

# **BUILDING ENVIRONMENTAL ASSESSMENT METHODS: CLARIFYING INTENTIONS**

**Raymond J Cole**

School of Architecture, University of British Columbia

IN: *Building Research and Information*, Vol. 27 (4/5), 1999, pp. 230-246

## **1. INTRODUCTION**

As the realities of resource depletion and global environmental degradation become more evident, we can anticipate a maturing and strengthening of the public's concern and knowledge on environmental issues. This will translate into an expectation for greater environmental responsibility. As with other sectors, the building industry will be increasingly scrutinized and required to develop approaches and practices that address immediate environmental concerns and adhere to the emerging principles and dictates of sustainability.

Building environmental assessment methods have emerged as a legitimate means to evaluate the performance of buildings across a broader range of environmental considerations. The increase in development and application of such methods has provided considerable theoretical and practical experience on their potential contribution in furthering environmentally responsible building practices. Their most significant contribution to date has clearly been to acknowledge and institutionalize the importance of assessing building across a broad range of considerations beyond established single performance criteria such as energy. An important indirect benefit is that the broad range of issues incorporated in environmental assessments require greater communication and interaction between members of the design team and various sectors with the building industry, i.e., environmental assessment methods encourage greater dialogue and teamwork. Further, since assessment methods are implicitly a synthesis of current environmental knowledge related to buildings, they can play a significant role in focussing a broad range of research through a common filter [Cole and Larsson, 1998].

*Green Building Tool (GBTool)* was developed as a "second-generation" assessment method that built on the limitations of existing methods, and confronted areas of building performance assessment that were previously either ignored or poorly defined. The structure and scope of the assessment framework went through several changes over a two-year development period involving the collective input of the National Teams in fourteen participating countries. *GBTool* was primarily developed to describe and assess the thirty-four case-study buildings within the *Green Building Challenge '98 (GBC '98)* process. This application has enabled *GBTool* to be tested by a wide variety of users in the participating countries. Specific lessons from the *GBC* process are reported elsewhere [Cole and Larsson, 1998], providing the opportunity in this paper to explore some of the more fundamental issues emerging from this experience.

In the rapidly evolving field of building environmental research and practice, many players have different agendas and requirements. This inevitably creates different expectations of an

assessment tool. This was particularly evident during the *GBC* process where the National Team members involved in the assessment process consisted of academics, researchers and practitioners. A primary role of a building environmental assessment method is to provide a comprehensive assessment of the environmental characteristics of a case-study building. Broadly speaking, three distinct roles of building environmental assessment methods were evident during the development and testing of *GBTool*:

- Providing a common and verifiable set of criteria and targets so that building owners striving for higher environmental standards will have a means of demonstrating that effort, i.e., a mechanism to influence market receptivity and demand for higher environmental performance standards.
- Providing the basis for making informed design decisions, i.e., a design tool that can provide direction and guidance at all stages during the design development by highlighting priority issues and suggesting the possible trade-offs between options.
- Providing an objective assessment of a building's impact on the environment, i.e., a tool to evaluate energy and mass flows between built and natural systems and provide a common yardstick for measuring progress toward sustainability.

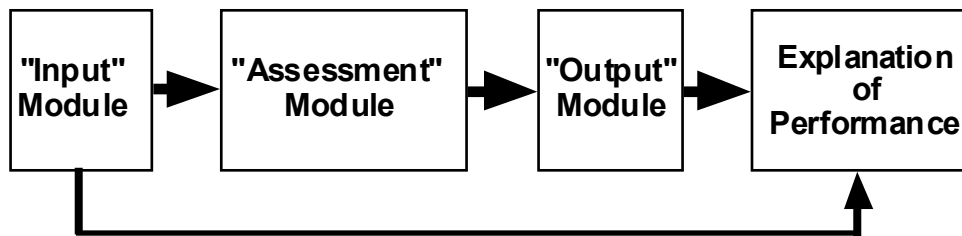
As with other assessment methods, *GBTool* offers some guidance in all three areas. This paper accepts the idea that evaluating the environmental merits of both completed buildings and of evolving designs is an important endeavour. However, it is the contention of the paper that having a clear idea of the overall intention of an environmental assessment and its anticipated audience is critical to its ultimate success. This requires making a distinction between the three roles identified above and making the distinction explicit in the structuring of the assessment method. The paper emphasizes this point by differentiating between what can be termed “green” assessment methods that offer relative assessments and “sustainable” assessments that offer absolute ones. While the distinction between “green” and “sustainable” is critical in framing the environmental assessment debate, it does not and should not preclude their coexistence in the structuring of current and future building environmental assessment methods.

