IEPPEC International Energy Policy & Programme Evaluation Conference

The performance gap of low-carbon building technologies and the role of public policy

David James McElroy, Energy Saving Trust, London, United Kingdom & School of Biological Sciences, University of Sydney, NSW, Australia

Jan Rosenow, Regulatory Assistance Project, Brussels, Belgium & Environmental Change Institute, University of Oxford, United Kingdom

ABSTRACT

In many instances, estimates of both energy savings and renewable energy generation in domestic buildings still rely on engineering models and building energy simulations rather than approaches that are based on measured parameters. And there are good reasons for this - evaluation budgets are often limited, with policy makers needing timely results for reporting purposes, whereas high quality measurement can be complex and is often associated with a substantial time lag. This trend is common despite there being ample evidence of the existence of performance gaps where the actual impact of low-carbon technologies is lower than predicted by models alone. This article adds to this body of evidence by drawing on a range of grey literature evaluations of low-carbon technologies (including energy efficiency measures, renewable heat, and renewable electricity) in the UK household sector. The paper focuses not only on the quantification of the performance gap but also qualitative factors often overlooked, such as installation issues or installer/user behaviour. The article concludes by recommending policy changes including the development of evaluation standards, the experimentation with pay-for-performance programmes, ensuring that installation standards for low-carbon technologies are being enforced, and taking reasonable steps to ensuring that end users are able to use any new technology effectively.

Introduction

Across the globe, efforts are under way to reduce the energy consumption of buildings by increasing the amount of renewable electricity generated on-site and decarbonising heating and cooling systems. For example, in the European Union (EU), the Energy Efficiency Directive is a key driver for improving the energy efficiency of the existing residential building stock (Rosenow et al. 2016). Much of the discussion in the EU has focused on the energy efficiency targets and the types of policies chosen by Member States. Going forward, a robust evaluation, monitoring and verification system will need to be developed in all Member States to ensure that the expected impact of the policies implemented will materialise. This requires not only a sophisticated reporting framework, but also the ability to monitor the achieved impacts with a reasonable degree of certainty. If the impacts of low-carbon technologies are significantly over- or underestimated it is very difficult to track policy impacts reliably and assess whether or not additional efforts are needed to achieve carbon and energy targets. Another potential result of poor evaluation and verification leading to false conclusions regarding low carbon technologies is the erosion of public credibility of these technologies and their ability to deliver energy savings and low-carbon energy (sensu the "Cry-Wolf" effect documented in Breznitz, 1984).

In many instances, estimates of both energy savings and renewable energy generation in domestic buildings still relies on engineering models and building energy simulations rather than approaches that are based on measured parameters. And there are good reasons for this - measurement is costly and often associated with a substantial time lag (Cooney 2017), while evaluation budgets are limited and policy makers require timely results for reporting purposes (Vine et al. 2014). Furthermore, policy makers may be more interested in the physical building performance rather than occupant behaviour. However, there is a large body of literature that demonstrates the existence of a performance gap - the actual impact of low-carbon technologies is often lower than the expected when solely based on models (Adan & Fuerst 2016; Bordass et al. 2001; de Wilde 2014; Gram-Hanssen & Susse 2018; Hamilton et al. 2013; Harold & Lyons 2015; Hong et al. 2006; Johnston et al. 2015 2012; Loucari et al. 2016; Menezes et al. 2012; Scheer et al. 2013). The reasons reported for this are manifold and include poor quality of installation, unrealistic manufacturers' specifications of a technology's performance, unintended user interaction, the use of incorrect models, and inaccurate assumptions on the ex-ante situation (ibid).

This paper adds to the already considerable body of evidence by drawing on a range of unpublished and grey literature evaluations of low-carbon technologies in the household sector carried out by the Energy Saving Trust (EST), a British organization devoted to promoting energy efficiency, energy conservation, and the sustainable use of energy. This data has so far not been used in the peer-reviewed literature and covers a range of technologies including energy efficiency measures, renewable heat, and renewable electricity. Most of the existing literature on the performance gap focuses on energy efficiency and is concerned with the quantification of the scale of the performance gap, although there are now some contributions focusing also on renewable energy (Boyd & Schweber 2018; Frances & Stevenson 2018). This paper has a wider scope that includes renewable energy technologies as well as energy efficiency improvements and focuses not only on the quantification of the performance gap but also qualitative factors often overlooked. The structure of this paper is as follows: (1) a summary of the exiting body of literature on the performance gap; (2) the chosen methodology is detailed, explaining the source of the data and how the selected evaluations were analysed; (3) the results are presented; and (4) the findings are discussed in terms of what they mean for evaluation practice, policy and further research.

