposed to automated daily energy performance feedback within a social-learning environment, and causal attribution to the intervention has been established beyond reasonable doubt. This chapter returns to some of the questions that arose in chapters 4-6; in particular, the importance of motivation, goal setting and peer interaction, to the development of a framework for presenting energy efficiency feedback to building operators. It then considers whether skill and competency levels increased over the course of the trial, and the likelihood that energy savings will persist beyond its completion. Despite well-established evidence that individual behaviours are heavily influenced by social and institutional contexts and reactions, and hence not readily promulgated at scale (see chapter 3), the analysis leads to specific conclusions and recommendations for structuring and delivering energy efficiency feedback to building operators.

14.1. Motivation and goals

An overarching objective of this study, as Geller (2002, p. 535) has expressed it, was “to help people get so personally committed to environmental protection that they would use self-management techniques to increase their pro-environment behaviour”. Unlike Geller, our interest is not with eliciting “pro-environment behaviour” or environmental outcomes \textit{per se} (even though reducing energy use has direct environmental benefits) but the concept is the same: we hoped to find evidence of people using self-management techniques to drive energy savings from the buildings under their control. According to Geller, if a person is to transition from someone who is not engaged in thinking about environmental protection (read: \textit{saving energy}) to one who actively strives to achieve energy savings, they will generally need assistance as they pass through three phases: 1)
awareness raising, 2) deployment of motivational techniques, and 3) provision of appropriate support. The first two are the subject of this section; the third will be addressed in the next section where we consider the role of the social learning environment.

14.1.1. Awareness is necessary, but not sufficient
Reasons for why awareness-raising is not, in and of itself, generally regarded as a motivation-raising technique were presented in chapter 5. However, in some circumstances it can be. The failure of the initial communication prototype discussed in chapter 8 is one example of how a meaningless or poorly targeted message will likely be ignored. But where targeted information is delivered in order to raise the awareness of building operators who have the appropriate knowledge and suitable strategies to draw upon, heightening of awareness can be expected to motivate action (Sadler, 1989, 1998, 2010; Van der Schaaf et al., 2013), as was the case with the intervention presented in this study.

14.1.2. No need for “rewards” or “incentives”
Incentive theories suggests people are motivated to do things because of rewards (Bernstein, 2011). However, as discussed in chapter 4, just like awareness, incentives can motivate and engage people or demotivate and alienate them, depending on the social conditions in which they find themselves (Ryan & Deci, 2000; Stern, 1993). Deci et al. (1999), for example, found that tangible rewards tend to reduce intrinsic motivation because reward contingencies weaken people’s ability to self-regulate and transfer their responsibility for motivating themselves. Also, as has been found by Bénabou & Tirole (2004) among others, offering of rewards will in some circumstances be interpreted by the recipient as conveying something negative about the nature of the task or their ability to perform it. So while a person may perform the task to a high standard in order to meet...
a target and receive a reward, they may do so in spite of an overall reduction in their intrinsic motivation.

Noting those insights from the literature, it was decided that no direct incentives to save energy would be offered to participants in this study. The few indistinct reward opportunities that had been in place over many years remained (e.g. a tacit recognition that managers who save energy are helping their employer’s business and may be looked upon favourably for promotion; a small portion—less than 1 percent—of the typical facilities manager’s annual remuneration ascribed to bonuses relating to energy metrics), but there was no evidence that these had ever driven behaviour to any significant degree. Furthermore, the daily feedback provided in this study did not include financial information. Electricity has a price, obviously, so FM’s understand that saving energy is saving money. But, as is generally the case in organisational settings, there was no direct or reliable way for FM’s to connect the knowledge that their buildings might be saving energy with the prospect of a reward for them personally (Carrico & Riemer, 2011).

Given there were no exogenous rewards on offer and assuming that most FM’s were already intrinsically motivated to operate their buildings well, i.e. reliably and efficiently, the energy performance information delivered as timely and unambiguous feedback was, of itself, sufficient to conjure an intrinsic reward that motivated action.

14.1.3. Explicit goals are unnecessary

Facilities managers did not have explicit energy saving targets. They were not asked to commit to any specific goals (either in relation to performance or learning) and none were set for them by the researcher or their employer. In these respects the organisational environment was ‘typical’ of that experienced by most facilities managers (Goulden & Spence, 2015). In saying this, Investa’s commitment to environmental ‘sustainability’ and
performance is fundamental to its brand (Investa Property Group, 2016a). Whether this corporate commitment actually trickles down to individual decision-making processes is uncertain, however. For example, an implicit but poorly-founded assumption built into most residential energy feedback studies is that participants have a goal to save energy. Often this is not true and even among those who seek to achieve savings there is much variability in how competing goals are prioritised (McCalley & Midden, 2002).

