Routledge
Taylor & Francis Group

Building Research & Information

ISSN: 0961-3218 (Print) 1466-4321 (Online) Journal homepage: https://www.tandfonline.com/loi/rbri20

Building evaluation: practice and principles

Adrian Leaman , Fionn Stevenson & Bill Bordass

To cite this article: Adrian Leaman , Fionn Stevenson & Bill Bordass (2010) Building evaluation: practice and principles, Building Research & Information, 38:5, 564-577, DOI: 10.1080/09613218.2010.495217

To link to this article: https://doi.org/10.1080/09613218.2010.495217

Building Research & Information (2010) **38**(5), 564–577

Routledge
Taylor & Francis Group

INFORMATION PAPER

Building evaluation: practice and principles

Adrian Leaman[1,2], Fionn Stevenson[3] and Bill Bordass[2,4]

[1] Arup, 13 Fitzroy Street, London W1T 4BQ, UK

[2] The Usable Buildings Trust
E-mail: adrianleaman@usablebuildings.co.uk

[3]School of the Built Environment, Oxford Institute for Sustainable Development, Oxford Brookes University, Gipsy Lane, Headington, Oxford OX3 0BP, UK
E-mail: fstevenson@brookes.ac.uk

[4]William Bordass Associates, 10 Princess Road, London NW1 8JJ, UK
E-mail: bilbordass@aol.com

Based on experiences of carrying out building-performance studies in non-domestic buildings in the United Kingdom and around the world, the question is addressed of how these might apply in the emerging area of housing evaluation studies. Principles are offered covering both non-domestic and domestic buildings. The research area and approach are defined, and types of feedback and their effectiveness are explored, along with the sorts of methods that should be used and some wider topics including duty of care and some of the implications of 'real-world' research. Key lessons from fieldwork are presented.

Keywords: building evaluation, building performance, energy assessment, feedback, occupant surveys, post-occupancy evaluation, real-world research

Sur la base de l'expérience acquise en réalisant des études sur les performances des bâtiments dans des bâtiments non résidentiels au Royaume-Uni et à travers le monde, la question est abordée de savoir comment celle-ci pourrait s'appliquer au secteur émergent des études d'évaluation des logements. Des principes sont proposés, couvrant à la fois les bâtiments résidentiels et non résidentiels. Le domaine de recherche et l'approche sont définis, et les types de feedback comme leur efficacité sont examinés, ainsi que les types de méthodes qui devraient être utilisés, mais aussi certains thèmes plus larges, incluant le devoir de diligence et certaines des implications découlant des recherches en «situation réelle». Les principaux enseignements de ces travaux de terrain sont présentés.

Mots clés: évaluation des bâtiments, performances des bâtiments, évaluation énergétique, feedback, enquêtes auprès des occupants, évaluation après occupation, recherches en situation réelle

Introduction

> For the last thirty years, in much of the English-speaking world … when asking ourselves whether we support a proposal or initiative, we have not asked, is it good or bad? Instead we inquire: Is it efficient? Is it productive?
>
> (Judt, 2009)

Questions of efficiency and productivity are the usual subtext of post-occupancy evaluation studies of building performance, hereafter called 'building evaluation' for short. Given a design brief, which should state how a building is intended to perform in practice, how well does it actually work? There are at least three perspectives:

- occupants, and how well their needs are met
- environmental performance, normally energy and water efficiency

Building Research & Information ISSN 0961-3218 print/ISSN 1466-4321 online
http://www.informaworld.com/journals
DOI: 10.1080/09613218.2010.495217

- whether the building makes economic sense, such as value for money or return on investment

Surprisingly, hardly any modern buildings do well in all three categories (Standeven *et al.*, 1998).[1] In fact, many of them perform poorly, so people are embarrassed to publish the results. As a result, designers do not learn from past mistakes, and managers and others commissioning buildings unwittingly help perpetuate the same mistakes because they do not know any better. Therefore, although the goal is often improved efficiency and productivity, the effect can easily be the opposite.

Wherever one looks in building-performance studies, one tends to see under-achievement. This is not because researchers are unduly gloomy, but because this is what they find in reality. Of course there are exceptions, outstanding exemplars, but often it is not known which these are because buildings are not routinely monitored, so good cases remain unnoticed. Award-winning buildings are no guide either, as identified by Brand (1994): the received wisdom of the judges often disguises the true picture, as in the case of the celebrated Laban Dance Centre in Deptford, London (Pearson, 2007). The current authors are aware of many others through our own studies and through material sent to us by others, most unsurprisingly unpublished. The difficulty of getting new concepts to work is underlined by Short *et al.* (2009).

Building performance evaluation is not just about the 'holy grails' of efficiency and productivity. There are other considerations. To continue the above quotation:

> This propensity to avoid moral considerations, to restrict ourselves to issues of profit and loss – economic questions in the narrowest sense – is not an instinctive human condition. It is an acquired taste.
>
> (Judt, 2009)

In building performance work, many such 'other' considerations are hidden, taken for granted, or just too difficult to handle, so they are put on one side. The pursuit of quantification obscures qualification. What about contexts and individual circumstances? What about design quality? What about perceived value? How best is the public interest served in the face of commercial self-interest? Where does duty of care to individual building users fit in, or indeed to the wider considerations of sustainable development? How can individual needs of users at the extreme ends of, say, the comfort spectrum be dealt with? To address these questions, the authors try to develop some of these points as best we can in relation to our own experience of carrying out such studies. We are not attempting to be comprehensive. We are simply trying to elaborate on an area that has received relatively little attention in the past.

The perspective used is building evaluation in the widest sense, *i.e.* non-domestic buildings as one broad group, and domestic buildings as the other, usually single-family dwellings, but also, for example, residential homes for the elderly. By far the greatest effort so far has gone into non-domestic building studies; and especially offices and educational buildings. Far fewer housing studies have been carried out, although this is now beginning to change (Wingfield *et al.*, 2009; Shipworth *et al.*, 2010), stimulated by recent environmental monitoring programmes funded largely by government and charities, and by the support of journals such as this one.

What kind of research?

Building evaluation work does not fit long-standing classifications: it spans the professions (architecture, services engineering and facilities management being the most prominent); it is multidisciplinary, often to a confusing extent (design, psychology, economics, planning, sociology, engineering, etc.); it draws on laboratory research and physical measurement, but it is predominantly about empirical field work, visiting and studying real buildings in use and talking to real people.

This makes it, to use Colin Robson's term, 'real-world research' (Robson, 2002). Figure 1 sets out some of its main features, developed further below.

- Solving problems
  Building evaluation monitors performance to discover and then try to solve problems. This is not knowledge for its own sake, but knowledge with results aimed at helping designers and managers make more informed decisions to help improve the building being studied and, of course, to spread the knowledge further to improve future buildings.

