What do we *really* know about wellbeing?

Property owners want proof that they are delivering workplaces conducive to occupant wellbeing. Small wonder that consultants are attempting to deliver that proof through various forms of certification. But can wellbeing assessments truly improve operational outcomes, or is the scientific evidence inconsistent and contrary? **Dr Roderic Bunn** explains why we need to exercise considerable caution

Property owners, architects, and engineers are rightly obsessed with creating healthier workplaces. One can't fault that. While the ambition is utterly worthy and well-motivated, we need to recognise that proof of the links between comfort, health and productivity have eluded researchers for decades. Commonsense states there must be relationships but empirical certainty is hard to come by.

Decades of academic research into relationships between physical environments and occupant health, comfort, and productivity have found many associations between occupant satisfaction and sources of discomfort. These include poor indoor air quality, low air change rates, thermal discomfort, excessive noise, and so on. But in that time no reliable metrics have ever been developed, let alone ones that can quantify the consequential improvements in occupant health and productivity. The research evidence is partial at best and often contrary.

Claims of causal associations, sought so desperately by construction professionals and their clients, suffer from a chronic lack of repeatability. Those who claim reliable replication tend to hide their calculations behind a veil of commercial secrecy. Some academics, consultants, and property owners act as if the scientific evidence is compelling, if not conclusive. Institutions set up to promote wellbeing (often funded by the industry they support) often amplify claims without reality-checking them. Small wonder that the act of disputing such claims, publicly, draws condemnation and ostracism. It seems that one challenges the wellbeing religion at one's peril. If you want to know why, just follow the money...

To understand how we got to this point it's worth reflecting on the research trajectory. Over the last 30 years the most significant academic work in the field of occupant comfort and satisfaction has stemmed from large cohort studies[1]. Survey data gathered from, say, 100 buildings have been data-mined, re-defined, extracted, or concatenated into tailored databases by subsequent generations of comfort researchers. Statistical sausage machines churn out calculations purporting to demonstrate correlations between environmental conditions and comfort levels, either self-reported or measured in some way. In some cases correlations are strong, such as absenteeism or sickness rates against some measure of indoor air quality, while others are marginal. Almost all statistical analysis is compromised, to a lesser or greater degree, by confounding variables or limitations in the research processes. Databases may comprise samples so small they cannot be considered representative of the population from which they were derived. Sometimes the claimed correlations seem somewhat random, or – worse – suspiciously contrived.

Some researchers muddy the waters by disaggregating survey data to suit a particular research interest. For example, in the last 15 years it's been fashionable to segregate wellbeing and comfort data into 'green' and 'non-green' building categories, despite evidence that suggests this categorisation is questionable at best, if not downright daft. This is because the categorisation on what is, or is not, 'green' is typically based on environmental certification rather than measurements of how the buildings actually perform. When the surveyed respondents in 'non-green' buildings are found to deliver wellbeing scores that are as good as – or even better than – those occupying the 'green' buildings, the researchers struggle to interpret their results.

The root of the problem is a category error of course, but inherent researcher bias lies behind it all. As the Swedish statistician Hans Rosling memorably said about health data, miss-categorisation doesn't necessarily stem from academic ignorance, but from pre-conceived ideas.

Even when wellbeing researchers report eye-catching and ostensibly compelling findings – often promoted and amplified by industry bodies that have a stake in a positive research outcome – those findings usually can't be repeated. Correlations

reported between measures of productivity and researcher-defined measures of wellbeing (say, acoustics or biophilia) often aren't supported by subsequent research. Conversely, some factors that are found to be loosely or poorly correlated in one study may be found in a subsequent study to possess a degree of statistical strength. It's pot-luck in many ways: you choose your academic journal and you select the paper you like.

The reasons for these differences in wellbeing research are many and varied: a variety of different research methods and tools, differently-constructed population samples, or a choice of modelling tool whose internal assumptions and boundary conditions are highly developer or vendor-specific. Esoteric statistical tests are also becoming more prevalent in wellbeing studies, the outputs of which befuddle industry practitioners who might rightly wonder whether the researchers are using smoke and mirrors in order to prove their hypotheses. Such fears may well-grounded, particularly where wellbeing research has been commercially-sponsored. Research datasets that underpin wellbeing analyses are not usually published, so you can't see how they were constructed let alone judge whether they have been manipulated.