The literature contains much evidence pointing to the benefits of individuals explicitly committing to their own goals through a rigorous goal-setting process (L. J. Becker, 1978; Kluger & DeNisi, 1996; Locke & Latham, 2002; McCalley & Midden, 2002). Goals mobilise and direct effort towards goal attainment (Gärling et al., 2002). As discussed in chapter 5, typically they work because people apply more effort when feedback shows performance falling short of their goal. However, most people will reduce (or just maintain) their level of effort if performance exceeds the goal (Kluger & DeNisi, 1996).

As there were no goals explicitly being set in this study, they apparently weren’t a factor driving participants’ efforts to achieve savings. But likewise, unchallenging goals did not hold them back either. Given the potential for these opposing motivational forces to cancel each other out across a sample of buildings and operators, could the absence of goals, therefore, have made little difference to the results? This question is purely hypothetical and the answer would depend on the sophistication of the goal-setting process adopted. Without a clear understanding of the level of savings achievable from an intervention such as the one adopted here, it would not be possible to know in advance where to set the goal for each building / operator. Setting goals or saving targets in those circumstances could be considered risky (they may be set too low) and it underpins the
importance of research to better understand the potential to save energy in buildings by presenting energy efficiency feedback to their operators.

One additional goal-related consideration to note is that neither ‘performance’ nor ‘learning’ goals were established for this study. As previously discussed (in chapter 5), high performance goals can drive people who have knowledge and training and skills, whereas learning goals are more appropriate for situations where there is a deficit (Earley & Perry, 1987; Locke & Latham, 2002). It is likely that the FMs at Investa achieved significant savings with an implied and open-ended performance goal floating in the background because they were already well equipped with the requisite skills and knowledge to utilise the feedback and participate in the forums that were introduced (see discussion in chapter 13). The information may not have been utilised as effectively if the FMs had been of a lower calibre. Under such circumstances a set of learning goals might be beneficial; however, a different research design would be required to explore that question.

14.2. The role of social learning

The social learning environment was integral to the study design and considered just as important as the feedback component. This is because the seeds for FMs’ ideas had to be drawn from somewhere. Other methods for prompting deliberative thought processes and purposive behaviour, such as coaching for example, may have been more effective but the facilitation of a peer network for discussing building performance and exchanging ideas was considered the most transparent and independent mechanism available. Furthermore, as discussed in the previous chapter, the level of researcher detachment afforded by the peer-to-peer engagement model is regarded as helpful for assessing external validity of the findings.
It was considered almost inevitable that FMs would start asking questions about their buildings’ performance upon exposure to automated daily feedback. And logically this meant that they would turn to their peers for ideas and suggestions. The structuring of ‘action learning groups’ met very little resistance because, in all likelihood, they would have formed anyway – albeit much less formally and potentially less productively in the sense that focus would probably have been less on “learning” and exchanging of ideas about improvement strategies and more on sharing opinions about the merits and value of the feedback itself. It is also hard to get busy FMs together regularly as a group as it requires coordination so without the structure of a regular fortnightly meeting it’s more likely they would only have met up individually.

Although we cannot isolate the social learning factors for separate evaluation, we can nonetheless draw some useful findings about their contribution to the overall results by relating some of the key lessons from the literature.

**14.2.1. Cues and trust are important**

Many reasons for adopting a ‘social’ stimulus to facilitate engagement with the feedback were discussed in chapters 6 and 9. One of the most important was the idea that information tends to be most trusted when it comes from existing professional and social networks (Stern & Aronson, 1984).

As there was no attempt to apply other social engagement methods it is hard to say anything definitive about the effectiveness of the adopted methodology compared to alternatives. But we do know that the interactions between FMs were highly supportive and participation rates at meetings were almost 100 percent. We know from the literature that people tend to seek out and emulate actions, beliefs, values and attitudes of those peers they respect and can relate to (Braaksma et al., 2002; Goldstein et al., 2008; Kim,
We also know that for tasks requiring creativity or complex problem solving, strong learners gain most from observing the performance of models whereas people with lower learning aptitude are more likely to succeed with direct instruction (Groenendijk et al., 2013; van Gog & Rummel, 2010). It was found that by matching peers and having them bring relevant content to meetings for discussion, FMs were able to receive the appropriate cues and apply them directly to their own buildings regardless of learning aptitude. This accords with Albert Bandura’s Social Learning Theory and in particular, the idea that behaviour modelling (e.g. introducing social cues) is an indispensable aspect of learning if a person lacks background knowledge on a task.