Performance gap

Defining the performance gap

The performance gap can be defined as the difference between actual (measured) and predicted (modelled) performance of low-carbon technologies. This is usually discussed in the literature with regard to energy use of the whole building rather than specific elements and mostly in the context of building fabric improvements. With a focus on the construction of new buildings, de Wilde (2014) analyses the root causes of the performance gap and groups them into three categories including:

- causes linked to the design stage including inaccurate modelling of expected performance (de Wilde 2014; Rosenow & Galvin 2013; Sunikka-Blank & Galvin 2012).
- causes linked to the construction stage such as poor construction of buildings and/ or installation of specific measures (Bordass 2004; Bordass et al. 2001; Zero Carbon Hub 2010); and

- causes linked to the operational stage including inappropriate operation of technologies, and unintended user interaction (Branco et al. 2004; Buso et al. 2015; Guerra Santin & Itard 2012 2010; Haas et al., 1999; Linden et al. 2006).

Causes linked to the design stage

Inaccurate modelling of expected performance is a common issue observed in the literature (de Wilde 2014; Menezes et al. 2012). This can be caused by the model itself, but also the people involved in the modelling (Dwyer 2013). For example, a common issue with building models is the overestimation of the ex-ante energy consumption based on unrealistic assumptions. This phenomenon that has been coined the 'prebound effect', such as when residential building models typically overestimate heating energy consumption prior to an intervention by as much as 30% (Sunikka-Blank & Galvin 2012). A reason for this is sometimes that the energy performance of the walls of old buildings is underestimated (Hens et al. 2007). This can lead to overestimates of the amount of energy saved through refurbishment, as householders cannot save energy that was not already being consumed prior to the intervention (Rosenow & Galvin 2013).

Causes linked to the construction stage.

Even if technologies achieve the expected performance under laboratory conditions there is a range of potential factors that can lead to underperformance at the operational stage. Dependent on the installer's level of training and attention to detail, poor workmanship can result in the inappropriate installation of technology which may hamper performance or even cause operational failure. A good example is cavity wall insulation whereby poor distribution of the insulation can result if the cavity is not carefully inspected for areas blocked off by debris (BRE 2016). Another example is in photovoltaics, when panels are overshadowed by trees and as a result generate less renewable electricity (Frances & Stevenson 2018). After the installation of low-carbon technologies regular maintenance is required for some types of technologies. Poor maintenance of technologies can result from insufficient instructions and training (Bordass 2004), thus potentially leading to additional performance issues.

Causes linked to the operational stage

Unfamiliarity with a new technology and old habits may prevent users from interacting with the technology in the way intended. Many examples of this phenomenon exist, such as in Hong et al. (2006) where people continued to use inefficient single room heaters even after a more efficient heating system was installed, or in households that override modern ventilation systems by ventilating the house through opening windows, with strong consequences for performance (Anderson et al. 2013). Or even more simply, is that when controls are misunderstood users may avoid more energy efficient features in favour of incorrect manual control, such as seen in how householders use programmable thermostats (Meier et al. 2011). Occupancy and the number of occupants in particular is a key factor that is often overlooked, although studies clearly show that the importance of this factor for energy use (Guerra Santin et al. 2009).

Significance

For the United Kingdom (UK), there are several studies providing estimates of energy savings and the performance gap from single measures and/ or combinations of those in domestic buildings with some based on actual measured data (Adan and Fuerst 2016; Hamilton et al. 2013; Hong et al. 2006; Johnston et al. 2015 2016) and others on modelled data (Loucari et al. 2016). Similar analysis also exists for Ireland (Scheer et al. 2013; Harold & Lyons 2015). Reviewing the literature, Adan & Fuerst (2016) conclude that "the scope of these studies is limited by the data sets used as they were unable to conduct matching on key explanatory variables, making the results sensitive to latent differences in the characteristics of treatment and control groups." They also point out that these studies focus mainly on the impact of installing a single energy efficiency measure, disregarding the impact of installing a combination of different energy efficiency measures. Whilst there are uncertainties around the data, the existing studies clearly suggest the existence of a performance gap. A study analysing energy demand of new domestic buildings (Johnston et al. 2015 2016) finds that the performance gap can be substantial - the difference between the measured and predicted fabric performance was in some cases greater than 100%. This is supported by other studies: Zero Carbon Hub (2010) found that while some new homes' energy consumption was in line with expected levels in some cases the performance gap exceeded 100% while in assessing solid wall insulation in UK homes, Loucari et al. (2016) state that the performance gap could be as high as 65%.