Having said all this, there are many researchers working in the field of occupant satisfaction and wellbeing who are conducting research honestly, thoroughly and objectively. The difficulty is working out who they are. Be in no doubt, statistical deception is rife in the social, physical and medical sciences. This is largely driven by an incestuous relationship between academic research bodies, the funders of research (who overtly favour the potential for commercial exploitation), and academic journals biased towards persuasive research outcomes. The tacit pressure to prove an hypothesis is often stronger than the value of reporting a null hypothesis. The demand for research citations can also pressure researchers to exaggerate, distort or even falsify their research findings. Accusations, no matter how well-evidenced, tend to be met with fierce indignation by those trying to build or protect their academic reputation[2].

This bias seems particularly strong in wellbeing research. Perhaps it's due to the power of organisations that stand to benefit commercially from statistical associations between, for example, high ventilation rates and improved productivity. It doesn't take too much imagination to work out who they might be.

Questionable wellbeing research often stems from laboratory studies, particularly those erroneously defined as 'office environments'. Recent claims of statistical associations between comfort and productivity have derived from cognitive tests carried out on volunteers placed under a range of environmental conditions in the hermetic world of a climate chamber. Quite apart from the fact that a laboratory has limited relevance to the real world, statistically-significant degradation in human cognitive abilities only tends to be found between extremes of whatever environmental variable is being measured, such as thermal conditions, levels of carbon dioxide, or concentrations of certain volatile organic compounds (VOCs). There is usually no discernable difference in human performance within the range of acceptable conditions that might reasonably be expected to prevail in a typical workplace.

Researchers who draw straight lines between two extreme conditions rarely, if ever, have the data to justify a claimed percentage difference in human cognitive ability (such as a measure of productivity) at any point along those lines. Readers of such research may rightly smell a rat when the researchers report statistical associations in one type of cognitive test (e.g. arithmetic) but not in another (e.g. reasoning skills). Subsequent wellbeing research may even report the opposite, for a wide variety of methodological or sampling reasons. Again, repeatability is the problem here.

In laboratory wellbeing studies there can be inherent problems with sample sizes, i.e. the number of people being tested. Climate chambers usually can't hold more than a dozen people. This, statistically, is a very tiny sample. A classic work-around is to concatenate human data from repeated tests to make the sample population look much larger than it really is, thereby implying that the whole sample is statistically valid when its disaggregated samples might be in single figures.

Irrespective of problems with sample sizes, laboratories contain way fewer confounding factors than are operating in real buildings. By working in a climate chamber the researchers are able to tactically strip out confounding variables they don't want. This gives them an unsullied and direct relationship between the physical characteristics they do want and the wellbeing factors they're studying. The problem then is that the results can't export back to the complexity and messiness of the real world. It doesn't stop them trying, however. The various wellbeing institutions don't seem awfully bothered by this, and broadcast the results willy-nilly.

On top of all that we also need to recognise a fundamental and rather inconvenient truth: that the innate human ability to adapt, cope, and tolerate a wide range of comfort conditions will always

confound attempts to show that one condition is statistically better than another. This is usually in terms of some kind of measured output – typically some arcane measure of wellbeing and/or analogue for actual productivity.

All these problems – and there are many, many more in a similar vein – don't seem to be holding back wellbeing research. There seems to be a kind of desperate desire to believe in a categorical relationship between wellbeing and environmental conditions in buildings, even though one cannot be reliably or compellingly delivered. Property owners with deep pockets and both eyes on maximising rental yield seem willing to pay for the answer – any answer – as long as it supports expectations. Construction professionals seem only too happy to say they can deliver one.

As a result there's no shortage of wellbeing assessment tools developed purely to meet the demand. Everyone wants a slice of the action, perhaps before the opportunity disappears; rather like, one suspects, the 1980s anxiety over sick building syndrome.

All of which leads to a key question: can a wellbeing assessment support an uplift percentage for improved productivity, or, indeed, any other outcome metric one might like to invent? Can the complex cocktail of variables that underpin occupant comfort be poured into a statistical centrifuge and separated out to the extent that one can say that doing this, this, and that, will equate to a 15 per cent improvement in wellbeing; and, by extension, support a claim than an office building is 18.5 per cent more productive than the one next door? If it can, so the thinking goes, the benefits are considerable: maybe lower absenteeism and sickness rates, reduced staff turnover, and less haemorrhaging of expensively-procured staff skills and expertise. And, of course, a better bottom line.