Although the paper provides an explanation and rationalisation for the emphasis and structure of *GBTool*, its primary objective is to expose and clarify a broader set of issues related to the development of building environmental assessment methods.

## **2. KEY ASPECTS OF ENVIRONMENTAL ASSESSMENT**

*Figure 1* shows the key features that are either implicit or explicit in all existing building environmental assessment methods:

- The primary component is the “assessment” module in which performance scores are assigned to the various environmental criteria being scrutinized within the assessment process. The scope and structure of this module tends to form the major part of the discussion of assessment methods.
- A considerable amount of information about the case-study building and its context is required to facilitate an assessment. These are represented in *Figure 1* by the ‘input’ module. Although the input module *serves* the assessment module, the practicalities of data collection ultimately dictate the number and type of environmental criteria evaluated during an assessment.
- The results of an assessment must be summarized and communicated. Weighting is the mechanism by which a very large number of performance criteria are reduced to a smaller and more manageable number and is a critical part of the output module. The output forms the basis for interpreting the assessment results and should logically dictate the structuring of both the assessment and input modules.
- An output profile is not particularly valuable in and of itself. It must be accompanied by an explanation of the reasons why the overall score is what it is. This links back to the information contained within the input module and through it back to strategic decisions in the building or building design.



*Figure 1: Key Feature of Assessment Method*

The output module is currently a relatively poorly understood aspect of assessment frameworks. This is itself indicative of the fact that environmental responsible building design practice is still in its infancy. By moving into relatively uncharted areas, the uncertainties are also reflected in the current definitions of the goals and intentions of building environmental assessment methods.

### **3. “GREEN” OR “SUSTAINABLE”?**

Environmentally progressive building practice is currently described using a variety of different tags: “green design”, “ecological design” or “sustainable design.” Although discussions regarding the most appropriate terminology to describe environmentally progressive buildings can deteriorate to meaningless semantics, the distinction between the notions of “Green” and “Sustainable” is critical in structuring environmental assessment methods.

### 3.1 GREEN BUILDING PRACTICE

Existing building environmental assessment methods attempt to measure improvements in the environmental performance of buildings relative to current typical practice or requirements. Similarly, design guidelines are typically structured to offer direction on how to improve upon current design practices and only implicitly acknowledge sustainability as a goal. The assumption is that by continually improving the environmental performance of *individual* buildings, the collective reduction in resource use and ecological loadings by the building industry will be sufficient to fully address the environmental agenda. The choice of the term “green” building assessment was seen as a useful term to convey this intent within the development of *GBT*ool.

The notion of “green” permits is useful in the context of building assessment in that it can be extended to distinguish and discuss varying ‘shades’ of green. A “deep green” building may, for example, refer to one designed from the outset to maximize the use of solar energy, daylighting and natural ventilation, as well as harvest rainwater, treat any wastes on-site and use environmentally sound materials in the most efficient way. “Light green,” by contrast, may refer to buildings that have incorporate one or more green features such as high-efficiency windows, high recycled-content carpets or automatic shutoff systems for lights but are otherwise fairly conventional.