Methodology

Amongst their work on energy efficiency and conservation, EST and its partners offer a bespoke verification service, by which manufacturers receive independent evaluation of their product's performance claims. Whilst each service is different, depending on the needs and context of each individual product, the verification process broadly comprises of a replicated field trial where quantitative product performance data (e.g. energy or water use and environmental variables such as temperature) are monitored for 12 months from the installation of the technology/start of the trial and interpreted in combination with contextual information, such as the reported behaviour of users. Here the results and insight gained from seven of the most recent field trials done by EST and partners are presented, excluding only a solid wall insulation field trial for which there was no reporting done and an exploration of domestic hot water use, for which no metered data was collected (EST 2015). The field trials presented here cross a range of domestic technologies/upgrades (condensing boilers, heat pumps, solar thermal, solid wall insulation, retrofitting and small scale wind turbines) across three utilities (electricity, gas and water). The specific methodologies for each trial are taken from the grey literature and provided in Table 1. All trials involved monitoring performance for 12 months after the installation/beginning of the trial and some degree of qualitative engagement with installers and users.

Table 1: Methodologies for seven field trials examining various technologies in domestic homes

Citation	Subject	Sample	Data Collected
EMC & EST 2008	Amount of hot water used (and the energy used to heat it)	68 regular boilers and 39 combi boilers in 107 houses of unspecified type	Water use and temperature
Orr et al. 2009	In-situ efficiency of condensing boilers	67 condensing boilers in a mixture of house types including 18 flats with the remainder terrace or stand alone buildings	Energy use, heat output and internal/external temperature
EST 2009	domestic wind turbine performance	38 roof mounted turbines, 19 pole mounted turbines	Energy use, energy generation, wind speed
EST 2011	performance of solar thermal hot water systems	88 terrace or stand alone homes	Solar radiation, solar collector temperature, input/output water temperature, output volume and energy, energy use
Dunbabin & Wickins 2012	Seasonal performance factors of air and ground source heat pumps (Phase 1)	54 ground source heat pumps and 29 air source heat pumps	Energy use, loop temperature, heat and hot water output, internal/external

			temperature
EST 2013	Seasonal performance factors of air and ground source heat pumps (Phase 2)	21 ground source heat pumps and 21 air source heat pumps, 15 of which were also used in Phase 1	Energy use, loop temperature, heat and hot water output, internal/external temperature
TSB 2013	Bespoke retrofitting of properties with renewable technologies and high levels of insulation	100 (37 with quantitative data) terrace or stand alone homes	Energy and water use, internal temperature/humidity. Floor space

Results/Discussion

Performance gap

A performance gap was identified in all seven field trials (Table 2). The largest performance gap was seen for building mounted wind turbines (EST 2015) which often fell short of the commonly quoted load factor of 10%, with no urban or suburban building mounted sites achieving load factors of more than 3%. In some cases performance was so poor that the installation was found to be a net consumer of electricity due to the inverter taking its power from the mains supply while the turbine was idle. In the hot water field trial (EMC & EST 2008) delivery temperature was found to be much lower (52.9°C ± 1.5°C 95% confidence interval for regular boilers and 49.5 ± 2.0°C for combi boilers) than the widely assumed value of 60°C. Furthermore, the average temperature rise of water as it passes through the system was reported as consistently lower (36.7°C) than the 50°C that was assumed in the BRE Domestic Energy Model at the time, leading to predictions being inflated by approximately 35%. Condensing boilers were found to perform closer to prediction than in the hot water field trial, albeit 5.3% less efficient than manufacturers' claimed performance with 75% of boilers recording annual electrical consumption of greater than the BRE standard assessment procedure assumption of 175kWh/year (Orr et al. 2009). For heat pumps, the issue was not whether they could meet the manufacturer's assurances, but that their system efficiencies (the amount of useful heat the heat pump produces compared with the amount of energy used to run the system) were considerably varied (air source heat pump range = 1.2 - 2.2; ground source heat pump range 1.55 - 3.47), meaning that many systems did not perform as well as expected (Dunbabin & Wickins 2012; EST 2013). Similarly, performance in solar thermal systems was found to be extremely variable, with the solar fraction (the amount of energy provided by the solar heat collector divided by the total energy input required, in this case not including heat loss from the primary circuit) provided ranging from 9 - 98% with a median of 39% (EST 2011). This variability was also evidenced in the solar thermal trial by the wide range of parasitic energy consumption (the electricity needed to run the system) from zero (in those systems supported by photovoltaic infrastructure) up to 180kWh per annum (EST 2011). In terms of a performance gap, the delivery temperatures from solar thermal systems were much lower than 60°C at 49.1°C, showing values similar to those observed in the condensing boiler trial (EST 2011). Finally, in the retrofit trial the performance gap manifested through the actual opportunity scope being smaller than assumed in original forecasts (a "prebound" effect, *sensu* Sunikka-Blank & Galvin 2012) based on the information supplied by applicants as part of the application process (TSB 2013). This is because most properties actually had lower air infiltration rates than suggested (most less than 10 $m^3/m^2/hr$ @ 50Pa). Conversely, 13 of the 87 pre-retrofit had considerably higher air infiltration rates (greater than 15 $m^3/m^2/hr$ @ 50Pa) which meant that these properties had greater scope for performance than initially anticipated (TSB 2013).

