Building Performance: Research Methods

Adrian Leaman

... supporting www.usablebuildings.co.uk

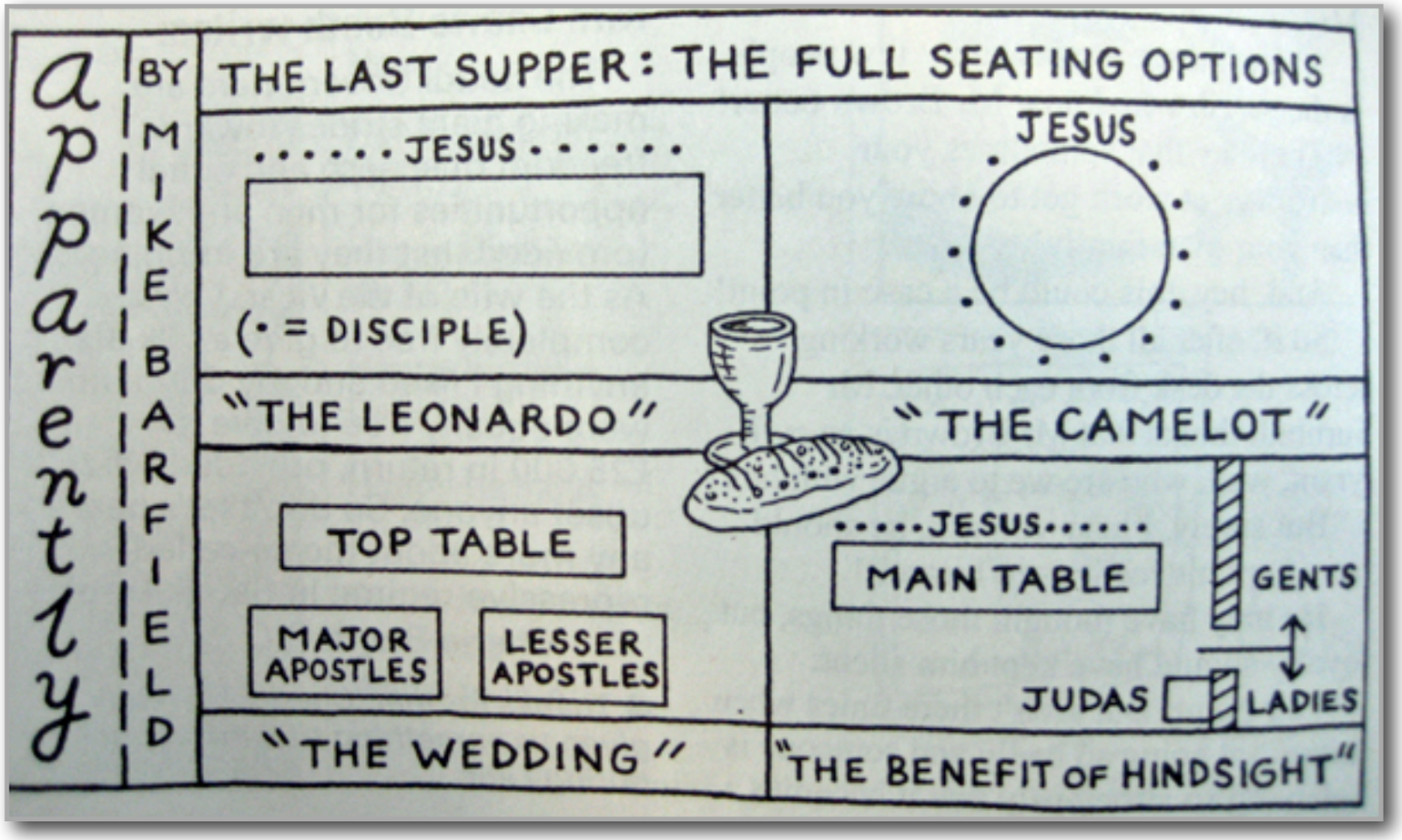
This is about ...