- Predicting effects
  It helps understand consequences, especially where complex systemic processes are at work. Some of these are serendipitously good – 'emergent'. The most serious are unintended. However, many of them are predictable given sufficient and shared knowledge of preconditions and contexts. Others are more difficult to predict, *e.g.* unanticipated operating modes for innovative technologies, or alterations in user behaviour.

- Robust results
  Building-evaluation studies cannot directly control inputs as in laboratory science. Inputs are givens:

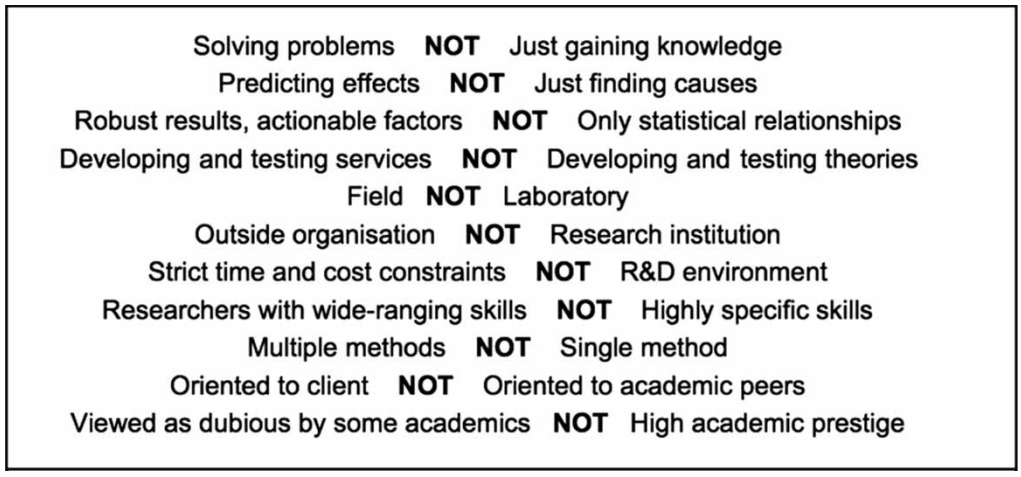

Figure 1 Building-performance studies are an example of real-world research. Source: adapted from Robson (2002), p. 12 (box 1.1). R&D, research and development

the circumstances of the building, its operation and context. The methods employed must give repeatable, believable, convincing results so that they satisfy scientific scrutiny. However, this is not experimental science, so certain criteria need to be relaxed to allow for its real-world character.[2]

- Developing services
  Which services depends on who is carrying out the work and why they are doing it. Motives differ from genuine attempts to improve performance to marketing ploys. It is worth remembering that the first question one should ask of any research project is 'Who is paying?' because there will usually be strings attached, often invisible.

- Field not laboratory
  A building-evaluation study comes with the implied promise to the occupants that if is there is a problem with the building something will be done about it.

- Outside organization
  Whenever a study is carried out, the researcher always needs the cooperation of the occupiers. Getting access is often the hardest part. Access is becoming increasingly difficult because of growing security concerns. Developing a dialogue with the on-site facilities manager or owner is often the key to success. Security clearance is now more onerous.

- Time/cost constraints
  The shorter the time spent in the building, and the less you disturb the occupants, the better. However, problems thrown up by initial observations are often most easily resolved by on-site observation, discussion and measurement.

- Wide-ranging skills
  As building evaluation is multidisciplinary, it follows that its practitioners must draw on a range of specialisms. Practitioners do not, however, need to specialize in all of them. We have found that the most effective teams tend to have members who are familiar with design processes, but not necessarily themselves designers. Teams benefit from an environmental services perspective, on the one hand, and a human needs perspective, on the other hand, a balance which is rarely achieved. Designers are not the best people to carry out building-evaluation studies on their own buildings, but can be enlightening in relation to other buildings, providing certain professional agendas are set aside. Practitioners will also need to be statistically literate, with well-developed communication skills.

- Multiple methods
  The authors have found the Post-occupancy Review of Buildings and their Engineering (Probe)[3] approach to building evaluation to be a valuable core for all studies, a package of techniques which starts with an energy-assessment technique[4] and an occupant survey,[5] supplemented, if appropriate, with a fabric air-tightness test[6] and a water consumption analysis.[7] This can then be further augmented by diagnostic techniques. The authors have developed a 'feedback techniques portfolio' that signposts 25 of them.[8] Core techniques used should be tried and tested, and with a set of empirical benchmarks derived from previous studies available to allow comparisons to be made. There are also many other approaches to building evaluation that are well established and widely taken up. This journal has published many of them.[9]

- Client oriented
  Reports from building evaluations will go to people who will not necessarily understand, or be comfortable with, jargon or specialist language. On the other hand, over-simplification, especially with statistical graphics, may create even more problems by disguising or misleading. We advocate 'no unanswered questions', so everything that is included in the reports and graphics is carefully, but not long-windedly, explained. Visual images and diagrams that pinpoint a clear conclusion are vital. It is also usually better to split the findings into several sub-reports, with an overview report aimed at the wider, non-specialist audience.[10] This also means that clients can choose which parts they wish to release to others.

- Dubious to academics
  One reason for slow progress with building evaluation is that some academics regard case studies that draw on a variety of material and methods as being either too challenging or merely anecdotal. Our experience is the opposite. There is nothing better than a vivid case study to communicate lessons learned and underwrite decision-making.

What sort of feedback?

Figure 2 shows five types of feedback: from initial justification to normal use in the life cycle of a building. The broad arrow at the top could also be shown as a spiral, indicating how the experiences gained from one project should ideally be carried on to projects in the future. Building evaluation fits into levels 2 and 3, labelled *Hindsight* and *Foresight*. In theory, levels 1–3 should also feed in to level 4, *Knowledge management*, but this is rarely achieved successfully partly because of the fragmentation of activities into professional sub-specialisms. Academic activity is usually at levels 5 and 6, thus more removed from design and management processes and decisions at levels 1 and 2. Academics tend not to recognize post-occupancy evaluation as 'research' because its findings do not always produce 'new' knowledge. There is a particular problem with the use of the word 'innovative' here. Evaluation studies often go unfunded because they are perceived as 'more of the same'. The underlying techniques may be the same, but innovation is in the way feedback, lessons learned, and quality-control procedures are embedded into future design and construction practice. Client requirements, practice, techniques, and technologies move forward and individual cases always produce something new, albeit in a standard analysis framework. Building evaluation also has an *Insight* role where the results can be applied directly to the work being done and building being studied, as, for example, in the *Aftercare* phase of Soft Landings (see below). A good example of this The National Trust headquarters building in the UK, where insight from the building evaluation helped fine tune building performance and reduce energy consumption (Nevill, 2007).