Table 2: Performance gaps observed in seven field trials of domestic energy efficiency technologies and home improvements.

Citation	Subject	Metric	Performance gap
EMC & EST 2008	Amount of hot water used (and the energy used to heat it)	Delivery temperature[1]	Much lower than the widely assumed value of 60°C: 52.9°C ± 1.5°C for regular boilers; and 49.5 ± 2.0°C for combi boilers.
Orr et al. 2009	In-situ efficiency of condensing boilers	Efficiency	5.3% less efficient than manufacturers' claimed performance; 75% of boilers had an annual electrical consumption of greater than an assumed 175kWh/year
EST 2009	domestic wind turbine performance	Load factor[2]	Technology often fell short of the commonly quoted load factor of 10%, with no urban or suburban building mounted sites achieving load factors of more than 3%.
EST 2011	performance of solar thermal hot water systems	Solar fraction[3]	Considerable variation ranging from 9 - 98% with a median of 39%
		Parasitic energy consumption[4]	Considerable variation ranging from 0 - 180 kWh per annum
EST 2011	performance of solar thermal hot water systems	Delivery temperature[1]	Much lower than the widely assumed value of 60°C: 49.1°C
Dunbabin & Wickins 2012; EST 2013	Seasonal performance factors of air and ground source heat pumps (Phase 1)	System efficiency[5]	Performance was considerably varied: ASHP range = 1.2 - 2.2; GSHP range 1.55 - 3.47
TSB 2013	Bespoke retrofitting of properties with renewable technologies and high levels of insulation	Opportunity scope/air infiltration rates	considerably higher air infiltration rates (greater than 15 $m^3/m^2/hr$ @ 50Pa) which meant that the scope for performance was greater than anticipated

Notes: 1: the output temperature; 2: amount of energy used as a proportion of the total possible energy that could be used ; 3: the amount of energy provided by the solar heat collector divided by the total energy input required, in this case not including heat loss from the primary circuit; 4: the electricity needed to run the system; 5: the amount of useful heat the heat pump produces compared with the amount of energy used to run the system

Contextual factors

According to the field trials presented here, performance gap can be caused when something in the physical environmental of the installation differs from what was expected/modelled, or functions as an entirely unanticipated key factor. For example, discrepancies in the hot water trial were partly caused by the initial cold water feed being cooler than anticipated (EMC & EST 2008). In the solar thermal trial it was found that the solar fraction was also directly affected by the demand on water, which, among other things, is a function of how many other cold water appliances are sharing the same domain (EST 2011). In the least, solar thermal performance is of course affected by insolation, which is dependent on weather, which can be unpredictable (Kreith & Kreider, 1978). Many similar issues were found in the retrofit trial, where performance was indirectly hampered by issues that affected the installation of the new infrastructure including internal and external spaces issues (i.e. accommodation space for larger hot water cylinders, or the need to retain alleyways, bin storage space or delivery access) and delays caused by local phenomena such as bat infestation, asbestos and wet rot (TSB 2013).

Installation factors

The performance gap can also be exacerbated when the quality of the installation is low. All of the boilers installed as part of the condensing boiler trial were oversized by factors ranging from approximately 1.5 to 10, potentially causing flue heat loss and purge losses from cycling to increasingly result in a reduced efficiency (Orr et al. 2009). Orr et al. (2009) suggests that if installation guidance for optimal boiler size is followed, there should be a correlation between installed boiler size and heat demand. Orr et al. (2009) did not observe this pattern, suggesting that the boiler installer likely chooses which size boiler is installed, suggesting that their decision is based on personal beliefs that might be independent of house size, household usage patterns and heat loss. Installation issues were particularly prevalent in the retrofit trial, primarily due to the fact that retrofits tend to need bespoke services of which the installation team has no direct experience (TSB 2013). Nearly a quarter of the homes upgraded as part of the retrofit trial (22 of 100) reported a lack of the skills needed as a challenge for their particular retrofit (TSB 2013).

User factors

Another factor influencing whether the performance gap is realised is in how the technology is ultimately used. If the technology is used incorrectly, a reduction in efficiency and performance is often inevitable. For example, parasitic energy consumption in the solar thermal trial can become prohibitively high (up to 180 kWh/annum) in households that have their pumps at an unnecessarily high setting (EST 2011). In that trial, one particular household was found to have their backup heating source set to input a large amount of energy in the morning. An expert site visitor applied the correct setting but reported that they were confident that the householder would return it to their previous setting, highlighting how the human factor can veto good performance, even in spite of clear instruction and guidance. Similarly, the analysis of electrical and gas consumption of boilers carried out as part of the condensing boiler trial indicated that a key factor in electrical consumption are the pump operating times which are dependent upon the settings of the thermostat, thermostatic radiator valves and/or other controls that can be manipulated by the user (Orr et al. 2009). The literature supports the idea that the human factor is important in how water and energy are used in the home, with consumption being influenced by a range of factors including occupancy and occupants' age, income and space heating preferences (Harlan et al. 2009; House-Peters et al. 2010; Santin et al. 2009), all of which might ultimately differ from the specifications in the designers' energy use models.