With all this in mind, it's worth considering what those confounding variables are, and what role they may play. Let's begin with the ways in which researchers obtain the perceptions of building occupants.

Most research in wellbeing tends to rely on data generated through self-assessed occupant surveys. There are some thoroughly-validated and proven ones available – for a price, of course. However, the UK construction industry's habit of sparing no expense to do things on the cheap (plus an industry-wide infection of the 'non-invented here' syndrome), means that good occupant surveys are vastly outnumbered by a great many rapidly-written and largely untested ones, originating both in academia and commerce. It's said that any fool can write a wellbeing survey. Many are indeed doing so, with partial experience and even less interest in testing their survey for reliability and validity before they launch it on an unsuspecting market.

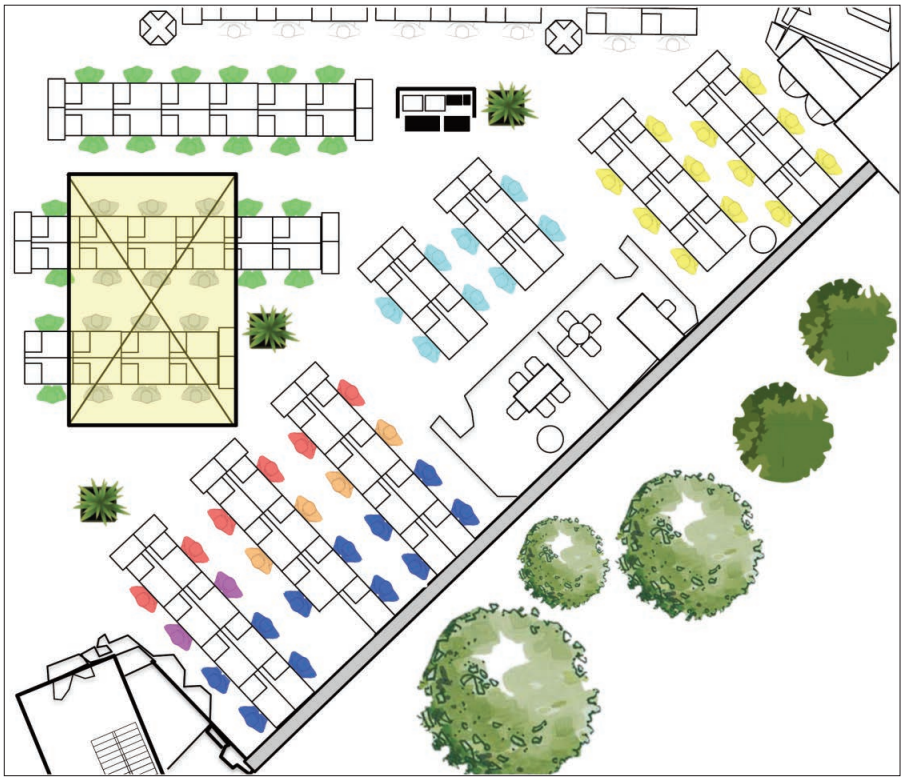
What all surveys tend to have in common is a means by which building occupants can score their comfort (*aka* wellbeing) perceptions. Commonly this is via web or paper-based surveys, and/or forms of structured interview. While the ways of doing this would take too long to explain here, more detailed advice is available[3]. Done well, such surveys can be highly informative. Done badly – particularly with poorly-worded questions and/or small population samples – such surveys will generate highly misleading statistics on wellbeing. But, how would you know?

Questions on technical subjects that are posed to non-technical people have to be ones that respondents can reasonably be expected to answer. For example, most people can differentiate between hot and cold conditions or whether they have enough daylight, but abstract concepts like 'morale', 'healthiness' or 'wellbeing' can easily confuse survey respondents and lead to wide variation in responses. Furthermore, while people may be able assess conditions quite easily as being 'too much' or 'too little' of something tangible (e.g. light or noise), asking people how they *feel* about those factors, in other words asking them to place a personal value on them, is not only more subjective but highly dependent on whether they've ever thought about those things, let alone be able to express an opinion that's reliable and useful.