Feedback potentially falls into four types:

- *making the case*: the project objectives and the brief

- *the design and building process*: including appointments, design, project management, construction, coordination, cost control, build quality, commissioning and handover

- *the building as a product, the outputs*: what it is like, what it costs, its fitness for purpose, and how professionals and public react to it

- *the building's performance in use, the outcomes*: technical, for the occupier, for users, financial, operational and environmental

This list presents a formidable array of requirements. In practice, the third and especially the fourth points are what are now taken to be 'post-occupancy evaluation'. The expense of a retrospective study of the first and second points partly accounts for their relative rarity. So too does muddle over whether 'design' and the 'design process' is being studied, or whether it is the building in its final form, or the end product in use. Ideally, it would be all of these. In practice, it has proven much more realistic to concentrate on the building's performance in use. That said, many of the conclusions from performance-in-use studies have profound implications for briefing, design, construction, commissioning and handover; and the associated products and services. In fact, conclusions from performance studies point to the need to revisit and reform just these areas. The Soft Landings initiative[11] is an example of how feedback is being incorporated into processes with new professional support systems so that the lessons learned from previous evaluation studies become 'embedded' in design decision-making processes. Soft Landings tries to close feedback loops, with the ultimate aim of improving the quality of process and product, and, of course, performance in use. There is an element of self-fulfilling prophecy here: one wants to identify the right types of feedback to create the most effective performance improvements. This begs knowing: 'What is the most effective type of feedback?'.

What is the most effective type of feedback?

In our experience, nothing betters case studies of named buildings backed by thorough data collection,

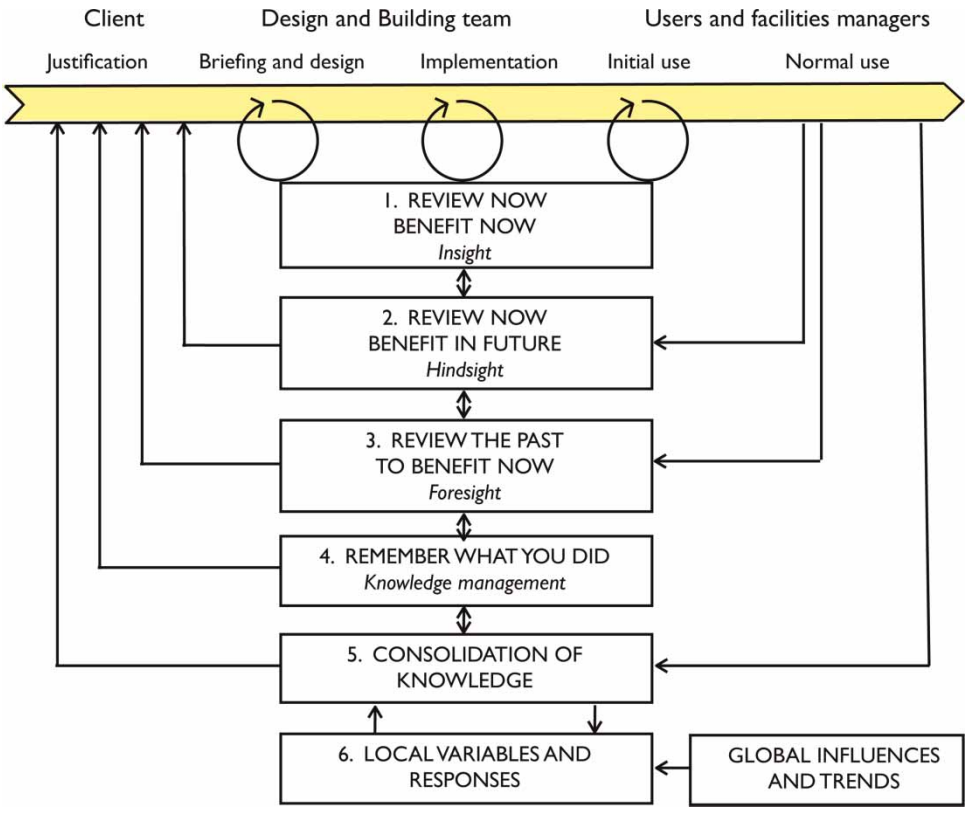

Figure 2 Five types of feedback

benchmarked against a national sample, finishing with a list of lessons learned, preferably including reflections on the results by the parties directly involved, and especially the design team. Circumstances should be clearly explained so readers can judge for themselves the likely effects of any influential factors current at the time. Ideally, cases should tell a meaningful story with some surprises, and be written up in a balanced, non-judgmental style.

Unfortunately, case studies by themselves still do not carry much weight with academic researchers unless there are enough of them to provide some sort of statistical data with generalizable outcomes. This results in building performance research that investigates only a few factors across a broad study but without the depth or understanding of a case study. A major study in the UK (Carbon Reduction in Buildings–CaRB) has recently investigated a number of factors across a broad study of energy use in housing, but even here, findings concerning the occupants have been relatively limited due to the lack of qualitative in-depth data typically found in a case study (Shipworth *et al.*, 2010). We believe this is a short-sighted approach and that research funders need to recognize the incremental value of case studies for building-up effective data. A single case study will nearly always throw light on new issues (as well as reminding one of old ones still in need of attention!) and create hypotheses that can be tested in other ways, *e.g.* on other case studies; in discussions with design and building teams; and in helping to structure new research. A good example of this is the case study of a new prototype 'zero-carbon' house in the UK (Stevenson and Rijal, 2008; Stewart Milne Group, 2009). For the first time in the UK a national house-builder has been prepared to publish 'bad news' occupancy feedback to help the industry move forward. It has directly informed the design of major new research by three national house-builders in the UK with significant government support.[12]

Studies are much less useful if cases are anonymous with findings cherry-picked to include only favourable outcomes and avoid embarrassment. In some cases 'research' has been exploited as a disguised form of marketing. Worse still, if results are misleading, especially where statistics are concerned, they are useless. Figure 3 is an example of an effective reporting of energy statistics from individual case studies using standard definitions and protocols. These are end-use consumption data from energy surveys of schools and university buildings with associated benchmarks, and using carbon factors which are clearly reported.