The success of the social learning environment created for this study is also predicted by the theory of reasoned action and the concept of ‘social proof’, both of which were discussed in chapter 6. In an uncertain situation people generally look to what others have done to determine the appropriateness of their actions – both for cues and also because they expect to derive benefits from the positive judgements of others (Ajzen & Fishbein, 1980; Cialdini et al., 1999). This was in evidence at the fortnightly meetings where participants came prepared to “talk about something that happened” in their buildings over recent days and then in the meeting that followed where peers often reported their experience replicating the initiative they’d recently heard about. Very few references were made to buildings and FMs in other groups, even though they were closely associated. This was interesting to note, and while additional input from outside sources may have been helpful it’s uncertain whether it would have been beneficial given the propensity for people to focus on differentiating characteristics rather than those they have in common (Abrams et al., 1990).
14.2.2. Pressure from hierarchy not required

Senior management ‘direction’ and ‘support’ for actions relating to energy use were unchanged throughout the entire project.

Most businesses use a hierarchical management system to organise their activities and ensure the priorities of the business are communicated from the leadership to front-line employees. Investa is no different to other real estate firms in this regard. The reporting line in place for facilities managers at the time of this study is illustrated in Figure 14-1, below, with shading indicating which participants were active in the study.

![Diagram of Investa hierarchy](image)

**Figure 14-1.** Investa hierarchy in place at the time of the study (active participants shaded).
Facilities managers had so-called ‘dotted’ reporting lines to other employees in the hierarchy (property managers, asset managers, etc., not shown), meaning that they were expected to follow their directions as well, but they were accountable to their line managers. As General Manager, Sustainability, Safety & Environment I obviously held a senior position in the organisation. While my position had influence, it did not have any direct authority over any of the study participants (other than the graduate engineer). Aside from the roles shaded in Figure 14-1, the only other person in the hierarchy having direct knowledge of the study was the Group Executive who authorised it (see Appendix A). His role was completely removed from operational matters and it is unlikely that he would have even discussed it with anyone subordinate in the structure, other than me.

The business expected FMs and senior facilities managers who were at the lowest levels in the hierarchy to focus on business objectives that were emphasised by their superiors, including property managers and asset managers (for a discussion about the role of the facilities manager in organisational energy use, see Goulden & Spence, 2015). And none of these more senior people were aware of the trial. It can therefore be stated with confidence that nobody with any authority in the hierarchy sought to influence the trial outcomes and, therefore, engagement with useful energy performance feedback can be expected without management support in an implicitly supportive environment, i.e. that this finding is generalizable.

14.2.3. Peers a good source of formative assessment

Information about the performance of each building and the success, or otherwise, of facilities managers’ efforts to influence that performance was delivered each day via the Pulse email messages. But there was no formative assessment contained in the automated feedback to guide FMs in developing techniques to attain progressively higher standards
of achievement. This important function was presumed to be performed within the social learning environment, as without it there would have been nothing to short-circuit the randomness and inefficiency of trial-and-error learning (Sadler, 1989).

A combination of automated feedback and direct instruction (coaching) from authoritative sources may have delivered larger savings more quickly, and this will be discussed in the next section. But it seems the interaction with peers did accelerate learning. An indication of this is provided by comparing results with those from a similar study which I conducted with operators of 53 buildings on the University of California, Berkeley campus from late 2014 onwards. The daily messages received by facilities managers at UC Berkeley were almost identical to those described here; the only difference being that the Berkeley FMs did not participate in facilitated social learning groups. Results for the first 220 work days (same duration as this study) of the Berkeley study are presented in Figure 14-2 below.