For example, this writer conducted a straw poll of office staff to define – without them conferring – what 'wellbeing' meant to them. Figure 1 shows a word cloud of the responses. Some redundant verbs and conjunctive terms were removed to highlight the defining characteristics of the wellbeing perceptions. Of course, word clouds aren't scientific by any stretch of the imagination, but they are good enough to illustrate the problems with defining wellbeing.

Figure 1: Definitions of wellbeing obtained from a people in an open-plan office.

Figure 2: Seating positions in a typical open-plan perimeter work area.



It's immediately evident that the factor 'health' was mentioned most, but it's also interesting that, even with this small sample, people mentioned a wide variety of factors. Many people related wellbeing to light, air, space and noise, but functional factors including maintenance, space, and infrastructure also got a mention.

The example clearly demonstrates a lack of consensus among respondents on what constitutes wellbeing. It seems that each respondent had a different idea of wellbeing swirling around in their head when they answered the question. This makes total sense: human experience is forged by many factors: gender, age, where they are in the building, whether they are trapped at their desk all day or free to move, whether they have access to natural light and so on. They may have the ability to control environmental conditions or they may be reliant on centralised control. They might be sitting in a corridor zone being assailed by others walking up and down talking on mobile phones. Furthermore – and more critically – they may not share the definition of 'wellbeing' with that held by the researcher who created the questions. And if they didn't, how would the surveyor know that?

Assuming that wellbeing can be defined in such a way that a researcher can be certain that respondents to a survey understand it the same way, there are still major problems with treating those responses as spatially homogeneous. Consider the corner of an office layout as illustrated in Figure 2. This is based on the layout of real (mixed-mode ventilated) office, but is typical of many.

The occupants have been characterised by their location to openable windows (the dark blue seats), those who are part-time (pink seats), those assigned hot desks (yellow seats), and those in deep-plan office space with their backs to nominated circulation routes (green seats). A few colleagues who enjoy daylight via a rooflight are shown under a highlighted yellow box. People represented by bright blue seats sit adjacent to a management office and a meeting room. These people may have little visual or acoustic privacy and no access to windows or environmental controls. They could be regularly disrupted by colleagues using the printer hub, and by people wandering (noisily) to and from the meeting rooms.

Let's consider ventilation. Satisfaction with natural ventilation tends to depend on the ability to manage the trade-offs between ventilation and draughts, room temperature, radiant temperature, natural light and glare, and possibly views out to the external landscape. Arguably such trade-offs can be managed best by those seated in the dark blue seats and closest to the control devices (windows and blinds). Furthermore, a democratic consensus over the position of perimeter devices might be reached easily between two to four people – what one could call the control workgroup. However, control will get progressively more difficult if the group increases in size to include the orange seats. It may break down entirely if desk positions grow to include the red seats. Those people may perceive conditions very differently to those in the dark blue seats, but they would be too far away from the perimeter to exercise control. Four seats

perpendicular to a window (eight desks in all) may be the limit at which democratic consensus over window and blind operation can be realistically achieved. In some cases the number may be lower than that.

Rows of desks longer than four deep will probably render consensus impossible in all but the most cohesive of workgroups. People on the outer rim of such groups may suffer in relative discomfort. They may even become relatively resentful of the building, which also might breed intolerance of noise or other discomfort factors that staff nearer the window are able to tolerate simply because they have the means to adjust conditions.

Dissatisfaction with thermal conditions and indoor air quality may also depend more on exposure and dose rather than the exceedence of a level defined, say, by institutional guidance. For example, people in a building for short periods (e.g. part-time staff) may be able to tolerate moderately adverse conditions, as may hot-deskers if they can choose their desk or, perhaps, rotate position on a daily basis to share the benefits (and drawbacks) of a window location. Staff with permanent desks, however, may become relatively intolerant, and may even be driven to exhibit what are known as 'escape behaviours' when conditions get too onerous. This may lead to behaviour such as working in an atrium, or hijacking an empty meeting room - an act that could conceivably be mis-interpreted as 'agile working'.