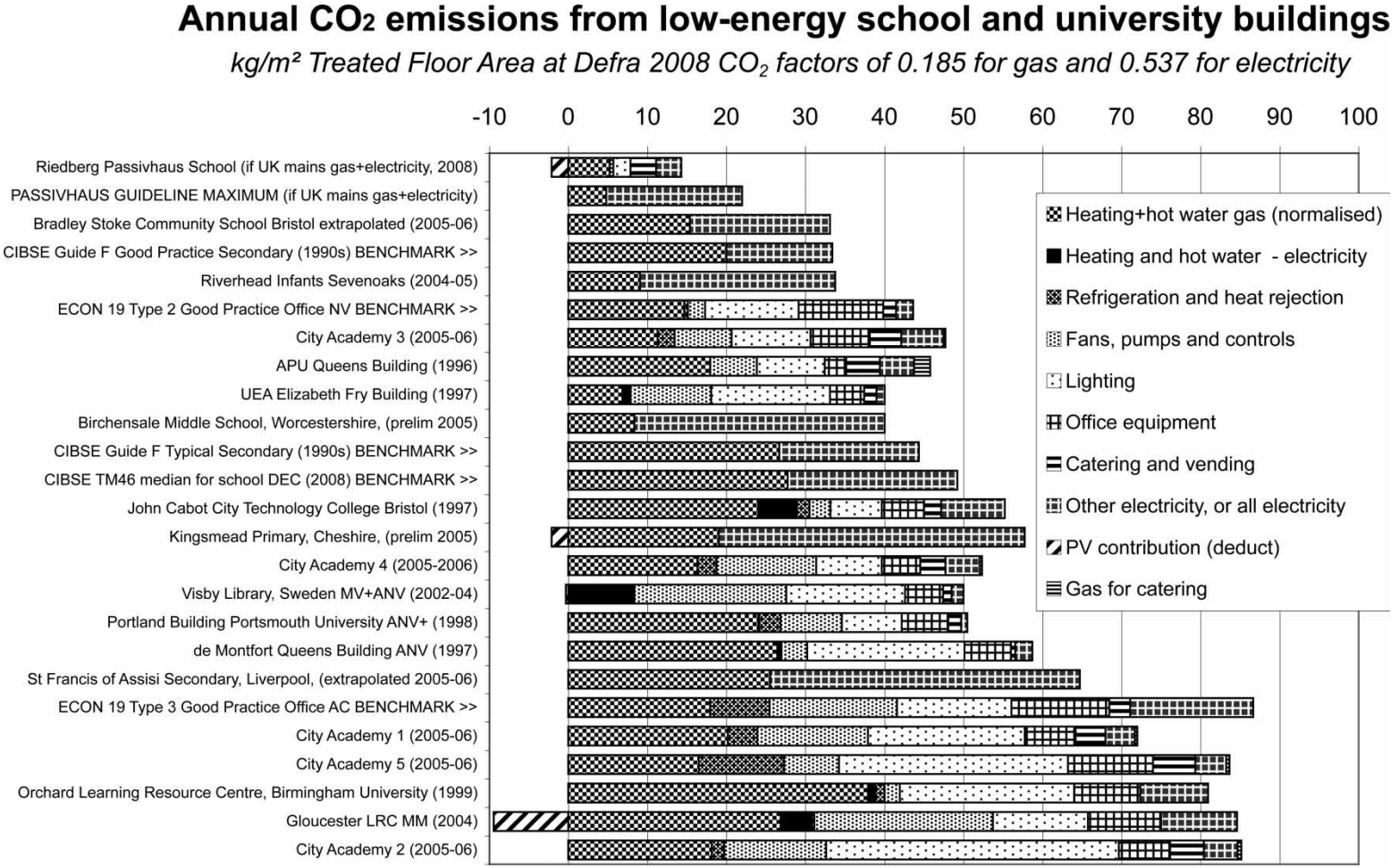

Figure 3 Format and style for communicating building energy performance. Source: authors

Effective feedback needs to provide objectivity, and lead to action and insight. It should:

- improve the performance of the studied building: this is nearly always possible, but needs motivation and commitment
- improve the services of those who provided it: this is always possible, but needs connection, motivation and knowledge management at the organizational level
- contribute to a wider knowledge base so that insights are disseminated and are more than anecdotal

Work by the UK government is progressing in this respect, with a feedback strategy including lessons learned that has helped to improve significantly the overall performance of its own buildings (Office of Government Commerce (OGC), 2009).

What methods should be used?

A large number of techniques have been used for building evaluation with hundreds reported in a review carried out in the 1990s by Baird *et al.* (1996). This is one of the problems. It is hard to know where to start and stop. In our experience, the following usually serve their purposes well:

- Expert walk-throughs, with informal discussions. These are quick and usually effective, but experts can be fooled. Inexperienced researchers may miss important things, and are also unable to provide immediate comments and feedback to their hosts, whose patience tends to run out more quickly. Learners therefore benefit by visiting sites with more experienced researchers, especially for the initial visit or where uncertainties arise.
- Measuring technical performance of building fabric, services and systems. Here it is important to concentrate on what really matters and not be side-tracked.
- Assessing environmental performance, usually energy but increasingly water and indoor air quality.
- Occupant survey questionnaires. People often miscast surveys of users as merely 'subjective', but, as Gary Raw said, 'People are the best measuring instruments. They are just harder to calibrate' (private conversation, 1995). The most important thing with occupant surveys, especially in smaller buildings, is to get a high response rate.
- Structured discussions interviews with participants. If needed at all, these are usually best when the results of occupant and other surveys are available and can form a basis for discussion and identify issues and pinch points. In our experience, focus groups that include a peer group of people can work well in non-domestic buildings. However, in housing, individual interviews are better, as focus groups can easily settle on certain gripes and be dominated by peer pressure.
- A visual record of matters related to the above five points. A photographic record of design features, including videos and thermographic images where appropriate can help highlight features and identify problems.

In choosing methods, these requirements stand out:

- Techniques should be relatively inexpensive and not too intrusive or time-consuming.
- Methods should cover basic user needs such as occupant comfort and control, use of space, storage, heating, lighting, cooling, noise, perceived health and productivity at work, image, location and safety. Methods should also allow people to expand on things that are not covered in detail in the surveys, such as usability or other aspects of design quality. This helps to keep the length of the questionnaire down and ensures that topics are not missed.
- Benchmarking against empirically derived (not theoretical or modelled) yardsticks is now a standard requirement for energy and occupants. However, benchmarks are hard to achieve because they need to draw on at least 30 studies in any geographical area to be really meaningful. Benchmarks do not yet exist for housing studies because so few of them have been performed consistently.
- Methods should incorporate a 'drill-down' approach so that more detail can be called upon or added, if needed. This is explored in Figure 4. A typical study might be at level 2, but progressing to levels 3 or 4. For example, if the occupant survey uncovers problems with thermal comfort, the diagnosis can be improved by reviewing the responses, say, floor by floor, and following up with discussions, technical checks, and possibly more detailed comfort surveys. Because of the large number of variables and their interrelations, it is usually fruitless to attempt to analyse all of them, but potential for secondary analysis should be available.
- 'Drill down' allows single topics or groups to be studied in more detail. A 'graduated response' allows the whole analysis to move to a deeper level. For example, the CIBSE TM22 Energy