- **A brief overview of our approach to research underpinning building performance evaluation**

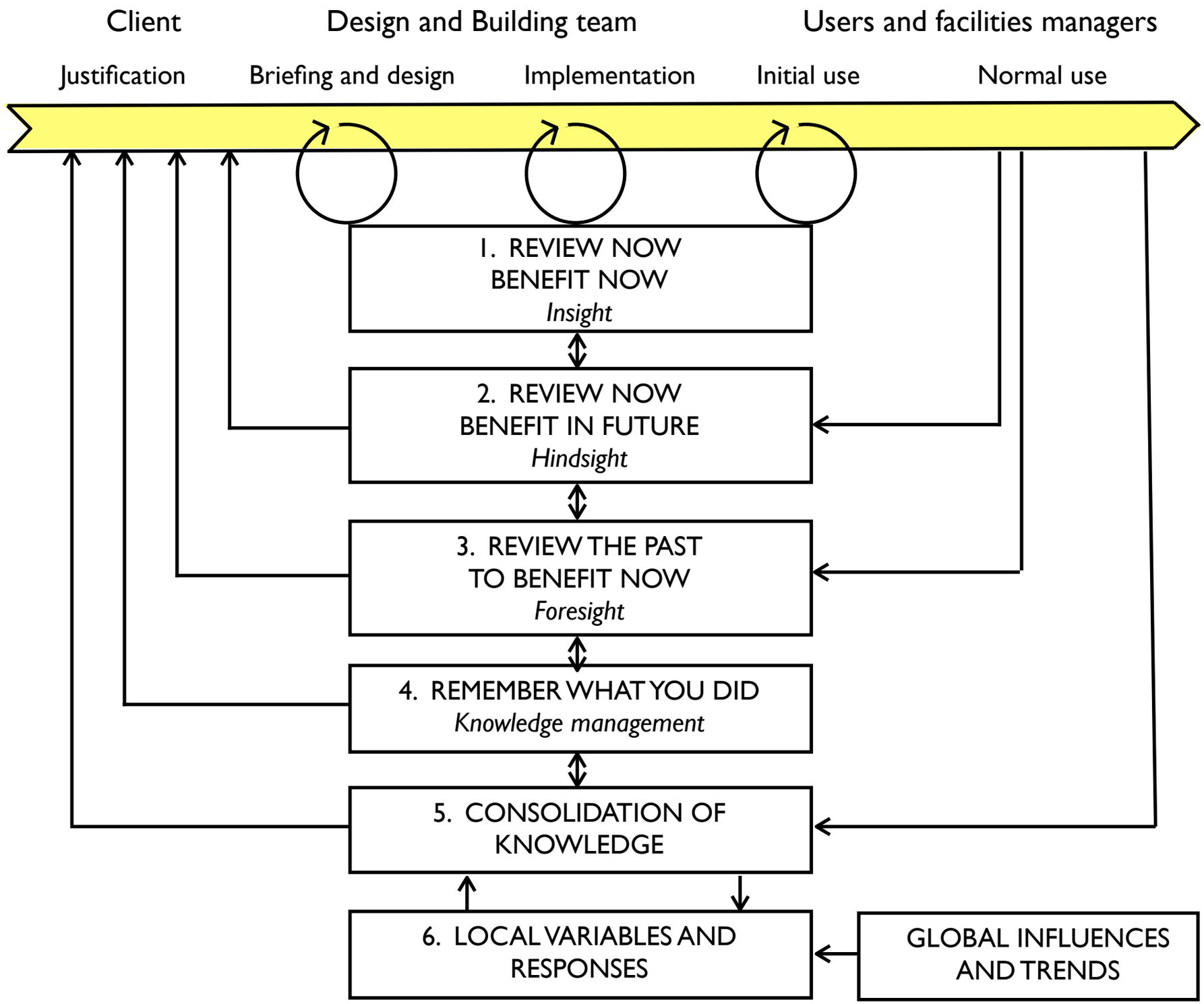
- Findings are supported by **survey information,** including the **Probe** series, in Britain and worldwide.

- ☑ More detailed information is available from the Probe, Books, Events and Publications pages of www.usablebuildings.co.uk.

In a nutshell ...



Types of feedback



Source: Bill Bordass 2004

Our main aim

Make survey methods capable of **widespread take-up** by design practices, building occupiers and researchers so that performance monitoring becomes both a **routine** part of the design process and a **mainstream** part of the educational experience of designers.

"Building research should never be more than one step away from a design or management decision."

after Bill Allen

Real-world research

Solving problems	…	Just gaining knowledge
Predicting effects	…	Finding causes
Looking for robust results and concern for actionable factors	…	Statistical relationships between variables
Developing and testing services	…	Developing and testing theories
Field	…	Laboratory
Outside organisation (e.g. business)	… rather than …	Research institution
Strict time and cost constraints	…	R&D environment
Researchers with wide-ranging skills	…	Highly specific skills
Multiple method	…	Single method
Oriented to client	…	Oriented to academic peers
Viewed as dubious by some academics	…	High academic prestige

Source: Adapted from Box 1.2 of Robson R., Real-world Research London, Blackwell, 1993.