Given the role of the human factor in the function of efficient technologies, and that many home owners may not have, either in principle or as a part of their home in particular, knowledge of how a new technology works (which indirectly affects behaviour as reviewed in Huijts et al. 2012) or how much of a resource they actually use (EST 2011), it is likely that undesirable behaviour can explain at least some of any observed performance gap. Ultimately, this highlights the need for a clear, simple and compelling engagement with users during the installation process. In the heat pump trial, customers exhibited varying levels of understanding of how to best use the various controls in order to achieve the best performance from the equipment (EST 2013). Similarly, based on the comments made by householders in the solar thermal trial, the level and quality of the advice given by installers to householders on how to modify their water use in response to the installation did not appear to be very consistent, though this did not seem to affect peoples' satisfaction with the system itself (EST 2011). Thus, the heat pump and solar thermal trials indicated that the inductions received were sometime inadequate, which may also explain some of the performance gap.

Policy recommendations

Accounting for the performance gap in evaluation and policy

The results of the verification services documented here clearly evidence the need for robust and replicated field trials of all new and current technologies where ever they have not yet been done. This will help consumers and manufacturers understand how new technologies work once they have been integrated in to homes, not just in an ideal context (i.e. in the laboratory). At the moment, there is substantial heterogeneity across the world when it comes to evaluating energy savings and renewable energy technologies (see for example Wade & Eyre 2015). This includes practices of how to account for the performance gap. The inconsistent approach to measuring energy savings and monitoring and verification leads to considerable uncertainties as to whether the benefits of energy efficiency and renewable energy policies will materialise to the extent anticipated by policy makers.

In the energy efficiency field, following the implementation process of the Energy Services Directive in 2006, similar issues were discussed in the literature (Boonekamp 2006; Thomas et al. 2012). There are also detailed global standards for monitoring and verification such as the International Performance Measurement and Verification Protocol (IPMVP; EVO 2012) that address the performance gap, but those standards are mainly being used for larger projects rather than in the residential sector. Given that manufacturers' estimates can be wrong (as seen in the performance gaps presented here) and that the methodologies that manufacturers use to predict performance can vary so widely (for example, the power curves and ratings of turbines presented in EST 2009 were calculated using different methods) it is virtually impossible for customers to compare the performance of different products and make an informed decision on how to invest. This literature can form the basis of a clear and consistent approach to monitoring and verification of energy savings across the EU. Rather than specifying in detail how exactly the performance gap should be accounted for, a set of high-level principles for policy evaluation would be a good starting point. Such high-level principles would need to set out the key parameters to be analysed in evaluations, quality standards for carrying out monitoring of low-carbon technologies once installed, key aspects to be covered by post-installation audits, and appropriate methods for evaluation, monitoring and verification. In the United States, this is a common approach (e.g. TecMarket Works Framework Team 2004). If applied, such principles would ensure that awareness for the performance gap is increased and, over time, policy evaluations account for it more systematically.

Pay-for-performance

In recent years there has been accelerated experimentation with so called "pay-for-performance" energy efficiency programmes (Szinai et al. 2017), although mainly in the United States and particularly in California. Pay-for-performance programmes reward actual measured energy savings based on metered energy consumption. Many, but not all, of such approaches evaluate savings using some form of meter data or utility bill data collected before and after an energy efficiency intervention. Payments of any subsidy are linked to the savings achieved over time - this minimises the risk of providing upfront payments for technological interventions that do not deliver energy savings in line with expectations. Traditionally, energy efficiency programmes reward specific technologies assuming a certain amount of savings. The incentives offered do not differentiate between the actual outcomes achieved. In a world of increasing digitalisation and with more abundant energy data (for example from smart meters) and new innovative monitoring and verification methods discussed under the heading M&V 2.0 (Franconi et al. 2017), pay-for-performance programmes provide a promising policy approach to reward real savings rather than just modelled savings. By definition, such a framework would set incentives to reduce the performance gap. So far, most of the savings delivered through pay-for-performance programmes stem from interventions on the commercial, public, and industrial sectors (Szinai et al. 2017). However, with the costs for both obtaining and analysing energy data come down there is the potential to also use pay-for-performance programmes in the residential sector. First examples of pay-for-performance programmes targeting residential buildings have begun to emerge in Europe: Germany recently launched the *Einsparzaehler* (energy savings meter) programme, a pilot scheme to test the feasibility of a pay-for-performance programme in Germany. Policy makers should continue to experiment with such instruments given the increasing coverage of buildings with smart meters. Finally, innovative retrofit programmes such as *Energiesprong* (Visscher et al. 2016) use performance guarantees to ensure the performance of the improvements over a long-term (minimum 30-year) period. This is based on the Dutch net zero-energy renovation concept (Rovers 2014), which has been applied in social housing. Tenants pay an increased rent but are not charged any energy costs at all after retrofit if their energy consumption remains within certain specified bounds. This approach only works if the performance gap is small. The downside of this approach is that the incentive to save energy for monetary reasons no longer applies if tenants stay within the defined bounds.