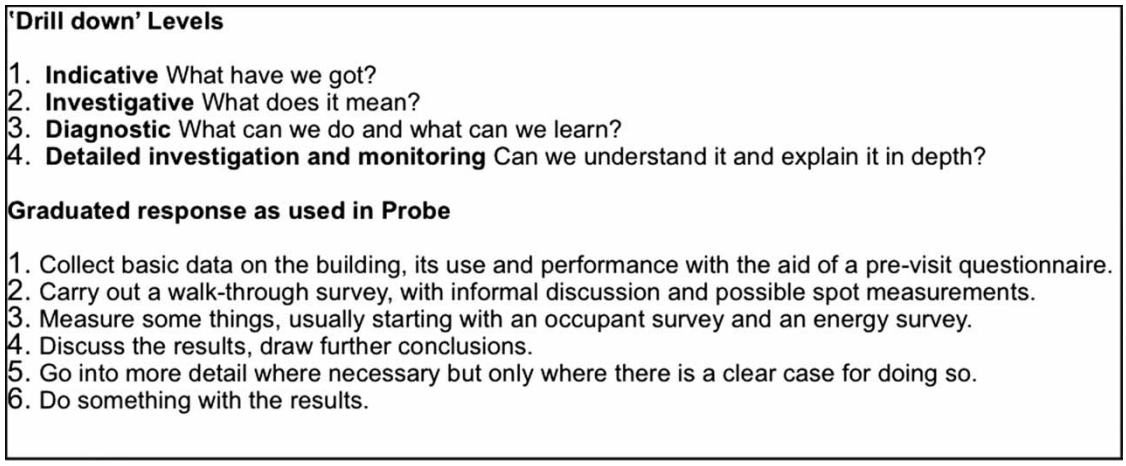

Figure 4 Graduated response. Source: Authors.

Assessment Reporting Method has three levels of increasing depth, rather like a computer game.

- Focus on the building in use, rather than on the design process. Obviously, lessons learned from the building in use should feed into the design process for future buildings. The design process, though, is not normally the object of a building performance evaluation study.

How to deal with context

Buildings are self-evidently settings or 'contexts' for human activity and behaviour. Activities can be studied at different levels of spatial resolution from individual behaviours at one end of the scale through to society at the other, each level nesting inside the previous one rather like Russian dolls. Each level forms the context for the next level in the system, so there is always some element of relativism present. Researchers will often want to allow for this relativism by 'controlling for context'. For example, the degree-day is a widely used statistical measure that helps to take account of changes in energy use for heating or cooling. Thus, comparisons can be made more readily, *e.g.* from month to month or year to year. Degree-day corrections are also used to compare the performance of buildings in different locations. However, buildings are profoundly adapted to context (Nicol and Humphreys, 2002), so, for example, buildings in colder countries tend to be better insulated. Adaptation also occurs even where buildings are nominally to the same regulatory standard. For example, if one is situated in a windy place, then there is a tendency to take more care about controlling draughts. Therefore, in a sense, controlling for context happens anyway.

Also, if contextual differences are evened out, this creates a risk of disguising the very things that need to studied in the first place! For example, if a researcher wishes to understand the effects of different lighting systems on, say, educational performance, then it is not sufficient just to set up a sample of buildings with different lighting systems. Difficult as this is, the researcher also needs to check whether the other features of the building (the thermal performance or noise, for instance) are also affecting occupants' perceptions and the ways they use and adapt to light. Viewed statistically as 'independent' and 'dependent' variables, buildings are 'cat's-cradles' of interdependence and complexity. Multivariable statistical techniques such as factor analysis, although superficially attractive for dealing with many such interrelated variables, are often problematic in interpretation because they tend to summarize groups of like variables rather than highlight individual differences or features.

For these, and other reasons too long-winded for here, building-performance studies should seek to expose and reveal contexts rather than controlling for them. It is thus important to describe fully the circumstances operating at the time of the study. For example, if staff have just been made redundant in the week before an occupant survey, then this should be made clear so that the reader, not just the researcher, can judge any likely effects on results.[13] Our advice is, 'normalize at the last possible moment'. In other words, use absolute (raw or untransformed) measures wherever possible, and if there is a need to control for context, then explain the process clearly, as with degree-day normalization.

Another perspective on context is shown in Figure 5. 'Context free' refers to principles, rules and processes that may be applied anywhere, irrespective of location. 'Context dependent' are factors locally determined. 'Physical' represents the features of a building's physical form; 'behavioural' is user activities within a

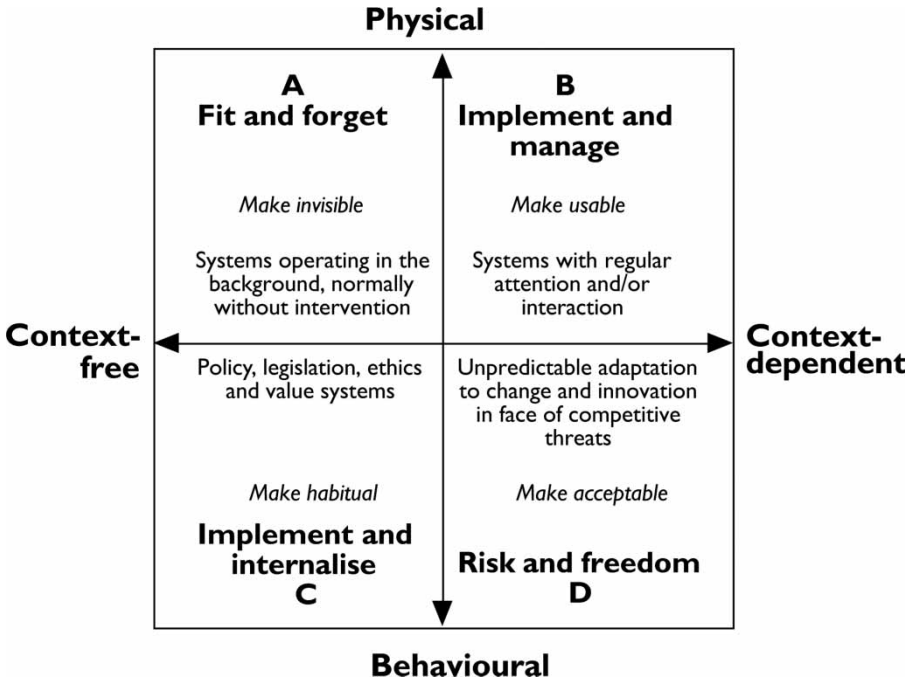

Figure 5 Perspectives on context

building. How these relate to each other is a way of characterizing:

- things that are supposed to work in the background with little or no human intervention, '*Fit and forget*', quadrant A
- things needing regular attention or intervention, '*Make usable*', quadrant B
- formal and informal rules that help with safe, comfortable and smooth running, '*Make habitual*', quadrant C
- things that emerge from existing use and situations as they develop, '*Risk and freedom*', quadrant D

The authors' experience is that the best buildings work well in all four quadrants. Buildings that can be said to be truly 'flexible' and 'adaptable' will have included consideration of each of the four strategies somewhere in the briefing, design and fit-out processes. Figure 5 splits into context free (quadrants A and C) and context dependent (quadrants B and D). There is an uneasy dynamic at work here. On the one hand, the imperative of the marketplace is to try and force everything into the context-free, 'fit-and-forget' and 'habitual' categories for which there are supposedly standard solutions. However, the real world resolutely brings everything back to context dependency, so that, for example, management and maintenance of installed systems are usually much more important than people are led to believe in marketing hype. This was certainly the case for a number of innovative housing schemes in Scotland where poor maintenance regimes and a lack of access to parts compromised heating systems (Stevenson and Williams, 2007). In the words of a building management system supplier, 'We sell dreams and install nightmares.' There is a lot of money to be made in creating dependencies from false promises of technologies that are then inappropriately applied.