---

17 This project was initiated by UC Berkeley’s Energy Office and Buildings Alive, the company I co-founded in November 2012 to explore questions arising from research described in this thesis. It commenced while I was jointly hosted by UCB’s Center for the Built Environment (CBE) and the Building Technologies and Urban Systems (BTUS) department at Lawrence Berkeley National Laboratory (LBNL) during my PhD candidature.
When comparing to the results presented at Figure 11-5, the absence of a clear improvement trend could be taken to suggest the social learning environment was in fact the primary contributor to savings observed in this study. That would be an oversimplification. The circumstances of the two trials were vastly different, with the UC Berkeley buildings split evenly between teaching and research facilities (no commercial offices, and most buildings had relatively few opportunities for operator control) and all operated on a public university campus with significantly different climate, culture and performance objectives. The analysis of the UC Berkeley trial was not nearly as rigorous as the one presented here and it is provided for illustrative purposes only, although a summary of the exercise and results through to March, 2016—including the explanation for the apparent erosion of savings—has been published for a general audience by the U.S. Department of Energy (Hartke, 2016) and also UC Berkeley’s energy office (see Roussac, 2016).

The Pulse messages provided the “high quality information” necessary to strengthen each FM’s capacity to self-regulate performance (Nicol & Macfarlane-Dick, 2006). Put simply, without clear and timely feedback delivered each day, FMs would not have had a reliable basis for discussing the energy-saving effects of their activities. Augmented feedback provided through the social learning environment complemented that received through cause-effect (i.e. the Pulse messages) and therefore helped generate positive motivational beliefs and self-esteem (Nicol & Macfarlane-Dick, 2006; Sadler, 1998). In effect this introduced a ‘learning’ dimension which, as Locke & Latham (2002) suggest, enhanced performance by helping to avoid evaluative pressure and performance anxiety.
14.3. **Skill levels increased**

It was anticipated that the daily feedback and social learning environment would support the development of evaluative expertise and deliberate self-monitoring among facilities managers (see chapter 5), and the positive results lend support to this hypothesis. In this section we consider some examples of the types of actions undertaken and consider whether the rate of improvement in energy performance might have been even faster had an alternative framework been adopted.

**14.3.1. Evaluative expertise emerged**

Experience from the first few weeks following the intervention at building BID0001 gives a good snapshot of the kinds of activities that took place in most of the participating buildings (see Roussac, 2012; Roussac & de Dear, 2012). Within a week of the *Pulse* feedback system being implemented the FM at BID0001 had made the decision to reduce space temperature setpoints by one and a half degrees from $23^\circ C\pm1.5$ down to $21.5^\circ C\pm1.5$. That particular facilities manager had very little familiarity with building mechanical services, having been promoted from a security attendant role only a year or so prior to this study. On Monday 22 August, 2011, the first workday after implementing the change and only one week after commencing the trial, the daily *Pulse* message arrived in his email inbox with the message: “Great news, yesterday [Building BID0001] used 8.5% less energy than expected.” The screenshot of BID0001’s HVAC electricity submeter shown in Figure 14-3, below, indicates the reduction in the morning 'spike’ that occurred on Friday 19 August. The FM was very pleased by this and noted that the change had not resulted in any discernible increase in tenant comfort complaints.
Figure 14-3. Screenshot of BID0001's HVAC electricity submeter indicating a reduction in the morning 'spike' following a lowering of space temperature setpoint (Roussac & de Dear, 2012).

That initial change was introduced at the end of the Sydney winter. Just over a week later the FM decided to raise the space temperature setpoint by 0.5°C (back up to 22°C±1.5) in an attempt to better balance heating and cooling demand and thereby delay the chiller plant’s start. This simple refinement delivered a gratifying result for the FM whose Pulse email message on Thursday 1 September stated: “Great news, yesterday [Building BID0001] used 5.1% less energy than expected”. Even more significant from the point of view of the FM was the chart (illustrated in Figure 14-4 below) showing the near-complete elimination of the morning ‘spike’ from the building’s energy usage profile for Wednesday 31 August and its closest 'like day' (Friday 26 August, 2011), both of which came after his tuning initiatives.
Figure 14-4. Impact of fine-tuning temperature setpoints at BID0001 on 31 August, 2011 and Friday 26 August, 2011, as illustrated in the ‘Pulse’ message.

The two examples cited above are of deliberate actions undertaken by a facilities manager who appeared to set himself an optimisation challenge and refine his approach iteratively. There were other instances where learning appears to have been more ‘accidental’. A few days later, on Monday 5 September, 2011 the FM at BID0001 opened his morning Pulse email to read the following: “Great news, last Friday [Building BID0001] used 22.9% less energy than expected.” This was by far and away the most significant saving he had produced and it prompted this written reply to the message:

“[On] Thursday (1st Sept) chillers had annual service clean, chillers were locked out on Friday. 23% savings seen for Friday with no ‘too hot’ complaints.”