What's the problem?

Despite nearly forty years of development, **building performance assessment is still not routinely applied** to mainstream building design practice.

Methods used for assessment can be **unwieldy, costly to implement and hard to manage** across more than one building study.

Requirements vary widely from formal hypothesis-testing to simple rules of thumb.

- ☑ **First in UK:** The first serious studies of building performance assessment in the UK were by Tom Markus and colleagues in the early 1970s at Strathclyde University.

- ☑ **Latest: Preiser W and Vischer J. (eds), Assessing Building Performance** Practical advice on assessing and monitoring building performance, Butterworth-Heinemann, 2004.

Techniques …

… are often created for **single-building studies**, or single-project, multi-building studies, with little thought for further application later, especially with respect to benchmarking, and use across multi-building types;

have to consider a **large number of variables**, with the significance of any of them often not fully clear at the outset, so it is hard to know what to leave out;

do not take enough account of **similar techniques already created** for similar study purposes, with a tendency to disregard serviceable methods which are 'not invented here';

and ...

underestimate how much of those resources have to be used for **back-office tasks** like scaling, categorisation and data typing, statistical analysis, software management, and quality control of data inputs and outputs, including range and error-checking.

over-optimism about the analytical and presentational capabilities of off-the-shelf software;

disagreements over **where best to use physical measurements** and where to measure.

and ...

some tend to look for **theories and causes** while

clients and building designers often need more **rough-and-ready guidance** (e.g. What should I do in this situation? What should I pay attention to? What should I ask? What should I avoid? How well did that idea work? Can I improve it next time?);

So ...

Studies can **run excessively over time** and budget.

Tendency for **every study to be treated as a prototype**.

Emphasis on individual building studies rather than on benchmarking and strategic conclusions across a larger sample of buildings.

Difficulties resolving results produced by different survey teams and protocols.

Not well attached to the procedures, timescales and budgets of clients, designers, builders and managers.

and …

'Scale' problems in the use of techniques when moving from single-building studies to studies with larger numbers of buildings of different sizes and functional types.

Over-reliance on standard software and graphical analysis products, with low standards of data presentation.

Tendency to over-simplify, thereby losing needed detail, or the reverse, a tendency to …

create data mountains, with excessive amounts of detailed information which is hard or wasteful to analyse.

Feedback Portfolio

www.usablebuildings.co.uk/fp/index.html

Our approach I

Studies which:

are **relatively less expensive to commission**;

provide unbiased results quickly;

are **useful across a wider range** of interest groups, (i.e. studies which are not necessarily restricted to single-issue topics or overly narrow professional audiences)

cope with the exigencies of the **real-world**, making allowances for the everyday requirements of building managers and occupants when the studies are carried out;

Our approach 2

Studies which are:

statistically **robust**, providing believable results that meet stringent statistical and quality control criteria, but are not bogged down in jargon and unnecessary detail;

useful in 'mix-and-match' clusters, so that the techniques can be put together in different combinations or separately as stand-alone studies;

compatible with other similar studies;

less off-putting to carry out for those who may wish to embrace them but may be frightened off by technicalities.

Less expensive I

'Need to know' not 'nice to have'. The standard BUS occupant survey now has 20 background questions and 45 building performance questions.

Modularised questionnaires so that sections can be added and subtracted easily.

A single database for the analysis of all questionnaire variants for every survey undertaken. The current 2007 version of the BUS individual building database has 4,500 database fields. Although the database is **relatively large and carries a lot of redundancy** because not all of the analytical power is needed for each study, **it is much easier to manage** across all surveys.

Less expensive 2

In-house software for all mainstream data analysis and presentation routines.

More **emphasis** on the production of reliable, attractive and easy-to-understand results and less to report writing.

A licensing system, so that users may apply the techniques for themselves (under licence) and then have them analysed.

Relatively **more effort paid to the management of the benchmarking database**. The current 2007 BUS buildings (benchmarking) database has 301 records (buildings, of which 50 are used for the latest set of benchmarks).

Less expensive 3

More emphasis on **routine metadata management tasks** such as consistent variable naming between surveys, data typing and scale definitions.

The BUS system includes **a variable and scale name database** which has 240 records (variable names). This services the 10 survey variants to ensure that all variable names, nomenclature, scale 'directions' and scale descriptions are absolutely consistent across all surveys.

Lack of clarity with definitions across individual building surveys, and confusion between individual building variables and benchmark variables can be a **major source of inefficiency**. This factor alone accounts for many surveys never proceeding beyond prototypes.

Probe

www.usablebuildings.co.uk

follow Probe menu item.