Installer standards and user training

A recurring theme throughout these field trials is the role that the installer plays in the performance gap with the domestic wind turbine trial pointedly concluding that industry standards must be agreed and implemented so that customers can realistically assess the potential for building mounted turbines to generate energy (EST 2009). In the first phase of the heat pump trial, heat pump performance was found to be dependent on the specification, design, installation and commissioning practices (Dunbabin & Wickins 2012), which led to a thorough review of installation and training guidance and the eventual revision of the Microgeneration Certification Scheme (MCS) installer standards (an industry-led and nationally recognised quality assurance scheme) to good effect (EST 2013). Generally, customers and policy makers should treat all novel technologies with caution, in particular domestic scale wind, until such a time that the product they are considering receives MCS accreditation or equivalent (EST 2009). The content of these accreditation schemes should be considered particularly carefully in terms of installer training, so that the pitfalls of having unprepared installers as observed in the more bespoke retrofitting projects (TSB 2013) are likely to be seen less often, assuming that the effectiveness of accreditation schemes documented for performance in engineering (Volkwein et al. 2007) translate to the retrofitting context. This is particularly relevant for more complex technologies such as heat pumps, where short courses without formal educational qualifications are simply not good enough (Gleeson 2016). Furthermore, minimum standards of user engagement and induction beyond a user manual should also be included if efficiency claims are to be allowed. This is because quality advice and understanding can help minimise the performance gap as seen here in the solar thermal trial where better advice on heating patterns, use of back up heating and the correct function of controls might have helped householders make the most of the new system (EST 2011) and encourage householders to capitalise on other low carbon opportunities (Owen et al. 2014).

Conclusions

The evaluations presented here corroborate with the literature in concluding that the performance gap is a common phenomenon that can have considerable impact on the ultimate effectiveness of energy efficient technologies. It is clear that structured guidance on how to measure the performance of products is needed to assist manufacturers in providing accurate and transparent advice to consumers on what technologies are appropriate for achieving any desired outcome. Such an evaluation framework, whether encouraged through regulation or an optional standard (such as IPMVP;

EVO 2012) should aspire towards international application and contain enough methodological flexibility to evaluate any technologies designed to improve the efficiency of buildings. Secondly, similar action should be taken to ensure that a) manufacturers develop installer standards for their products and b) end users are engaged with proper guidance on how to use the technology effectiveness. Until steps such as these or similar are taken the potential for technological innovation to help society achieve a sustainable future will continue to be undermined by the performance gap. Finally, innovative approaches such as pay-for-performance programmes deserve to be tested also in jurisdictions outside of the United States and in the residential sector. Rewarding energy savings based on standardised analysis of metered data potentially offers a policy solution to the complex evaluation, monitoring and verification problems that the performance gap creates.

References

Adan and Fuerst 2016 Do energy efficiency measures really reduce household energy consumption? A difference-in-difference analysis. *Energy Efficiency* 9(5), 1207–1219

Andersen, R., Fabi, V., Toftum, J., Corgnati, S.P., Olesen, B.W. 2013. "Window opening behaviour modelled from measurements in Danish dwellings". *Building and Environment* 69: 101-113.

Bordass, B., Cohen, R., Standeven, M., Leaman, A. 2001. "Assessing building performance in use 3: Energy performance of the Probe buildings." *Building Research & Information* 29(2): 114–128.

Bordass B., Cohen R., Field J. 2004. "Energy performance of non-domestic buildings – closing the credibility gap." In: *International conference on improving energy efficiency in commercial buildings*. Frankfurt, Germany.

Boyd, P., Schweber, L. 2018. "Unintended consequences: institutional artefacts, closure mechanisms and the performance gap." *Building Research & Information* 46(1): 10-22.

Branco, G., Lachal, B., Gallinelli, P., Weber, W. 2004. "Expected versus observed heat consumption of a low energy multifamily complex in Switzerland based on long-term experimental data" *Energy and Buildings* 36: 543–555.

Breznitz, S. 2013. *Cry wolf: The psychology of false alarms*. Hillsdale, USA: Psychology Press.