The saving was pure luck arising from a technician’s mistake. Under warmer weather conditions the FM would in all likelihood have been alerted to the problem by occupant comfort complaints. In this instance, however, he was only made aware of the issue when
he tried to work out why the weather-normalised feedback contained in his daily message was so unexpectedly good. This prompted him to reflect on the building’s HVAC control strategy and the opportunity that clearly existed to reduce unnecessary cooling in mild conditions (the maximum dry bulb temperature that day was 18°C and the minimum was 14°C).

Experiences such as these drove the conversation at the fortnightly group meetings where each attendee was encouraged to “talk about something that happened”. As a consequence, not only did participants develop expertise in evaluating their own buildings, they also began evaluating and learning from the experiences of FMs at nearby buildings. Within a fortnight of commencement, this prompted the FMs in Group A to begin asking their senior facilities manager to distribute the normalised graph overlaying each of the peer-group’s buildings (graph shown at bottom of Figure 9-6) which had initially been intended only for senior facilities managers. That graph was subsequently introduced to the daily format for all Pulse messages in October, 2011, and it became a popular means for FMs to scrutinise and verifying daily feedback messages once they became familiar with how their building’s profile typically compared with nearby buildings under different conditions.

14.3.2. Skill transfer could have been faster

For a combination of reasons, most notably the FM’s high level of engagement with the intervention, the Pulse messages and group meetings did not cease at the conclusion of the study but continued on under a commercial model where buildings were required to pay money for the feedback and learning support their operators received. Additional buildings owned and operated by different parties were also introduced to a methodology resembling the one described in this study, but with active coaching and facilitation.
provided in addition to the daily feedback messages and peer-group support. This post-intervention period has yielded additional insights that could not have emerged in a controlled field study research design, including the potential for even larger energy savings by supplementing the social learning with coaching and facilitation.

As noted in chapter 11, the study duration was between 200 and 300 workdays for the majority of buildings and it took approximately 50 days for significant energy savings to emerge (see Figure 11-4). Though energy savings continued to build, the rate of improvement slowed from about day 100 and may even have begun to plateau towards the end of the study.

Coaching and facilitation were deemed necessary in the post-study commercial environment as owners (including Investa) began scrutinising performance and sought justification for the fees they had begun paying. Their expectation was that in order to “breakeven” (i.e. to offset the fee) the feedback messages and “support” should quickly deliver savings in excess of 5 percent. As can be observed from Figure 14-5, below, which shows trends for the year following the conclusion of the study (non-Investa buildings commenced after 3 months), it appears that reductions in energy use under the commercial model were achieved at approximately twice the rate observed under the experimental conditions at Investa.
Figure 14-5. Difference between ‘predicted’ and actual electricity consumption for all Investa buildings (including some not involved in the study). Post-study results (i.e. after ~250 days) were influenced by coaching and facilitation. All non-Investa buildings had coaching and facilitation from the outset.

While the methodological rigour of the commercial trial falls well below that presented in earlier chapters, the findings do tend to indicate that the rate of skill transfer and energy performance improvement can be accelerated if participants (in this case FMs) have access to expert prompting and coaching. The results of this study may therefore be at the lower end of the range that might be expected from behaviour-based interventions. Clearly a commercial model that combined automated daily feedback on energy performance relative to expectations, social learning environment, supplemented with active prompting and coaching produced a rate of improvement that was roughly twice that observed from the tightly controlled daily feedback and social learning model. Coaching and technical support appears to have introduced a wave of additional savings that may have been left unaddressed otherwise.
14.4. **Will savings persist?**

In contrast to most published field studies of (predominantly residential) energy efficiency feedback interventions which are too short to determine whether savings can be sustained over the long-run (Delmas et al., 2013), this intervention ran for more than a year at the majority of buildings. It therefore provided an opportunity to observe the influence of operator fatigue, the emergence of ‘boomerang’ effects, and whether or not habits were changed or formed.

14.4.1. **No boomerang effect**

When an effort to persuade someone of something leads unintentionally to them adopting an opposing position it is called a ‘boomerang effect’. This was discussed in chapter 5 and has been found to arise in energy efficiency feedback studies in which a subject’s perceptions about what is appropriate or socially desirable within their peer group (injunctive norms) are not consistent with the objective to save energy (Bittle et al., 1979-80; Brandon & Lewis, 1999). To avoid such problems, the intervention adopted a simple injunctive message (“good news” / “bad news”), much like the emoticons deployed successfully by Schultz et al. (2007). A related challenge was to avoid boomerang effects arising from perceptions (misplaced or otherwise) about what is common practice or a normal level of performance among a facilities manager’s peers. This was addressed through the influence of the facilitated peer groups.