A further take on context is from Stewart Brand in his seminal book *How Buildings Learn* (1994). Buildings evolve with contextual processes always present, always changing, and always subtly altering background conditions and constraints. Jeremy Till has taken this step further with observations on the contingency of building processes (Till, 2009). Without adequate inputs of management and maintenance resources, buildings may quickly assume vicious circles of deterioration and dysfunctionality. This process usually starts with poorly executed handover and commissioning, so that chronic performance inefficiencies are built in from first occupancy. Once present, these can become embedded, and then quickly create conditions for chronic failures like occupant discomfort and poor energy performance. This results in a vicious spiral of further performance deficiencies. At their heart are failures to design for manageability, with the vigilance demands of the promised features, and technical systems being out of proportion to the resources available for ongoing facilities management and maintenance. Buildings that work properly in their context create virtuous

circles of improvement, which are reinforcing. If these buildings are also relatively simple, like the Elizabeth Fry Building at the University of East Anglia, Norwich (Standeven *et al.*, 1998), they are more likely survive in the long-term.

What about design quality and cost?

Irrespective of design intent, buildings always act in four ways: they modify climate, behaviour, resources, and culture (Hillier and Leaman, 1974, also developed in Lawson, 2005). Design effort, though, has not been equally spread. Too much emphasis is often given to the image buildings present and not enough to behavioural and usability aspects, for example; or to indoor climate and environmental impact, though this is now changing.

From the perspective of evaluation studies, each of the four categories is just as important as the other. However, culture and resources are underrepresented, because they are much harder to pin down and so more difficult, and thus expensive, to study. For example, it is relatively straightforward to collect rounded data on energy and occupant performance, as most of these can be gathered and verified by the research team on site in a relatively short period of time, at least for non-domestic buildings (in housing, data are more individualized and studies may take longer). However, cost data are much harder to pin down: they tend not to be stored in standard ways (or even at the individual building level), money comes out of different pockets, and organizations are often reluctant to provide or release it. The most useful studies of behaviour often triangulate qualitative and quantitative findings with cultural and socio-technical aspects. Gram-Hanssen (2010) successfully uses this approach in five case studies that explore occupant habits and meanings in relation to housing energy data.

How best is the public interest served?

A persistent area of difficulty is whether or not results are published. The present authors between them have carried out over 700 building studies, but, despite our best efforts,[14] fewer than 10% of them have been published. The main reasons are:

- Reluctance on the part of building owners, occupiers, developers and designers to have 'their' buildings cast in a critical light. Our experience, though, is the opposite. In the Probe series of post-occupancy studies, where designers and researchers embraced the strategic and detailed findings, it helped to give their work impetus and credibility.
- Cost and time. Producing a journal article of professional quality is a skilful and time-consuming undertaking. Often, it is too demanding to add the 'front-end' article on top of the research reports.
- Partial studies. Buildings may have had occupant surveys or energy surveys carried out on them, but not both together. This incomplete picture means that a full journal article is not possible.
- Self-interest beats altruism. Although most agree about the public interest of having articles with 'lessons learned' (usually for what to avoid even more than what to do), commercial self-interest in suppressing problems usually wins. For example, design practices may fear sharing knowledge with other practices perceived to be competing with them.
- Lack of a properly funded national umbrella organization for building evaluation to provide a formal network and promote the dissemination of findings.
- Doubts about the status of building-evaluation studies. With the notable exception of this journal, academic journals still find evaluation studies hard to classify.
- Audiences. There are at least four separate audiences. Designers wanting to avoid past mistakes; educators passing the knowledge on to students; existing and prospective building owners, occupiers, developers and managers; and policy-makers looking for the best way forward.

Duty of care

With any interventionist study, there is a duty of care to building occupants. A study that asks people to rate, for example, their comfort and health carries with it the implied promise that something will be done about conditions if they are below reasonable expectations or unsafe. If people report to the research team features of the building that they think are unsafe (*e.g.* one building where occupants commented that the glass stair treads also used as landings caused them to misjudge distances and trip), then this information must be passed on to management.

There are wider matters, *e.g.* respondents using an occupant survey as a safe conduit to report bullying in the workplace. Obviously, this has to be dealt with discreetly and sensitively. Equally, poor housing-management practice has been highlighted in interviews about occupant comfort. Such outcomes stress the importance of systematically reading through all

feedback and comments. These insights may be missed with automated, internet-based surveys.

The needs of individual building users can never be ignored because it can never be predicted quite what they will say. If respondents are given plenty of opportunity to comment, then they will almost certainly reveal things that are important to them, even if there is not a direct question on that topic in the questionnaire. We also find that however good the base building, there will always be a cohort of people (usually above 15% of occupants) who will be dissatisfied with the conditions in one way or another. From a tactical management point of view, it is usually better to try and deal with the needs of this group and improve things for them. Baird (2010) has case-by-case details on positive and negative occupant responses.

We recommend that the names of respondents are included on survey forms so that any detailed duty of care matters can be followed up. However, it is also important to anonymise database files and reports so that names of individuals are never referred to and that identities are always protected. In the UK, Data Protection Act requirements also mean that the names should never be used for any purpose other than the direct objective of the study.

Principles

The guiding principles shown below are drawn from our personal experience. Others will be able to add to them from theirs. The list may assist those wishing to undertake building performance evaluations and obtain feedback, particularly in the emerging areas of housing retrofit and 'zero-energy' and 'zero-carbon' new build.

- A study cannot go ahead without access and cooperation from the occupiers. For non-domestic buildings, the authors use a pre-visit questionnaire that helps to establish a relationship with the main contact on-site, often the occupier's facility manager. Dwellings need a different approach: they are smaller and seemingly easier, but it is often harder to obtain access. As a result, response rates may be lower than hoped for. Also, the wider variety of lifestyles people follow in their own homes affects performance outcomes and makes it harder to record and analyse data. Differences between home-owners and renters also need accounting for. In rented housing, there may also be communication difficulties with tenants, and sometimes safety problems for the researchers.