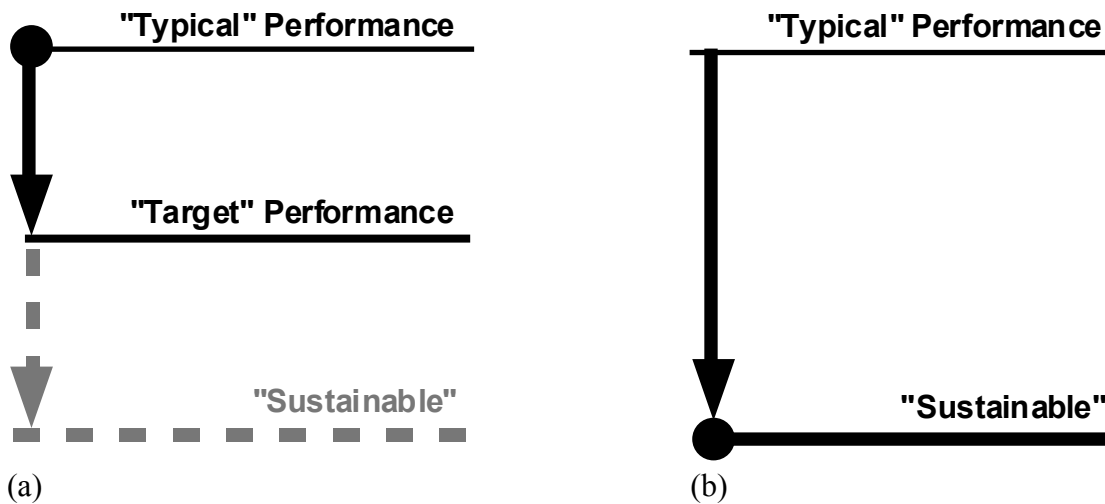


Figure 2: “Sustainable” and “Green” models of assessment methods

Assessment implies measuring how well or poorly a building is performing, or is likely to perform, against a declared set of criteria. *Figure 2a* illustrates the defining characteristics of a “green” assessment. Within the definition of green offered above, the primary characteristics of a “green” building assessment method logically emerge:

- Assessments are made relative to “typical” practice without having to define an ultimate goal.

- To define an assessment scale and assign scores to the performance, it is necessary to declare a demanding “target” performance level. This is a demanding performance that can be progressively increased as “green” design matures.
- Since “green” assessment methods are invariably used as a mechanism for encouraging building owners and designers to aspire to higher building environmental performance, the range of issues is considerably larger than that necessary to assess whether it is sustainable.
- Given their role of encouraging higher levels of environmental performance, “green” assessment models place a higher premium on comparing the performance at a regional and local scale where building owners and developers can demonstrate a marketing “edge” over their competition. “Green” assessments place less emphasis of international comparability.

The majority of existing “green” environmental assessment methods are voluntary in their application and have the primary objective of stimulating market demand for buildings with improved environmental performance. An underlying premise of voluntary assessments is that if the market is provided with improved information and mechanisms, a discerning client group can and will provide leadership in environmental responsibility and others will follow suit to remain competitive. Voluntary assessment protocols must serve two conflicting requirements -- they must function as an objective and sufficiently demanding metric to have credibility within the environmental community, while simultaneously being attractive to building owners who wish to have something positive to show for *any* effort that they have placed on environmental performance. Satisfying these twin requirements invariably compromises both the number of criteria that are assessed and where the benchmarks are set before performance points are earned. [Cole, 1998] The “acceptance” of existing assessment methods currently derives largely from their voluntary application. Given the practical (and incentive) constraints on setting demand targets and dependency on market acceptance, it is uncertain whether this mechanism will be sufficient to create the necessary improvements in environmental performance of buildings needed to meet broader national environmental or sustainability targets.

### **3.2 SUSTAINABLE BUILDING PRACTICE**

Sustainability has emerged as an overarching notion for the environmental discourse and must, therefore ultimately give direction to the structure and application of environmental assessment methods. Sustainability has environmental, social and economic dimensions, embraces all facets of human activity (e.g., industry, transportation, food production etc.), and spans local actions through to redressing the major inequities that exist between developed and developing nations. Given the political and economic interdependencies wherein the actions of one nation profoundly affect others, the notion of “sustainability” is perhaps only meaningful when applied at a global scale.