As demonstrated in Figure 12-4, and more generally in Figure 12-5, there is no indication that pre-existing energy performance levels influenced the rate or direction of changes to individual buildings’ energy use: higher performing and lower performing buildings both improved in this study. Boomerang effects were avoided by the methodology adopted for this study. Furthermore, given that self-reinforcing social norms develop over time,
boomerang effects are unlikely to emerge in contexts or with conditions resembling those created for this study.

**14.4.2. Fatigue not an issue**

As also noted in chapter 5, a meta-analysis of experimental studies between 1975 and 2012 that deployed information-based strategies for energy conservation found that almost 60 percent lasted for fewer than three months and for each additional month of treatment there was a small but significant increase in energy usage (Delmas et al., 2013). Three months corresponds to day 60 of this study (assuming approximately 20 workdays per month), i.e. only a few days after the significance of the energy savings was established. After that point it was found that the rate of improvement did indeed slow, but it did not reverse.

The fact that energy use did not begin to revert to pre-intervention levels after three months is not, of itself, proof that fatigue did not influence the results. As Figure 14-5 in the previous section suggests, the plateauing that occurred over the second half of the trial (see Figure 11-5) may have been fatigue related because savings picked up again post-trial with the introduction of additional support and coaching initiatives. This response should perhaps have been foreseen because, as Stern (1993) notes, “It is absolutely essential to treat interventions as dynamic and to monitor and revise them continually” (p. 1898).

The daily email messages acted as a motivator for engagement, and they also served as a persistent and effective daily reminder alerting FMs to slippages in their buildings’ performance. As such, even if FMs’ initial enthusiasm levels began to wane after three months (though there was no clear evidence of this), they were able to preserve gains with virtually no additional effort. The use of email as the delivery medium may have been
crucial in this regard. The delivery mechanism (the email inbox) was already an indispensable tool of trade for each FM and with the most important information carried in the daily message’s subject line, bad news could not easily be ignored.

14.4.3. Energy saving becomes habit

There is some evidence that the subjects of this study were, by and large, in the habit of seeking ways to conserve energy even before the intervention study commenced. For example, when considering whether to carry on with the daily feedback and fortnightly meetings post-study, Investa executives surveyed their FMs to determine how much additional time they were spending on energy management as a consequence of the intervention. The results surprised them: they found the average time saving was half an hour per day\(^\text{18}\). This they put down to the reduced effort required to review an email message compared to logging on to BMCS and sub-meter dashboards and analysing raw data, as many FMs claimed to have been doing previously. If FMs were, as they claimed, already in the daily habit of attempting to extract meaningful information from raw energy usage data, their pre-existing habits would have become far more productive when directed to contemplating what actions to take in response to unambiguous daily feedback messages.

This reinforces the findings about the (lack of) fatigue effect discussed above and brings us back to the Theory of Interpersonal Behaviour (TIB) which shows how intention and habit independently contribute to behaviour. The role of deliberate intention declines over time as experience increases and the influence of habit takes over (Triandis, 1977). There are three possible scenarios for this study, all of which point to energy saving becoming

\(^{18}\) The researcher had by this stage resigned from Investa and was not involved in this survey. The results were only reported by Investa verbally. There was no rigorous or reliable means of gauging the time spent on energy management before and during the intervention and this anecdotal finding is presented for interest only.
an entrenched habit. Either FMs did already have constructive (i.e. energy saving / investigation) habits prior to the commencement of the study and the results suggest the intervention made these habits more effective. Alternatively, if habits were not well established prior to the intervention, the intervention gave them an impetus to begin practicing effective energy conservation activities and over a relatively short period of time practice will have become habit due to the convenience of the task of checking their email each day (see chapter 4) – something they most certainly were already in the habit of doing. The third scenario is where FMs may have come with deliberate energy-wasting habits. There is much research to suggest an individual’s habits are greatly influenced by membership of a group (Lewin, 1951; McKenzie-Mohr, 2000) and given the collective results and social learning environment, energy saving behaviour became the social norm.
15. CONCLUSIONS AND RECOMMENDATIONS

This thesis which presents a study of thirty large commercial office buildings in the CBDs of three Australian capital cities, provides new insights into the scale of energy saving and greenhouse gas emission reduction opportunities arising from human behaviour-based intervention techniques.