Rapid turnaround I

As soon as a survey has been carried out, clients want to have the results, so **the sooner the better**.

The energy and technical module of Probe, including a detailed technical report, can be **concluded in a elapsed time of no more than ten days** after the occupier has completed their pre-visit questionnaire [8], including a detailed technical report. With annual energy use data already available, the spreadsheet-based results can produce a first cut in less than a day and detailed results in typically three days. **The major drawback with collecting energy data, however, is getting the energy consumption information out of the occupier's organisation and its energy suppliers in the first place.** In some extreme cases, this can take 2-3 months.

Rapid turnaround 2

The occupant module takes a day or less to set up, a day to carry out and less than a day to produce substantial analysis results. A report would add two more days, so **the total turnaround time is five days or less**.

The air tightness module, used in the latter stages of Probe, takes about 3 hours for blocking off holes and ducts, 1-2 hours for fixing the monitoring equipment to the building, and about 2 hours for the tests themselves. If any further investigation is required after the first measurement (for example, to investigate the reasons why the building did not pass the test) then this might take a further 3 hours. **The test usually takes less than a day all told.**

We found from Probe that ...

Secondary information generated by the two main survey modules (energy/technical and occupants) supplemented by specialist interviews where necessary, **gave us most of the information we needed** on these topics; and on other topics too.

the classified list of **written comments in the occupant survey identifies issues not specifically covered by the questions** themselves

an energy survey one also reviews the **quality** of construction, installation, commissioning, record information, operation, management and maintenance.

in an energy survey there is a chance to **talk to occupants**, facilities and maintenance staff whilst walking round, seeking information, checking plant and taking spot readings of light, temperature, air movement and so on.

Sampling occupants I

Robustness: techniques must still work reasonably well even if some of the assumptions on which they are based are mildly violated

Real world: It may not be possible for cost or access reasons to achieve the levels of sampling that might be desirable.

We usually try for the **largest sample that is reasonably available** (especially in buildings with less than 100 permanent occupants). In larger buildings, we try for a minimum of 125 responses.