Irrespective of the social and economic context, the health of the biosphere is the limiting factor for sustainability. The unabated growth in the “throughput” of energy and material to

satisfy human demand is critical. As Rees [1999] argues “empirical evidence suggests that resource consumption already exceeds the productive capacity of critical biophysical systems on every continent.” He further suggests that “waste production already breaches the assimilative capacity of many ecosystems at every scale.” Since the management of local, regional and global mass and energy flows is of vital importance, physical indicators describing these flows must logically dictate the emphasis of any methodology attempting to assess a sustainable approach to human settlement and building. Kohler [1999] criticizes *GBTool* and other existing assessment methods based on relative performance as both hiding “the real mass and energy flows which determine the effective environmental impact” and the “differences in impact between individuals and different countries.” Cooper [1999] elaborates on this limitation and suggests that unless methods for assessing the built environment are capable of measuring performance against carrying capacity criteria, “their ability to contribute to the sustainability debate is likely to remain limited.”

*Figure 2b* illustrates the role that an environmental assessment method would within the context of sustainability. The two defining points on the assessment scale would now be “typical” practice and “sustainable” practice. Assessments would be made of the extent of the progress that the building performance has made *toward* a declared, “sustainable” condition.

- It is possible to define sustainability goals at a *global* scale in terms of the relationship between resource use, assimilative capabilities of the biosphere, carbon sinks, albeit in general terms. Buildings are also clearly connected to systems at much larger scales, the characteristics of which may change radically over time. Without reference to their wider context, it is somewhat more difficult to define specific “sustainability” goals for individual buildings. The use of “sustainable” targets such as zero fossil fuel use, zero greenhouse gas emissions, zero potable water use and zero sanitary waste entering municipal systems, implies that all future buildings should become more “autonomous.” It is unclear at this time what environmental strategies are most appropriately addressed at the building or community scale, or whether autonomy is an appropriate goal at either of them.
- Assessments require an understanding of the *absolute* impact or stress that building design and operation place on ecological systems to ensure that it is within the productive and assimilative capabilities of the local, regional and global ecosystems. This would require an extensive understanding and quantification of the complex links between building decisions and ecological loadings -- a goal that will not be fully attainable in the foreseeable future and may, in fact, never be completely possible. As such, within a sustainability model, assessment would be directed at identifying the absolute amount energy and mass flows associated with buildings. Here, reductions in absolute resource use and ecological loadings by buildings would be indicative of a positive move toward ecological sustainability. Methods are emerging that set aggregate human resource use and ecological loadings against the finite productive and assimilative limits of the biosphere. For example, “ecological footprinting” estimates the area of productive ecosystems all over the world whose biophysical output is appropriated for the exclusive use of a defined human population. The ecological footprint concept is instructive in illustrating the immense disparities between urban and non-urban settlements, between

first and third world countries and revealing differences between alternative forms of housing (high-density apartments versus single family dwelling). [Rees, 1998].

- Sustainability goals are necessarily dependent on the actions of others. Sustainability goals for *buildings* can only be meaningfully defined having first made assumptions about global or regionally averaged sustainable rates of activity, e.g., a two-fold increase in human population, changes the necessary reduction in resource use and ecological loadings by a factor of two. Given the uncertainties and time dependent nature of these assumptions, sustainability goals must be set within declared scenarios.
- The number of criteria required to judge a building's performance with the context of ecological sustainability can be relatively few with the *proviso* that the performance indicators are carefully selected. For example, Lowe [1996] argues that most other physical and many non-physical indicators of sustainability are, for example, statistically and causally linked to carbon emissions. As such, strategies reduce carbon emissions to a sustainable level would carry a host of other improvements that would not have to be independently assessed.
- Physical indicators of sustainability would be normalised by some measure of the total sustainable level of activity described by that indicator. Satisfying the human principle of equity would suggest that denominators represent equitable shares of the total sustainable level, e.g., using per capita share of the total carbon sink capacity to normalise carbon emissions.
- Assessment methods set within the context of sustainability offer the powerful advantage of international comparability. Lowe [1998] points out that office buildings that consume less than 100 kWh(delivered)/m<sup>2</sup>/year or emit less than 30 kg (CO<sub>2</sub>)/m<sup>2</sup>/year would exemplary anywhere. Moreover he suggests that if, for example, the reason why buildings perform better in Denmark than equivalent buildings in other countries is because of efficient district heating systems, this offers important direction at the energy systems and urban design levels.