So what about those people sitting under the skylight? They may welcome the daylight, but they may also suffer cold downdraughts. If so, then their expressed wellbeing may be a function of offsetting the drawback of downdraughts by the possibly greater benefits of the extra daylight, and, possibly (if one believes some wellbeing research), emotionally-uplifting views of the sky. The opposite might apply, if, for example, the gender balance is in favour of females. It is known that females consistently report feeling colder than males by a whole integer on an uncomfortable /comfortable' 7-point survey scale used in many comfort questionnaires[4] (Figure 3).

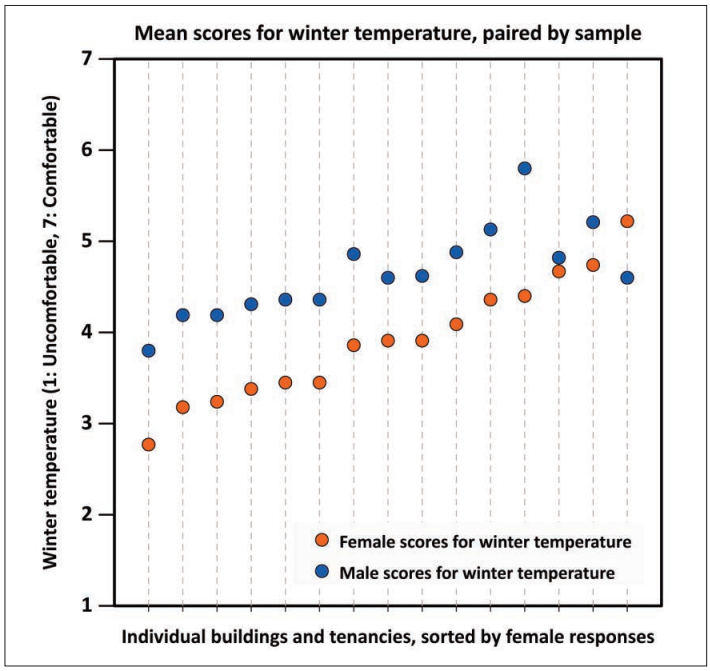


Figure 3: Females consistently reported being colder than their male office counterparts in 15 buildings and tenancies studied in a longitudinal comfort study[4]. When females are thermally neutral in winter, males are slightly too warm. The pattern in scoring is also found in scores for summer temperature to a slightly lesser degree.

Noise disturbance is often a major factor in perceptions of satisfaction, particularly in open-plan offices. People by an openable window will be closest to external noise but will also have the ability to manage its volume. Those people can trade-off noise against ventilation. Birdsong, for example, may be welcomed while road traffic will not. Again, those furthest away from the windows will have less say over that control decision. Furthermore, those seated in window seats will be interrupted less by cross-talk and general movement in circulation zones, relative to those in centrally-located desks. On top of that, noise conditions usually vary during the working day. In open-plan, we all know hubbub can be useful in cancelling out irritating sources of noise and in preserving acoustic privacy for private conversations. Both may become a problem at low occupant numbers, usually at either end of the working day and working week.

Theoretically, those seated in green seats seem to have the worst of all worlds. It may be safe to assume so. Or, it may not: their jobs may allow them to be more mobile than other workers. They may be nearer the printers or the beverage points – both of which have advantages and drawbacks. They may have plants that provide some visual delight, even if they do little for indoor air quality unless the internal light levels are above, say, 4000 lux[5] (which, of course, would introduce much bigger problems).

The example doesn't cover the effect of functionality variables, such as the availability of personal storage, or whether meeting rooms are readily available or always booked up, forcing irritated staff to improvise. There can be issues of occupant density and utilisation – how often all seats are occupied. At normal times occupant density may hover around one person/8 m^2, but at peak times social density could conceivably rise to one person/6 m^2, creating a cocktail of disturbance factors which might – might – start to hamper productivity and prompt aforementioned escape behaviours, possibly among a large number of people. As with working in an atrium, it's possible that escape behaviours could be miss-categorised

as agile working, and thus regarded as a healthy and productive response, when, actually, it's an undesirable reaction to discomfort. Categorisation will largely depend on an employer's beliefs and the pervading organisational work culture. In any case, a study of agile working that does not take density into account would generate somewhat questionable statistics.