Non-residential buildings account for a large portion of the energy usage and greenhouse gas emissions attributable to the built environment. And yet the potential for the operators of these buildings to contribute to global efforts to limit climate change through more effective management of energy use is rarely assessed in the literature. Because of a lack of knowledge about the relationships between building technologies and non-technological (i.e. human behaviour-based) factors, and the mechanisms that can promote and sustain operational savings, the scale of potential savings has been unclear to practitioners and policy-makers alike.

The automated daily feedback methodology developed for this study, and the social learning context which was adopted to support the participants, while based on established research and experience from other contexts, were novel in the domain of building operational management and overcame difficulties in recruiting subjects and controlling for variables that have been found to limit studies of non-residential building operations published to-date.

Almost all of the literature that relates to energy consuming behaviours in buildings comes from the residential sector, particularly from studies involving the occupants of residential dwellings. This literature provides valuable insights about feedback and the construction of learning environments that are applicable to operators of non-residential buildings. Likewise, the extensive literature on human motivation and behaviour,
application of feedback methodologies and the transfer of knowledge through social interactions spans many fields, and much of it is directly relevant to the operation of buildings. Yet few published studies have applied behavioural science to the problems of influencing or explaining patterns of energy use in non-residential buildings, and none have applied quantitative methods to assess the reduction potential as this thesis does.

It was found that building operators (i.e. facilities managers), when exposed to automated daily energy performance feedback in a social-learning environment, reduced energy usage relative to their buildings’ normalised baselines. Through rigorous intervention design and application of Triandis’ (1977) Theory of Interpersonal Behaviour, the analysis was able to focus on two independent variables (the facilitating conditions of ‘automated daily feedback’ and a ‘social learning environment’) and one dependent variable: normalised energy use (i.e. energy use relative to that expected under the given external meteorological conditions). A significant divergence between actual and ‘predicted’ daily electricity consumption began to appear after about day 50 of the treatment and a mean improvement of greater than 6 percent was recorded over the entire intervention period. A saving of approximately 10 percent compared to the pre-intervention baseline was observed at the conclusion of the trial across the twenty-two buildings (comprising approximately 620,000 m² or 2.8 percent of the Australian investment-grade office stock) that were included in the final analysis.

The structure and content of the daily messages despatched to building operators met requirements for highly effective feedback identified as important in residential studies but not yet applied in a non-residential context. In particular, the feedback consisted of simple, clear, reliable, high quality messages delivered consistently and frequently; referenced a baseline standard; contained information that was directly relevant to the
recipients; and built motivation, promoted experimentation, and supported constructive
dialogue with peers.

The supportive social learning environment encouraged participants to challenge their
peers in a non-threatening way such that they functioned as models for each other and
subconsciously developed evaluation capabilities (“intelligent self-monitoring”) predicted

These findings are significant for a variety of reasons. The portfolio from which the thirty
trial buildings were drawn was the subject of a mature energy efficiency investment
program that had yielded significant savings over many years, but energy performance
had plateaued during the three years immediately prior to the commencement of this
research project. Its plateauing energy performance combined with the demonstrated
commitment to environmental management reinforced the perception (both internally and
from outside) that there were relatively few ‘easy wins’ still available for FMs to improve
energy performance in this portfolio of office buildings. In other respects, the sample was
considered representative of the Australian investment-grade office building stock.
Therefore the results of this study indicate the potential for similar savings to arise from
similar interventions.

The similarity between the findings reported here and those from some residential studies
reported in the literature may suggest that a distinction is not necessary when it comes to
quantifying the energy saving possibilities from buildings: similar results may reasonably
be expected from other building typologies and for buildings in other regions. However,
caution should be applied when extrapolating the findings of this study and it is
recommended that action research projects be undertaken in many different settings and
with many different building typologies to permit broader generalisation.
The findings have implications for the design of policies and programs aimed at reducing energy use from non-residential buildings, particularly in relation to the roles of motivation, goal setting and peer interaction, the development of skill and competency, and the permanency of savings. They also indicate the potential for achieving energy savings from operational buildings at minimal cost and without resorting to capital-intensive upgrade projects (noting evidence that individual behaviours are heavily influenced by social and institutional contexts and reactions, and hence not easily generalised or up-scaled).

Findings indicate that FMs drew upon their intrinsic motivations to act in response to the treatment rather than any prospect (real or assumed) of exogenous rewards. Furthermore, well-established guidance from the literature regarding the importance of goal- and target-setting was not applied in this study. A target provides the opportunity to highlight a gap between a level of performance and a reference level (the target). The literature suggests feedback directed towards closing of a gap is more effective than merely altering the gap in some way (as was the case here). Research to assess the potential for saving energy from applying various target- and goal-setting methodologies to the management of non-residential buildings is required as this has not been addressed in the research literature to-date.