BRE. 2016. *Post Installation Performance of Cavity Wall & External Wall Insulation. Constructing Excellence in Wales.* Available at: http://www.cewales.org.uk/files/2414/7505/5257/Post_Installation_Performance_05_FINAL_ENG.pdf [accessed 15/11/2017]

Buso, T., Fabi, V., Andersen, R.K., Corgnati, S.P. 2015. "Occupant behaviour and robustness of building design." *Building and Environment* 94: 694–703.

de Wilde, P. 2014. "The gap between predicted and measured energy performance of buildings - A framework for investigation." *Automation in Construction* 41: 40-49.

Dunbabin, P., Wickins, C. 2012. *Detailed analysis from the first phase of the Energy Saving Trust's heat pump trial: evidence to support the revision of the MCS Installer Standard MIS 3005 Issue 3.1*. London: Energy Saving Trust.

Dwyer, T. 2013. "Knowledge is power: benchmarking and prediction of building energy consumption." *Building Services Engineering Research and Technology* 34(1): 5-7.

Energy Monitoring Company, Energy Saving Trust. 2008. *Measurement of domestic hot water consumption in dwellings*. London: Energy Saving Trust.

Energy Saving Trust. 2009. *Location, location, location: Domestic small-scale wind field trial report*. London: Energy Saving Trust.

Energy Saving Trust. 2011. *Solar Thermal Field Trial Scientific Report*. London: Energy Saving Trust.

Energy Saving Trust. 2013. *The Heat is on: Heat pump field trials phase 2*. London: Energy Saving Trust.

Energy Saving Trust. 2015. *At Home With Water 2: Technical Report*. London: Energy Saving Trust.

Efficiency Valuation Organisation. 2012. *International Performance Measurement and Verification Protocol: Concepts and options for determining energy and water savings Volume 1*. Washington: Efficiency Valuation Organisation.

Frances, Z., Stevenson, F. 2018. “Domestic photovoltaic systems: the governance of occupant use.” *Building Research & Information* 46(1): 23-41.

Franconi, E., Gee, M., Goldberg, M., Granderson, J., Guiterman, T., Li, M., Smith, B.A. 2017. *The Status and Promise of Advanced M&V: An Overiew of "M&V 2.0" Methods, Tools, and Applications*. Rocky Mountain Institute, 2017 and Lawrence Berkeley National Laboratory, 2017. LBNL report number #LBNL-1007125.

Gleeson, C.P. 2016. “Residential heat pump installations: the role of vocational education and training.” *Building Research & Information* 44(4): 394-406.

Gram-Hanssen, K., Susse, G. 2018. “Energy performance gaps: promises, people, practices.” *Building Research & Information* 46(1): 1-9.

Guerra Santin, O., Itard, L. 2010. “Occupants’ behaviour: determinants and effects on residential heating consumption.” *Building Research and Information* 38(3): 318–338.

Guerra Santin O., Itard L. Visscher H.J. 2009. “The effect of occupancy and building characteristics on energy use for space and water heating in Dutch residential stock.” *Energy and Buildings* 41: 1223–1232.

Haas, R., Auer, H., Biermayr, P. 1998. „The impact of consumer behaviour on residential energy demand for space heating.” *Energy and Buildings* 27: 195–205.

Hamilton, I. G., Steadman, P. J., Bruhns, H., Summerfield, A. J., Lowe, R. 2013. “Energy efficiency in the British housing stock: energy demand and the homes energy efficiency database.” *Energy Policy* 60: 462–480.

Harlan, S. L., Yabiku, S. T., Larsen, L., Brazel, A. J. 2009. „Household water consumption in an arid city: affluence, affordance, and attitudes.” *Society and Natural Resources* 22(8): 691-709.

Harold, J., Lyons, R. 2015. “The determinants of residential gas demand in Ireland.” *Energy Economics* 51, 475–483.

Hens, H., Janssens, A., Depraetere, W., Carmeliet, J., Lecompte, J. 2007. “Brick Cavity Walls: A performance analysis based on measurements and simulations.” *Journal of Building Physics* 31: 95–124.

Hong, S. H., Oreszczyn, T., Ridley, I. 2006. “The impact of energy efficient refurbishment on the space heating fuel consumption in English dwellings.” *Energy and Buildings* 38(10): 1171–1181.

House-Peters, L., Pratt, B., Chang, H. 2010. “Effects of urban spatial structure, sociodemographics, and climate on residential water consumption in Hillsboro, Oregon.” *Journal of the American Water Resources Association* 46(3): 461-472.

Huijts, N. M., Molin, E. J., Steg, L. 2012. “Psychological factors influencing sustainable energy technology acceptance: A review-based comprehensive framework.” *Renewable and Sustainable Energy Reviews* 16(1): 525-531.