#### 4. SCALE OF MEASUREMENT

As raised in *Section 3* above, all assessment methods implicitly embody a scale of measurement. Such a scale forms the basis for allocating *performance points* that are subsequently used to obtain an overall performance score. A primary emphasis of assessments is, therefore, to use the selection of the criteria to define the *direction* of environmental progress and to measure the *degree* of progress being made in improving the performance of buildings either relative to other similar buildings in the case of “green” building design, or natural production and assimilative capabilities of local and global ecosystems in the case of “sustainability.”

#### 4.1 REFERENCES AND BENCHMARKS

Irrespective of the goals of a building environmental assessment, it is necessary to characterise current performance levels. A common, but often un-stated, baseline for assessment is a 'typical' or 'average' performance and, as such, recognition is given for better

than industry norm performance. If scrutinized, this choice of benchmark is an extremely difficult one to both define and quantify across all assessment criteria in a consistent manner.

*GBTool* attempted to define explicit reference performance levels for all performance criteria. The original proposal was to have the National Teams establish “reference buildings” to establish benchmark performance levels. A reference building was considered as a building of the same size and type as the case-study building, but designed assuming industry norms. The use of the reference building concept is well established in energy simulation procedures in North America, but in *GBC '98* it was proposed to extend the concept to cover a wider range of issues. It is assumed that this reference building would characterize industry benchmarks for that building type and region across all applicable performance issues and provide a base for performance scoring that could be derived and stated with some confidence.

Though the notion of having explicit benchmarks as an essential requirement of an assessment method was unanimously accepted, the use of a *Reference Building* proved universally problematic during the *GBC* process, both conceptually and in practice. Creating a performance profile for a ‘second design’ was seen as unrealistic task, requiring time and effort that were not commensurate with the anticipated overall scoring goal. For those criteria that required a numerical reference (energy use, water use etc.,) National Teams typically chose national figures derived from national data bases and other statistically based sources.

Typically, performance information is normalized in some way to facilitate comparison. For example, energy use is typically compared on a per m<sup>2</sup> basis to normalize for size, or per degree-day to account for variations in climate. Propositions were made to introduce normalization for occupancy (i.e., by introducing a per person factor.) to account for differing use patterns and operating schedules. Normalization becomes less critical if the complete definition of the *Reference building* is used, since the case study building is compared to the performance of a similar sized building (eliminating /m<sup>2</sup> issues), in the same location (eliminating climatic differences) and same use (eliminating occupancy differences). However, problems may exist when the benchmark is derived as a statistical average value. Even though these statistics may be normalized for area (e.g., annual energy use/m<sup>2</sup> of floor area) the local climate conditions, occupancy patterns and operating schedule for the case-study building may be radically different from the average.

#### **4.2 TARGET PERFORMANCE LEVELS**

Assessment methods require the declaration of a target or upper level on the assessment scale. An important issue when assessing *relative* performance is whether the measurement scale is “open” or “closed.” An open scale has the advantage of accommodating performances in advance of those initially anticipated when defining the upper limit of the assessment scale. *GBTool* chose a closed assessment scale for all criteria and set demanding targets for each. This offered advantages in the application of weighting factors to summarize the performance results and the ability to provide a common format for the output profiles. However, as discussed in *Section 9.2* of this paper, this diminishes the ability to reveal and highlight priority issues.

Some of the performance targets in *GBT* were set in absolute terms such as zero CFC/HCFC emissions or 100% reuse of the floor area of an existing building. The majority were set at a percentage of current typical practice, e.g., 75% of reduction in operating energy use compared to that of a reference building. The choice was to be both demanding, yet within the bounds of attainability with current knowledge and existing technologies. An assumption implicit in having fixed target performance levels is that they can be “ratcheted up” in later versions of the assessment tool as experience develops.