This example is serves to demonstrate the sheer complexity of occupied spaces in real buildings – even in one little corner of an office – and the difficulties inherent in making one-size-fits-all wellbeing assessments. Moreover,it illustrates the stupidity of attaching importance to laboratory tests that measure the effect of only one or two comfort variables. It's simply not possible to weight the factors in a given space accurately. Nor is it practical to come to a reliable assessment of wellbeing by simple addition, subtraction or summation of variables.

So, what do we know about wellbeing in buildings? Not very much, it appears. We may be able to devise theories and make commonsense assumptions, derived from studies of human health or behavioural psychology, but proving associations – statistical or otherwise – between measures of wellbeing and outturn human performance continues to elude us. The only thing that can be said with any conviction is that occupant wellbeing – however it is defined – is highly variable and ultimately wholly context-dependent. Furthermore, if one adds the potential effects of age and gender on wellbeing perceptions (particularly relevant to thermal criteria) the situation gets even more complicated and murky.

So where does all this uncertainty and complication leave the construction industry and its clients, both of whom seek iron-clad remedies for improving wellbeing and productivity?

In their influential paper *Productivity: The Killer Variables Twenty Years On*[6]), Leaman and Bordass warn that the cats' cradle of causality and association [of comfort factors] will differ from one building to the next, making it dangerous to be over assertive about causation without careful appreciation of the local context. Nonetheless, they believe there are some fundamental truths. In general terms Leaman and Bordass believe that buildings – especially offices – tend to work best for occupants when the following are satisfied:

- An environment that is comfortable for the most of the time, with plenty of opportunities for changing things should conditions deteriorate
- Usable systems with clear communication of design intent
- Buildings that can respond rapidly to perceived needs, supported by a responsive management
- Shallow-plan forms,particularly with systems that are neither technically complex nor management-intensive
- An element of natural ventilation, such as mixed-mode schemes that are well-integrated with any mechanical systems
- Good spatial capacity for the occupying organisation, with appropriate zoning for control of heating, cooling, lighting, ventilation, and noise (the latter also to meet needs of acoustic privacy)
- A focus on low energy consumption. This tends to be a consequence of better briefing, procurement, management and monitoring, all of which helps to create a workplace that will foster better human performance.

Will concentrating on these factors help improve occupants' actual experiences and perhaps some measure of productivity? It seems reasonable and much evidence suggests so. Furthermore, for many contexts that are under a lot of pressure – particularly densely occupied spaces – the

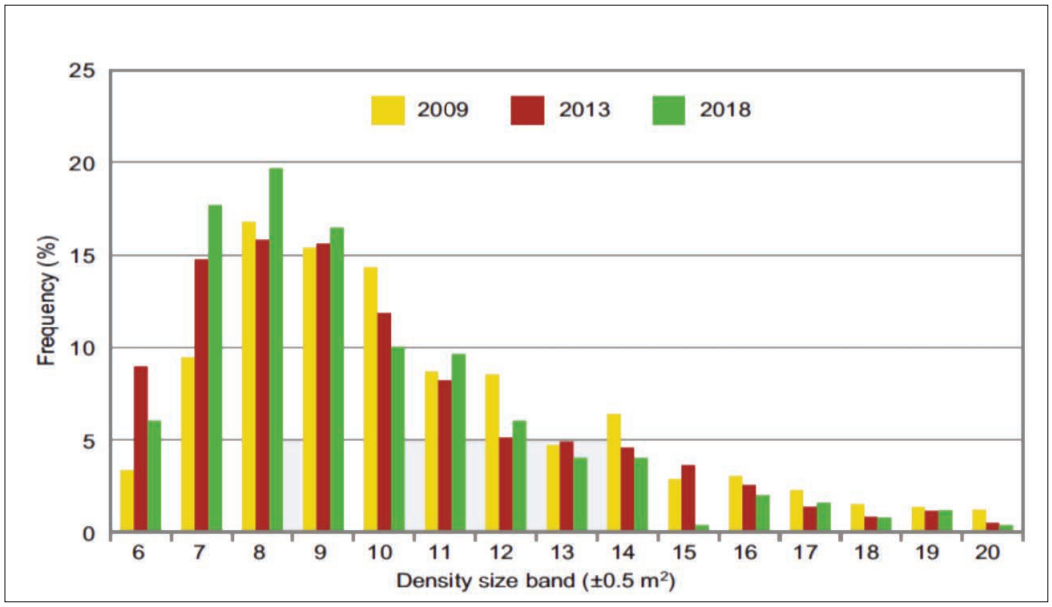


Figure 4: The increased densification of office space since 2009. Source: *Office Occupancy: Density and Utilisation.*(British Council for Offices, 2018).

elimination of discomfort might be the best that can be achieved.