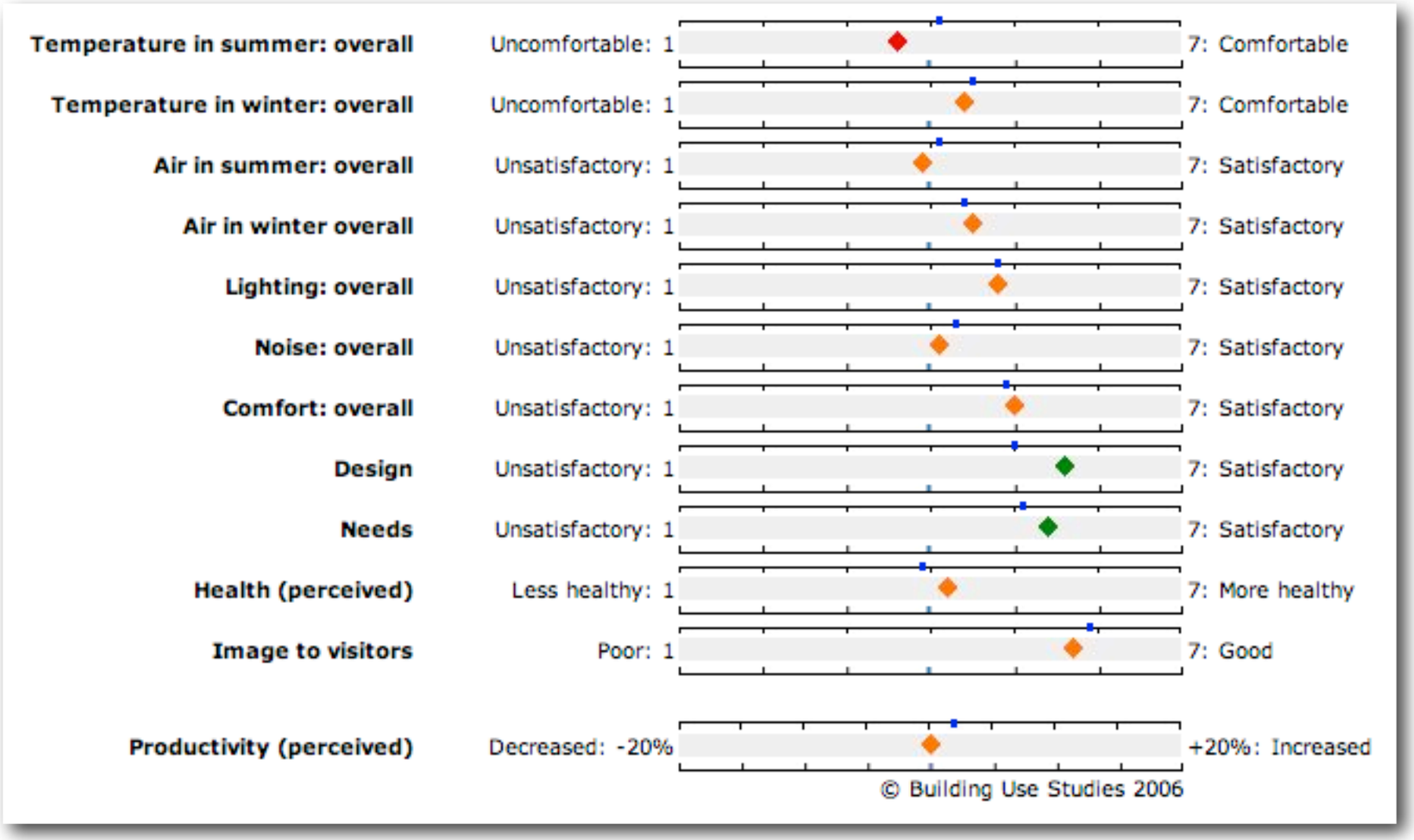
Sampling occupants 2

The mean sample size across all the buildings we have studied in the BUS occupant survey is 122.

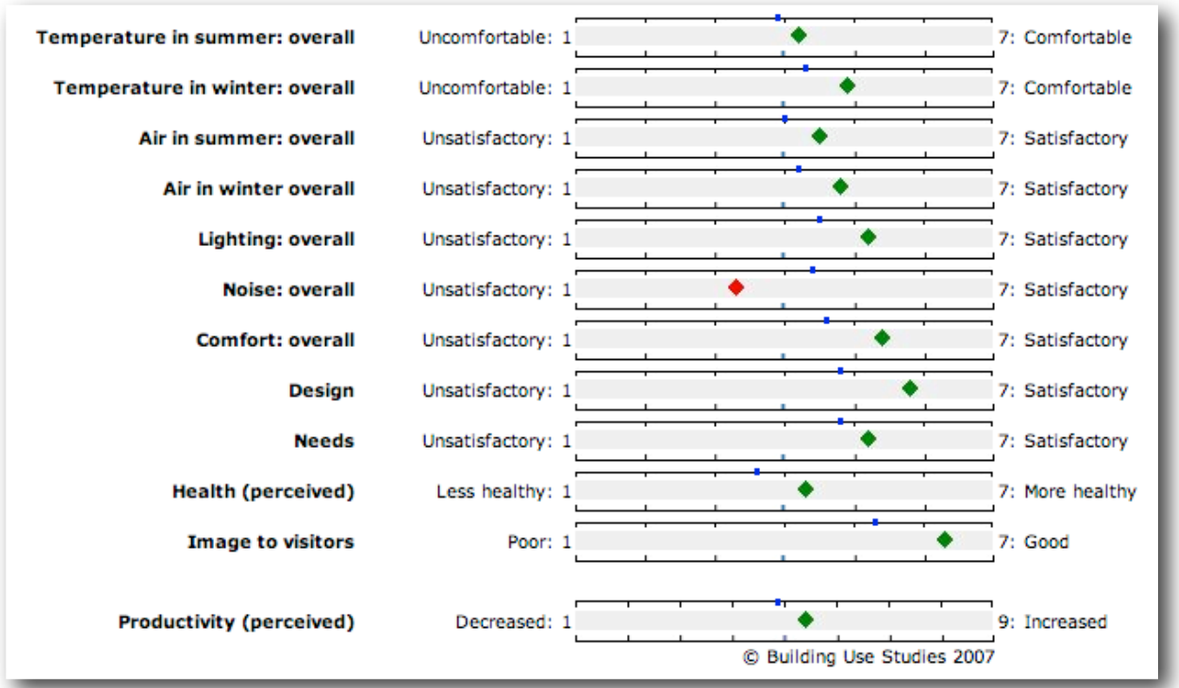
Response rates are almost always higher than 80 per cent of those sampled (where BUS administers the survey on site we normally achieve 90 per cent plus).

We normally stick to **paper questionnaires**, personally handed out to and collected, in the building because these reach everyone in the building (including staff who do not have computers) and produce much higher response rates than internet-based questionnaires.

Higher response rates are more important than large samples.