- Apply the highest achievable standards of sampling, data collection, data analysis and reporting to the numerical parts of the work. The findings will have to appeal to management and design decision-makers at one end of the spectrum and to statisticians at the other, so graphs and plots, for example, must tell a clear story, but also be rich in necessary detail for possible scrutiny, with underlying assumptions made plain. Try to resist pressures to oversimplify statistical graphics, *e.g.* graphic designers often want to reduce clutter and convert graphics into their preferred house styles and colours. However, it is often vital for proper comprehension for the full detail of the original to be retained.

- Try to benchmark results against a bigger sample of buildings reporting both absolute scores and their respective relationships to each other, and to benchmarks. Do not normalize results unless absolutely necessary, and only then at the last possible stage in the analysis process. For housing it will be important to develop benchmarks for typologies (flat, semi-detached, detached, etc.) and occupancy (owner versus tenant) in order to understand the impact of rapidly emerging energy policy in this sector and to capture relevant technical and usage detail. It is not necessary to have a particularly large sample in one study in order to obtain meaningful data (Energy Saving Trust (EST), 2008).

- Write up the findings in a jargon-free and accessible way, remembering that the target audience is a wide one and that not everyone knows the acronyms used.

- Where databases hold the results, be scrupulous about applying the terms of the local data protection legislation, and capture the stories not just the data, so that the contexts can be understood and revisited.

- Give as much effort to the management and maintenance of the database 'metasystem' (*i.e.* the collection of different studies) as to the individual studies themselves. Researchers and funders almost always underestimate how much effort is required to maintain the metasystem. This is the reason why many database initiatives fail in the long term.

- When carrying out surveys, apply the codes of conduct of organizations such as the Market Research Society.[15] Make clear to the client at the outset that all results will remain confidential unless they otherwise agree. If a written report is prepared, then this should also remain confidential to the client. Any published report should be prepared with the express permission of the client, and where necessary the building owner. The published report should include 'lessons learned' and should not exclude important findings. Clients

will be often be nervous about uncontrolled use of findings, once they get into the hands of a press that is motivated to sensationalize results, especially the problems.

- Encourage the client to identify a route for the 'lessons learned' to be incorporated into their knowledge-management systems and not just sit on a shelf. For housing developers and managers, this will often mean making a suitable link between customer services and the maintenance team.

- Modularize reports and data appendices so that it is easier to separate out sections of the report that can be seen by others and those that must be kept private.

- If it is difficult to publish specific findings, then try to offer at least generic findings that disseminate the lessons leaned without revealing their source.

- Aim for clear 'strategic' summaries of findings trying not to include too much technical detail, whilst also drilling down to any particular details which may have had an unexpectedly large effect on the outcomes, for better or for worse. Do not overwhelm the reader with too much information. Be clear at what level conclusions and recommendations are required or even needed at all: in some instances the data may speak for itself.

A way forward

Accumulated knowledge has been gathered about building evaluation over 25 years, mostly from the non-domestic sector. Despite some notable differences, many of the principles and techniques used and lessons learnt are relevant to the domestic sector, and sometimes directly transferable. This belief is based on initial housing evaluation studies recently carried out, and referenced above, which have benefited from the principles outlined here.

Building evaluation stands at a crossroads. Government policy-makers globally are realizing the importance of a stronger evidence base for building performance. This will allow the building industry to improve the performance and resilience of new, and particularly existing, buildings as rapidly as possible. Recently, for example, the UK government has deemed that housing should deliver a particularly large saving in carbon emissions as part of its low-carbon transition plan (Department of Energy and Climate Change (DECC), 2009). Clearly, this deliverable must be measurable, and yet to date housing evaluation has been patchy at best with only a nascent methodology.

There is a danger that government will fall into the trap that, in the hope of delivering rapid results on a large scale, they will promote initiatives that work on paper but not in practice. They may favour things that make profitable businesses but do not serve the public interest, with too narrow a focus. Already, it is apparent that while low-carbon is the desired outcome, a focus on 'zero carbon' is creating unnecessarily expensive, complicated and technically risky approaches which building evaluation work predicts are unlikely to work well or prove good value in practice. In UK housing, there has also been too much focus on heating and renewable energy, while burgeoning electricity use, improved building procurement methods, and opportunities for understanding and influencing behaviour have been largely overlooked.

Much of this has been heard before. It is now 40 years since the landmark building performance work at Strathclyde University, Glasgow (Markus *et al.* 1972); over 20 years since the first books on methods (Zeisel, 1984; Preiser *et al.*, 1988), and nearly 10 years since the special issue of *Building Research & Information* dedicated to the Probe studies (*Building Research & Information*, 2001). Most of this work has been about non-domestic buildings: but there is little housing building evaluation. Unfortunately, however, lessons are still not learned, in spite of the crying need to close the feedback loop and get our buildings performing radically better. People tend to say 'we know all that' and then blunder on to repeat the same mistakes. There is something systemically wrong.

The divisions of responsibility make it difficult to close the feedback loop from building performance in use to briefing, design and construction. In particular:

- In the UK especially, central and local government have been outsourcing their technical skills (Bordass, 2003). For example, the UK no longer has a Property Services Agency; in spite of a massive school building programme, there is no longer a Design Research Unit in the Department of Education; and the technical departments in local authorities are shadows of their former selves.

- Central government has outsourced its research too. In the UK, the Building Research Establishment (BRE) was privatized in 1997 shortly after its 75th anniversary. It is now a consultancy. No longer does government have a single authoritative source of disinterested information to which to turn.

- Government perpetuates a category error in seeing 'the construction industry' as the experts on building performance. In fact, the industry designs and alters buildings, but does not know much about how they perform in use. For the most part, it hands over the keys and has no continuing involvement or interest.

- In spite of the talk about whole-life costing (*e.g.* OGC, 2007), splits between capital and operating expenditure are rigid. It proves difficult in practice to set aside capital budgets to include aftercare, tune-up and feedback after building work is over. It is proving equally difficult to fund these activities from operating expenditure brought forward. For example, in the UK we have Private Finance Initiative (PFI) finance, design, build, and operate packages that might be expected to link things up. However, inside the package, responsibilities can be even more tightly divided up than ever, *e.g.* with the project being sold on after it has been built; and if feedback is obtained, it tends not to be shared.

How can the deadlock be broken? In the 1980s, the authors thought this could be achieved by publishing research results, but designers and managers were not part of this culture. In the 1990s, we used published post-occupancy evaluations as a window onto building performance more generally. This worked well, until government funding stopped, mainly because case studies were perceived as 'more of the same' rather than a fundamental contribution to ongoing feedback. We then worked with clients, but found that large repeat clients who could afford to fund the work mostly had a 'supply-side' mentality: their priority was to procure the building on time and to budget, not to get it to work as well as possible in use, and to learn from the experience.