Johnston, D., Miles-Shenton, D., & Farmer, D. 2015. “Quantifying the domestic building fabric ‘performance gap’.” *Building Services Engineering Research and Technology* 36(5), 614-627

Johnston, D., Farmer, D., Brooke-Peat, M., & Miles-Shenton, D. 2016. “Bridging the domestic building fabric performance gap.” *Building Research & Information* 44(2): 147-159

Kreith, F., Kreider, J.F. 1978. *Principles of solar engineering*. Washington: Hemisphere Publishing Corporation.

Linden, A.L., Carlsson-Kanyama, A., Eriksson, B. 2006. “Efficient and inefficient aspects of residential energy behaviour: what are the policy instruments for change?” *Energy Policy* 34: 1918–1927.

Loucari, C., Taylor, J., Raslan, R., Oikonomou, E., Mavrogianni, A. 2016. “Retrofit solutions for solid wall dwellings in England - The impact of uncertainty upon the energy performance gap.” *Building Services Engineering Research and Technology* 37(5): 614-634.

Meier, A., Aragon, C., Peffer, T., Perry, D., Pritoni, M. 2011. “Usability of residential thermostats: Preliminary investigations.” *Building and Environment* 46(10): 1891-1898.

Menezes, A. C., Cripps, A., Bouchlaghem, D., Buswell, R. 2012. “Predicted vs. actual energy performance of non-domestic buildings - Using post-occupancy evaluation data to reduce the performance gap.” *Applied Energy* 97: 355-364.

Orr, G., Lelyveld T., Burton, S. 2009. *Addendum report: In-situ monitoring of efficiencies of condensing boilers and use of secondary heating*. Cheltenham, UK: Gastec at CRE Ltd.

Owen, A., Mitchell, G., Gouldson, A. 2014. "Unseen influence—The role of low carbon retrofit advisers and installers in the adoption and use of domestic energy technology." *Energy Policy* 73: 169-179.

Rosenow, J., Galvin, R. 2013. "Evaluating the Evaluations: evidence from energy efficiency programmes in Germany and the UK." *Energy & Buildings* 62: 450–458.

Rosenow, J., Leguijt, C., Pato, Z., Fawcett, T., Eyre, N. 2016. "An ex-ante evaluation of the EU Energy Efficiency Directive - Article 7." *Economics of Energy & Environmental Policy* 5(2): 45-63.

Rovers, R. 2014. "New energy retrofit concept: 'Renovation trains' for mass housing." *Building Research and Information* 42(6): 757–767.

Santin, O. G., Itard, L., Visscher, H. 2009. "The effect of occupancy and building characteristics on energy use for space and water heating in Dutch residential stock." *Energy and buildings* 41(11): 1223-1232.

Scheer, J., Clancy, M., Hógáin, S.N. 2013. "Quantification of energy savings from Ireland's home energy saving scheme: an ex post billing analysis." *Energy Efficiency* 6(1): 35–48.

Sunikka-Blank, M., Galvin, R. 2012. "Introducing the prebound effect the gap between performance and actual energy consumption." *Building Research & Information* 40(3): 260-273.

Szinai, J., Borgeson, M., Levin, E. 2017. *Putting your money where your meter is. A study of pay-for-performance energy efficiency programs in the United States. Study prepared for the Natural Resources Defense Council and Vermont Energy Investment Corporation*. Available from: www.nrdc.org/resources/putting-your-money-where-your-meter [accessed 15/11/2017]

TecMarket Works Framework Team. 2004. *The California evaluation framework*. San Francisco: California Public Utilities Commission.

The Technology Strategy Board. 2013. *Retrofit Revealed: The Retrofit for the Future projects - data analysis report*. Swindon, UK: The Technology Strategy Board.

Vine, E., Sullivan, M., Lutzenhiser, L., Blumstein, C., Miller, B. 2014. "Experimentation and the evaluation of energy efficiency programs." *Energy Efficiency* 7: 627-640.

Visscher, H., Meijer, F., Majcen, D., Itard, L. 2016. "Improved governance for energy efficiency in housing." *Building Research & Information* 44(5-6): 552-561.

Volkwein, J. F., Lattuca, L. R., Harper, B. J., Domingo, R. J. 2007. "Measuring the impact of professional accreditation on student experiences and learning outcomes." *Research in Higher Education 48*(2): 251-282.

Wade, J., Eyre, N. 2015. *Energy Efficiency Evaluation: The evidence for real energy savings from energy efficiency programmes in the household sector. A report by the UKERC Technology & Policy Assessment Function*. Available from: www.ukerc.ac.uk/asset/EEBD91B1-C5A0-4F77-BF8B90201FF8A2C7 *[accessed 15/11/2017]*

Zero Carbon Hub. 2010. *Closing the Gap Between Designed and Built Performance*. London: Zero Carbon Hub.