Results not public domain



Energy surveys

We advocate:

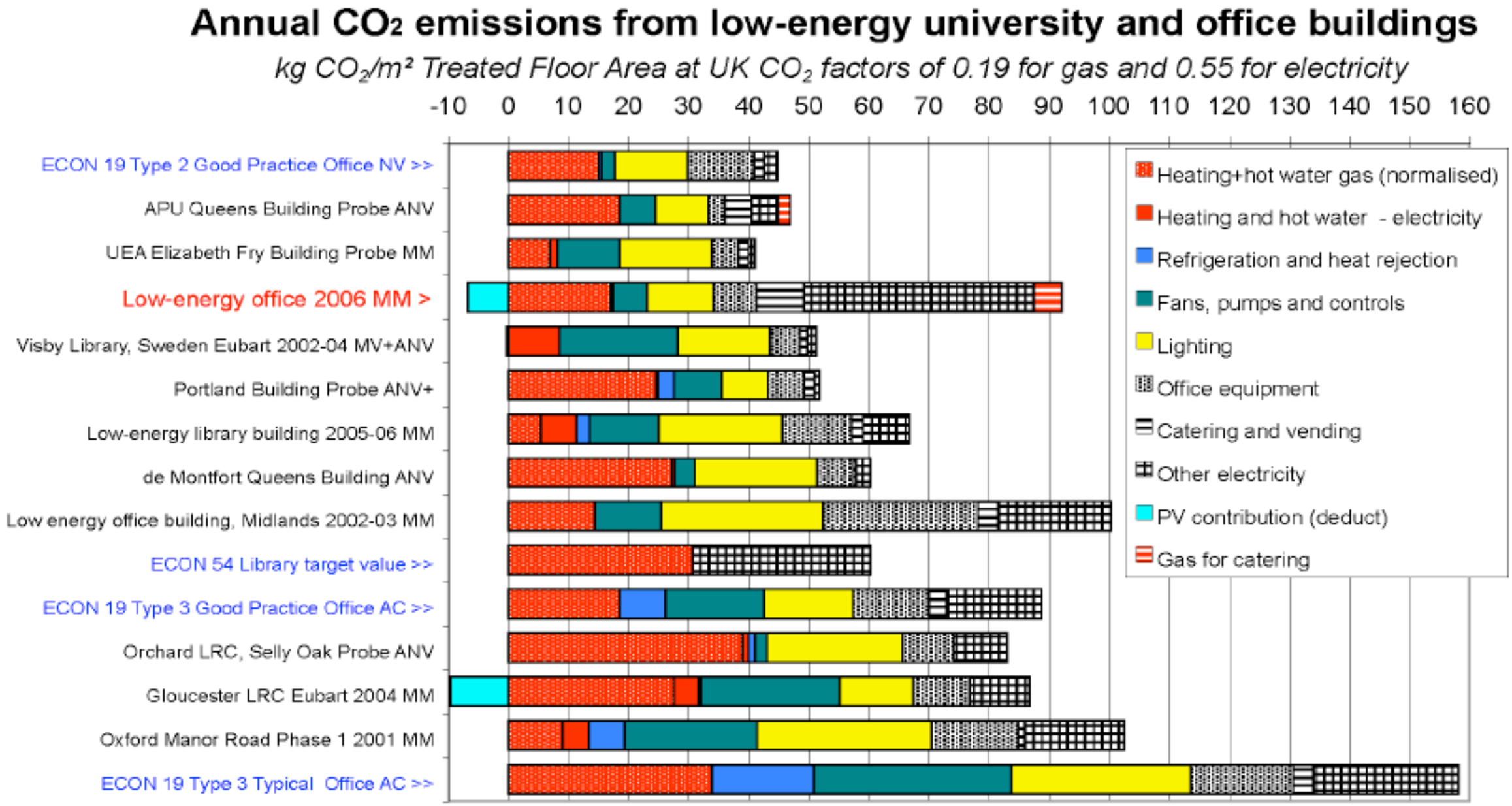
graduated response, that is, they can cope with different levels and qualities of input data. For example, energy use and equipment may only be partially available, or there may be uncertainty about the veracity of the data; and there is no point collecting details which you do not really need (we do it on a “need to have” basis instead);

in-built error-finding capability, where they can detect discrepancies and/or mistakes in the input data and make them clear in the outputs.

Why we need to permit corrections

What are the relevant emissions of the office in red?

92 kgCO_2/m², or 34 (85 or 27 deducting the PV); or … ?