Now we are working again with the industry, trying to foster a new professionalism that engages routinely with outcomes and consequences, and places more emphasis on integration, communication, and applied knowledge. Renewed impetus has been gained from 'green' building imperatives. However, even where teams are engaged and enthusiastic, this is still proving difficult, as the contractual and financial structures within which people have to work presume a particular way of doing things which often as not makes routine feedback difficult or impossible. If buildings are to become radically more sustainable, then a break with these restrictive practices is needed, for example, by getting clients, government, and the industry committed to Soft Landings-like processes that make follow-through, feedback, and building evaluation the norm.

References

Baird, G. (2010) *Sustainable Buildings in Practice: What the Users Think*, Routledge, Abingdon.

Baird, G., Gray, J., Isaacs, N., Kernohan, D. and McIndoe, G. (eds) (1996) *Building Evaluation Techniques*, McGraw Hill, New York, NY.

Bordass, W. (2003) Learning more from our buildings, or just forgetting less? *Building Research & Information*, **31**(5), 406–411.

Brand, S. (1994) *How Buildings Learn*, Penguin, London, (see also: http://video.google.com/videoplay?docid=8639555925486210852#).

Building Research & Information (2001) 'Post-occupancy Evaluation' (Special Issue), **29**(2).

Department of Energy and Climate Change (DECC) (2009) *The UK Low Carbon Transition Plan: National Strategy for Climate and Energy*, The Stationery Office (TSO), London.

Energy Saving Trust (EST) (2008) *Monitoring Energy and Carbon Performance in New Homes*, CE298, EST, London.

Gram-Hanssen, K. (2010) Residential heat comfort practices: understanding users. *Building Research & Information*, **38**(2), 175–186.

Hillier, W. and Leaman, A. (1974) How is design possible? *Journal of Architectural and Planning Research*, **3**(1), 4–11.

Judt, A. (2009) What is living and what is dead in social democracy? *New York Review of Books*, **56** (20; 17 December).

Lawson, B. (2005) *How Designers Think*, 4th Edn, Architectural Press, London.

Markus, T., Whyman, P., Morgan, J., Whitton, D. and Maver, T. (1972) *Building Performance*, Wiley, New York, NY.

Nevill, G. (2007) So how are you doing? *Building Services – CIBSE Journal*, **November,** 32–37.

Nicol, J. and Humphreys, M. (2002) Adaptive thermal comfort and sustainable thermal standards for buildings. *Energy and Building*, **32**(6), 563–572.

Office of Government Commerce (OGC) (2007) *Whole Life Costing and Cost Management*, OGC, London.

Office of Government Commerce (OGC) (2009) *The State of the Estate in 2009: A Report on the Efficient and Sustainability of the Government Estate*, OGC, London.

Pearson, A. (2007) Laban Dance Studio: a disappointing performance. *Building Services Journal*, **June,** 32–38.

Preiser, W., Rabinowitz, H. and White, E. (1988) *Post-Occupancy Evaluation*, Van Nostrand Reinhold, New York, NY.

Robson, C. (2002) *Real World Research*, Blackwell, London.

Shipworth, M., Firth, S.K., Gentry, M.I., Wright, A.J., Shipworth, D.T. and Lomas, K.J. (2010) Central heating thermostat settings and timing: building demographics. *Building Research & Information*, **38**(1), 50–69.

Short, C.A., Cook, M. and Lomas, K.J. (2009) Delivery and performance of a low-energy ventilation and cooling strategy. *Building Research & Information*, **37**(1), 1–30.

Standeven, M., Cohen, R., Bordass, W. and Leaman, A. (1998) Probe 14: Elizabeth Fry Building. *Building Services Journal*, **February,** 37–41.

Stevenson, F. and Rijal, H. (2008) The Sigma home: towards an authentic evaluation of a prototype building, in Proceedings of the 25th International Conference on Passive and Low Energy Architecture, University College, Dublin, Ireland.

Stevenson, F. and Williams, N. (2007) Section Nine: Case study lessons, in F. Stevenson and N. Williams (eds): *The Sustainable Housing Design Guide for Scotland*, Communities Scotland, Edinburgh, (available at: http://www.scotland.gov.uk/Topics/Built-Environment/Housing/investment/shdg/).

Stewart Milne Group (2009) *Sigma Home Brochure*, Stewart Milne Group, Aberdeen (available at: http://www.stewartmilne.com/media/Image/brochure/sigmaIIbuildsystembrochure.pdf).

Till, J. (2009) *Architecture Depends*, MIT Press, Cambridge, MA.

Wingfield, J., Bell, M., Miles-Shenton, D., South, T. and Lowe, R.J. (2009) *Evaluating the Impact of an Enhanced Energy Performance Standard on Load-Bearing Masonry Construction – Final Report: Lessons from Stamford Brook – Understanding the Gap between Designed and Real Performance*. PII Project No. CI39/3/663, Leeds Metropolitan University, Leeds.

Zeisel, J. (1984) *Inquiry by Design: Tools for Environment-Behaviour Research*, Cambridge University Press, Cambridge.

Endnotes

[1]For example, in the Probe series of post-occupancy studies, only one out of 20 buildings looked at – the Elizabeth Fry Building, University of East Anglia, Norwich, UK – could reasonably be argued to meet all three criteria.

[2]For example, the location of measuring instruments may be compromised by users' use of the space.

[3]For the Probe archive, see http://www.usablebuildings.co.uk/.

[4]Probe used the CIBSE TM22 energy assessment reporting method (see http://www.cibse.org/index.cfm?go=publications.view&item=43).

[5]Probe used the BUS Occupant Survey method now also incorporated into the Arup Appraise methodology and available either via Arup or via Building Use Studies.

[6]See http://www.bsria.co.uk/services/airtightness/.

[7]A version was developed by William Bordass Associates.

[8]For the Feedback Techniques portfolio, see http://www.usablebuildings.co.uk/.

[9]For a summary of feedback and strategy references published in *Building Research & Information*, see http://www.taylorfrancis.com/journals/access/rbri-Feedback-strategy.pdf/.

[10]The Probe studies released a published overview report, but the sub-reports on which it was based remained private.

[11]For further details on the Soft Landings approach, see http://www.usablebuildings.co.uk or http://www.bsria.co.uk/.

[12]See http://www.aimC4/. This was the aim of Vital Signs Project (1992–1998) (see http://arch.ced.berkeley.edu/vitalsigns/).

[13]Our experience is that the results of occupant surveys are not affected by events such as staff redundancies. However, others may think differently, so such information should be provided so that readers can judge for themselves.

[14]For example, setting up a charity, the Usable Buildings Trust, to disseminate results to a wider audience.

[15]The Market Research Society Code of Standards is a useful guide (see http://www.mrs.org.uk/code.htm).