Without a control group, conclusions about the value of the social learning environment compared to alternatives, such as coaching, can only be tentative as it served the dual purpose of (1) providing the “spark” that is necessary to start off a feedback loop, and (2) avoided the need for interference from the researcher which may have confounded the analysis. Nonetheless, we can say that peers proved to be a beneficial source of formative assessment, and that no additional stimulus (e.g. pressure from management) was
required other than the coordination of the fortnightly peer group meetings. Indications from the post-study continuation phase suggest an evaluation of the relative effectiveness of social learning versus, say coaching, would be a worthwhile area for further research.

We cannot say, based on the findings of this study alone, that the competency or ‘skill’ levels of FMs increased. However, the positive results support the hypothesis that the intervention contributed to the development of evaluative expertise and deliberate self-monitoring. Like motivation, competency is a ‘state of mind’ and was only able to be observed through the interaction between independent and dependent variables. An in-depth understanding of these and other intervening variables would require a qualitative research method and is recommended as a promising area for further research that may yield insights into how to accelerate the rate of skill transfer.

In relation to the question of whether savings are likely to persist, there was no evidence of a fatigue effect causing a decline or reversal in savings during the course of this study which lasted for a period exceeding 220 business days at all buildings – longer than any studies of energy behaviour interventions at multiple non-residential buildings reported in the literature to date. It is likely, therefore, that energy saving behaviours became habit. This suggests a degree of plasticity not previously ascribed to behaviour-based interventions in buildings, and challenges the predominant view among policy-makers that information measures targeted at behavioural change lack the longevity of measures involving investments in physical materials and equipment.

Clearly there is a limit to what can be achieved by non-technological approaches to saving energy in the operation of non-residential buildings, just as there is a limit to what can be achieved by technology alone. However, the line between “technological” and “psychological” is becoming blurred as technological advances make it possible for
feedback to operators to become more precise, insightful and ultimately useful. It is for this reason that the application of psychological and behavioural science to the design and implementation of information-based technologies represents perhaps the most exciting opportunity to drive energy savings and greenhouse gas emission reductions from the operation of non-residential buildings.

END
REFERENCES


219


Zhaojian, L., & Jiang, Y. W. Q. (2007). Tsinghua University, Beijing, China; Survey and analysis on influence of environment parameters and residents' behaviours on air conditioning energy consumption in a residential building. Heating Ventilating & Air Conditioning, 2007(8).
16 December, 2011

Assoc. Prof. William L. Martens
Associate Dean (Graduate Research Studies)
Faculty of Architecture, Design and Planning
Wilkinson Building (G04)
148 City Road
University of Sydney
NSW 2006

Dear Dr Martens,

Permission for Craig Roussac to use Investa Property Group data and information

Craig Roussac is an employee of Investa Property Group (Investa) and since July 2009 he has been undertaking higher degree research through the University of Sydney. His research is closely aligned with his work at Investa as it involves investigating ways to reduce greenhouse gas emissions from the operation of commercial buildings by enhancing the effectiveness of building operations teams.

One of the methods Craig has devised to influence the effectiveness of building operators is to provide them with simple automated messages on a daily basis which indicate whether a building used more or less energy the previous day than weather conditions would have otherwise suggested. If the building used less energy than predicted, building operators are encouraged to repeat whatever it was that contributed to the success. If the building used more than predicted, the operator has the opportunity to make a timely adjustment.

The system relies on data generated by Investa’s electricity supply meters and weather station data purchased by Investa from the Bureau of Meteorology. The messages received by building operators are produced on an Investa-owned data warehousing and transformation system developed by Craig and his colleagues working with a third party software developer. The output data is stored in the system and can be used to evaluate the effectiveness of the action research by analysing for trends.

The purpose of this letter is to confirm that Craig has approval to use Investa’s data for research and publication purposes. Furthermore, he is permitted to use information gathered from Investa employees by way of interview, observation and other methods that are consistent with Investa’s objective of improving the operational performance of our property portfolio, on the understanding that the identities of individual Investa employees will be protected.

Yours sincerely,

Campbell Hanan
Director and Group Executive
Investa Property Group Holdings Pty Limited
APPENDIX B: RDD SCATTER PLOTS FOR ALL 30 BUILDINGS