Things that can cause problems

or have been used to massage the figures

POOR INPUT DATA

- Incorrect building classification — *e.g to get a generous benchmark.*
- Poor energy use data — *seek annual Energy Statements from suppliers.*
- Not all energy sources included — *try to improve meter registration etc.*
- Poor floor area data for denominator — *use registered data if possible.*
- Floor area artificially inflated — *make checks or use registered data.*

OVER-EGGED CORRECTIONS

- Special areas and end-uses — *require strong burden of proof, usually sub-metering*
  Not sufficient just to exclude the item, to change the statutory grading, also report its energy efficiency and opportunities for improvement. This paves the way for specialist advice, e.g. ICT, to look into the equipment, not just the building services.
- Exaggerated hours of use — *a common problem, needs stringent checks.*
- Exaggerated occupancy densities, in and out of hours — *stringent checks, as above.*
- Extra allowances for having things like air-conditioning — *benchmark allowances under the statutory system should be for what a building DOES, NOT what it HAS.*

LOW CARBON SUPPLIES USED TO MAKE INEFFICIENT BUILDINGS LOOK GOOD

- Important to motivate building + management efficiency before low-carbon supplies: *otherwise these precious supplies will be squandered.*
- Avoid *"get out of jail free"* with purchasing measures like green electricity. *though these might be reported alongside on a voluntary basis.*

Error finding

The TM22 energy reporting method has a **'double-entry' system to help resolve discrepancies**. Metered supply data is calculated separately from the demand data and the balance is reviewed throughout the calculation process.

The BUS occupant survey method has separate software for **range-checking and variable name checking** the spreadsheet data input files. Log files of anomalies systematically list any suspicious input data. This approach is more systematic than using menu-driven spreadsheets, which can be more error prone.

Simple mistakes lead to gross errors

Compatibility

Ideally, techniques should be compatible across studies so that comparisons may be more easily made.

It can be helpful for **a set of core questions to be identical**.

> This applies especially when two complementary databases need to be brought together. It is best if there is **at least one data field in common**.

However, **tiny changes in formats destroy compatibility**, even between two similar studies in the same organisation!

In theory, **open source** arrangements based on licensing and the emergence of **XML** may help to overcome this.

Improving take-up I

Make it more attractive for people to use evaluation techniques routinely.

Many are put off by the research culture. Designers are often uncomfortable with statistics, data analysis and survey methods. So it is important to demystify techniques as much as possible, explain how they work clearly, especially the assumptions and terminology, and present the results attractively. Equally, the results should not be trivialised or over-simplified.

Performance assessment studies are often carried in dissertations and technical reports, and seldom reach mainstream audiences. When they are written up for academic journals, perfectly good studies tend to be rejected by referees because they often do not fit their stricter academic criteria.

Improving take-up 2

Studies have to meet a wider audience, from the corporate manager who only really wants to know how a building affects the perceived bottom line, through designers who will be more interested in detailed diagnostics of building features, to researchers who may be more theoretical in outlook.

Results need to be trustworthy and reliable, so people do not think that studies present only the favourable part of the whole story, leaving out the bad news (in which the greatest part of the scope for rapid improvement actually lies).

Improving take-up 3

Ideally, methods should also be structured by the main targets and headings set out in design briefs, so that it is easier to check whether targets have been met. **Building evaluation then becomes a routine part of quality control and continuous learning**, rooted in the real world.

Above all, **methods must be seen to be value for money**. In the past, post-occupancy work has been seriously under-valued by the construction industry and design professions, to such an extent that few have ever wanted to pay for it. Clients, too, do not see why they should have to pay yet more fees for what they regard as the construction industry's problem.

In a nutshell

Make feedback routine by making methods understandable, manageable, timely and relevant.

Results should be **never more than one step away from a design or management decision**.

Data should be **need to know, not nice to have**.

Understandable by all, **so no begged questions**.

Results must appeal to advanced practice so **highest common factor not lowest common denominator.**

Things that can cause problems

or have been used to massage the figures

POOR INPUT DATA

- Incorrect building classification — *e.g to get a generous benchmark.*
- Poor energy use data — *seek annual Energy Statements from suppliers.*
- Not all energy sources included — *try to improve meter registration etc.*
- Poor floor area data for denominator — *use registered data if possible.*
- Floor area artificially inflated — *make checks or use registered data.*

OVER-EGGED CORRECTIONS

- Special areas and end-uses — *require strong burden of proof, usually sub-metering*
  Not sufficient just to exclude the item, to change the statutory grading, also report its energy efficiency and opportunities for improvement. This paves the way for specialist advice, e.g. ICT, to look into the equipment, not just the building services.
- Exaggerated hours of use — *a common problem, needs stringent checks.*
- Exaggerated occupancy densities, in and out of hours — *stringent checks, as above.*
- Extra allowances for having things like air-conditioning — *benchmark allowances under the statutory system should be for what a building DOES, NOT what it HAS.*