An underlying premise in the *GBC* process is that a common set of features for building performance assessment procedures can be defined that are applicable to all buildings in all regions. Further, if these ‘core’ criteria are made explicit, they can provide a clear starting point for developing customized methods for specific building types, geographic regions and specific intentions. The customizing of the assessment scales by the various National Teams was a further illustration of the different agendas that currently define building environmental assessments. Whereas some National Teams either accepted the relative default assessment scales, other replaced them with absolute performance values. The setting of performance goals for individual criteria is more complex than initially envisaged and it is evident that greater consideration must be given their aggregate effect. This, by implication, requires an explicit recognition of the context and the involvement of a wider range of participants in deriving the goals and targets.

### 4.3 SCALING INCREMENTS

There are two general approaches to summarizing the results of a performance assessment:

- A simple designation of a number of points for achieved performance in each of the various environmental areas, using different scoring systems for each and without concern for the significance of one criterion relative to the others. By assuming that all the performance criteria are of equal importance, a simple aggregation is used to provide a total score.
- Using a common scale as the basis for assessing for all and applying weightings to acknowledge the different significance of each criterion prior to deriving the aggregate score.

Existing assessment methods typically use a different scale of measurement for different performance issues and often identify a number of points or credits available for specific criteria without any explicit declaration of why or how they relate to others. *GBT* uses a consistent scale for all assessments and explicitly declares the benchmarks: the zero (0) on the performance scale for “typical” practice and five (5) for the most demanding performance. A negative value is included to account for performances worse than typical. This common –2 to +5 scale was used for all assessed sub-criteria and criteria. In the *GBC* scoring systems, a consistent score of zero (0) would represent a building conforming to typical practices and standards in the region. A consistent score of five (5) would represent the best technically feasible performance in those areas deemed significant within the *GBT*. This, by definition, represents a “green building” within the *GBC '98* process.

Most of the scores for the various criteria were presented on a scale that gave more points for a given increment in performance as the overall performance level increased. The rationale being that it becomes increasingly more difficult to attain performance improvements. One National Team changed the performance scale to linear scale in even 10% steps, but using the same intention as the original scale. Their argument was that assessments should be based on environmental impact not the difficulty to reach specified levels, otherwise costs should also be considered explicitly in the assessment. This again raises the distinction between voluntary assessment tools designed to encourage participation by building owners and developers, and assessment tools that attempt to give an objective measure of the environmental consequences of buildings. In the former case, the performance requirement for a '5' score is usually tempered as a result of cost and implementation difficulties, whereas the latter is defined as theoretically desirable performance irrespective of such problems of achieving it.

## **5. POTENTIAL OR ACTUAL PERFORMANCE?**

There is sufficient evidence to show that a building's performance in use is often markedly different from that anticipated or predicted during design. As such, an important decision in developing assessment methods lies in the choice to evaluate the "potential" or "actual" building performance.

- The obvious advantage for assessing the *actual* performance of the building in use is that it captures what resources are consumed, what ecological loadings are generated and the actual indoor environmental qualities and occupant responses to them.
- Beyond external factors such as specific weather conditions during a specific time period, actual performance depends on the actual behaviour of occupants, tenants and building operators. This brings into play many idiosyncratic operational factors that may not be generally applicable to other buildings.
- The assessment of potential performance is based on assuming normal or default patterns of occupant behaviour and building operation, making it easier to distinguish between improvements in the physical features of buildings and improved efficiencies in their use and operation.

Real performance data is clearly of significance within the sustainability agenda where the primary objective is to assess the absolute impact of buildings. It is also of considerable importance in providing experience and feedback to the design community as to what does and does not work in practice – a critical concern in a rapidly evolving field.

Although potential performance is less "real," it can still produce useful information to guide the future actions of developers, owners, designers and anyone else who is involved with the production of buildings. However, in recognition of the current mismatch between anticipated and actual building performance, there is an obvious need to recognize the relationship between strategic design and the ease with which a building can be managed and operated, i.e., making the design of the building management and operational systems part of the building design and procurement process (Bordass and Leaman, 1997).