The recent upward trend in occupant density reported by the British Council for Offices suggest that outturn performance might be more dependent on the numbers of people in a building, particularly given that high social densities will affect a great many variables – environmental and functional (Figure 4). Whatever is thought to constitute measureable improvements in wellbeing, may, in fact, be no more than a simple consequence of agile working policies that have reduced social density to more acceptable levels.

What this means is that assessments of occupant satisfaction (and whatever constitutes wellbeing) can only be done by being sensitive to the individual building contexts. The current fashion of trying to optimise building performance through the application of generic, rule-based, wellbeing-based metrics seems rather more risky, given the lack of empirical evidence and use of questionable measurement practices. Unfortunately, the flywheel of commercial wellbeing assessments is in full spin, and it will probably take years to slow down enough for people to consider more proven and practical alternatives. Those alternatives will likely only come from more robust and empirical case studies of real buildings, preferably free of the confirmation bias that tends to come from research motivated by commercial self-interest. Context-sensitive case-study research, unfortunately, is not practiced much in academia, which tends to prefer big data and statistical acrobatics. So a change is needed there, too.

I once asked a well-known social scientist what tools he thought most important when researching occupant satisfaction in buildings. Instrumentation? Physical measurements? Occupant surveys? "A functioning set of eyeballs," he said. He had a point. Buildings aren't nearly as mysterious as some people like to make out. Most problems with occupant satisfaction are readily evident, if you bother to look properly. The big question is whether what you've found actually matters. And that, as is often the case, is a matter of professional judgement. So-called 'wellbeing' assessments may not, in practice, possess a great deal of practical value. If they provide reassurance, like some kind of homeopathic remedy, then maybe that's justification enough. Just don't claim statistical proof of the benefits of wellbeing assessments in terms of improvements to health and productivity where such proof doesn't exist, and probably never will.

Dr Roderic Bunn is a building performance analyst and Soft Landings consultant. This article is an updated and extended version of one written by Roderic Bunn in November 2016 and available at www.bsria.com /uk/news/article/what-do-we-know-about-wellbeing/

Notes and further reading

[1]Large cohort occupant health and satisfaction research projects:

BASE (100 buildings, 1994 – 1998)
NIOSH (160 buildings, 1999)
EC-AUDIT (56 buildings, 1996)
SLOSH (4913 persons, 2003 – 2005)
HOPE (164 buildings, 2005)

Many research papers in the wellbeing field cite these projects or data-mine them. Note that the samples per building are not always as impressive or as statistically representative as the overall population databases may imply.

[2]Weblinks for further reading and deceptions in social, physical and medical research:
www.theguardian.com/global/2019/jun/30/house-plants-bloom-economy-wellbeing
www.theguardian.com/education/2020/feb/01/david-latchman-geneticist-should-resign-over-his-team-science-fraud
www.youtube.com/watch?v=Jt7gEAoUl8s (a talk by UCL's David Shanks on deceptions in social science research)

[3]CIBSE Technical Memorandum 61. *Operational Performance of Non-Domestic Buildings and Large Apartment Blocks*. CIBSE, 2020. (In production February 2020.)

[4]Gubb, C., Blanusa T., Griffiths, A., & Pfrang, C. (2018). Can houseplants improve indoor air quality by removing CO2 and increasing relative humidity? *Air Quality, Atmosphere & Health*. 10.1007/s11869-018-0618-9.

[5]Bunn, R. (2018). *Towards a theory of carrying capacity, evidence from long-term longitudinal case studies of occupant satisfaction in non-domestic buildings* (Doctoral dissertation, UCL (University College London)).

[6]Leaman A. and Bordass W., *Productivity in Buildings: The killer variables twenty years on*. Chapter 19: *Creating the Productive Workplace*. Clemence-Croome D. (ed). Routledge, third edition, 2017.