LOW CARBON SUPPLIES USED TO MAKE INEFFICIENT BUILDINGS LOOK GOOD

- Important to motivate building + management efficiency before low-carbon supplies: *otherwise these precious supplies will be squandered.*
- Avoid *"get out of jail free"* with purchasing measures like green electricity. *though these might be reported alongside on a voluntary basis.*

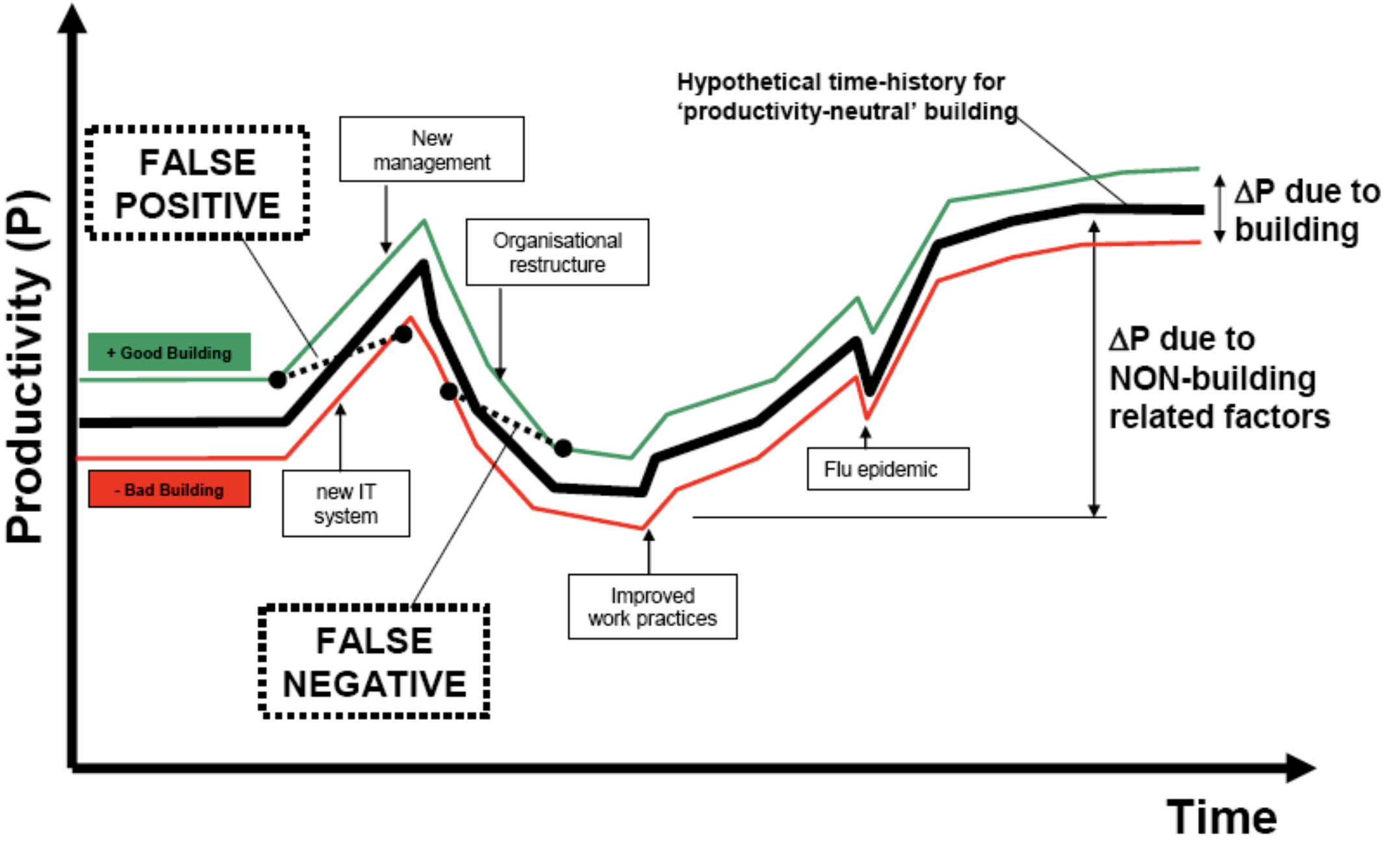


Figure 18: Conceptual diagram showing possible misleading effect of contextual factors on before-and-after productivity assessments