## **6. QUALITATIVE AND QUANTITATIVE CRITERIA**

A defining characteristic of the *GBT* is that it has embraced a broader range of performance issues than that found in existing assessment methods. Existing methods temper the range of assessment issues by remaining within the bounds of objective, scientifically acknowledged and verifiable issues. In this sense they only provide a partial view of environmental performance. However, moving into new areas where the measures of the performance are currently poorly defined, requires more qualitative descriptions of the measurement scale. Such scoring techniques can be easily criticised as lacking the objectivity necessary to establish trust in the assessment system:

- Criteria expressed qualitatively are open to wider interpretation by assessors and therefore, the assigning of points can vary considerably depending on those making the assessment.
- It requires a great deal of time, energy, and commitment from an unbiased third party to be successful.

A distinction can be made between assessing “green” and “sustainable” performance. Assessing sustainable performance, which is largely an issue of energy and mass flows, can and should be described in quantitative terms. The wider range of performance issues necessary within an assessment of “green” performance currently cannot avoid using more qualitative metrics to evaluate a building comprehensively. The critical issues here seems not that qualitative criteria should be excluded from building environmental assessments but that:

- Greater care and precision has to be given to the descriptions of the assessment scales for qualitative criteria to reduce misinterpretation.
- Within the presentation or summarizing of performance results, the qualitative criterion scores are kept distinct from the quantitative performance data that is assumed to be more objective, reproducible and therefore more reliable. This would avoid the perception that after a massive effort of data collection and input, the final performance scoring and profile can be potentially skewed by a subjective judgement.

## **7. ASSESSMENT OR DESIGN TOOL?**

Although conceived solely as an “assessment tool” to evaluate a completed building design, many of the participants raised the issue of using *GBT* as a “design tool.” The question emerges as to whether a single tool can function equally effectively as an assessment and design tool? And what compromises would be necessary to an assessment tool to enable it to be useful in design? The answers lie in the structure of the assessment framework and with the skill and enterprise of the user.

Selected criteria within assessment methods are currently being adopted as part of broader sets of design guidelines and specifications and are gradually diffusing through the design

community in this form. However, since environmental assessment methods present an organized set of selected environmental criteria, by default they communicate to building owners and design teams what are understood as being the most significant environmental considerations. As such existing assessment methods are used as design tools, even though they were not specifically designed to do so.

A considerable amount of building design-relevant information has emerged on broad range of environmental issues, far more than are currently incorporated in assessment methods. A tool designed to provide guidance on design would therefore require more detailed information than one intended for assessments but, by necessity, must still be practical in its application. Design tools for environmental assessment must:

- Be based on information that is accessible during design.
- Identify critical environmental issues and provide guidance on a range of possible design strategies to address those issues.
- Quickly assess the relative environmental benefits gained by using a particular strategy or set of strategies early in design development and compare alternative schemes, i.e., enable early scoring or preliminary scoring to facilitate timely decisions by the design team and clients.
- Permit the data needed by the assessment tool to flow seamlessly from the tools the designer uses across the design process, e.g., all the data on building area are imported automatically from the CAD tool.
- Make links with other design criteria.

Given the arguments of potential verses actual performance, the availability of information and the importance of nurturing regionally appropriate strategies, design tools logically relate more easily to methods that assess “green” performance than “sustainability.”

## **8. BOUNDARIES OF ASSESSMENT**

The scope and “boundaries” of an environmental assessment method are critical. *Figure 3* shows a conceptual framework that can be used to illustrate the scope and boundaries of environmental issues in current building assessment methods. It consists of three primary “dimensions”: *Criteria*, *Time* and *Scale*.

### **8.1 CRITERIA**

The *Criteria* dimension references the extended set of considerations within environmental assessment, distinguishing between *ecological* concerns (resource use, ecological loadings etc.) and *human* concerns (indoor environmental quality, economics etc.) Each of these sets of issues can be further subdivided into:

- Performance criteria that can currently be quantifiable and that can be confidently defined and assessed, such as energy use, water use etc. These are shown as solid lines.
- Performance criteria that currently only be described qualitatively such as loss of biodiversity etc. These are open to wider interpretation and therefore their assessment is less certain. These are shown as broken lines.

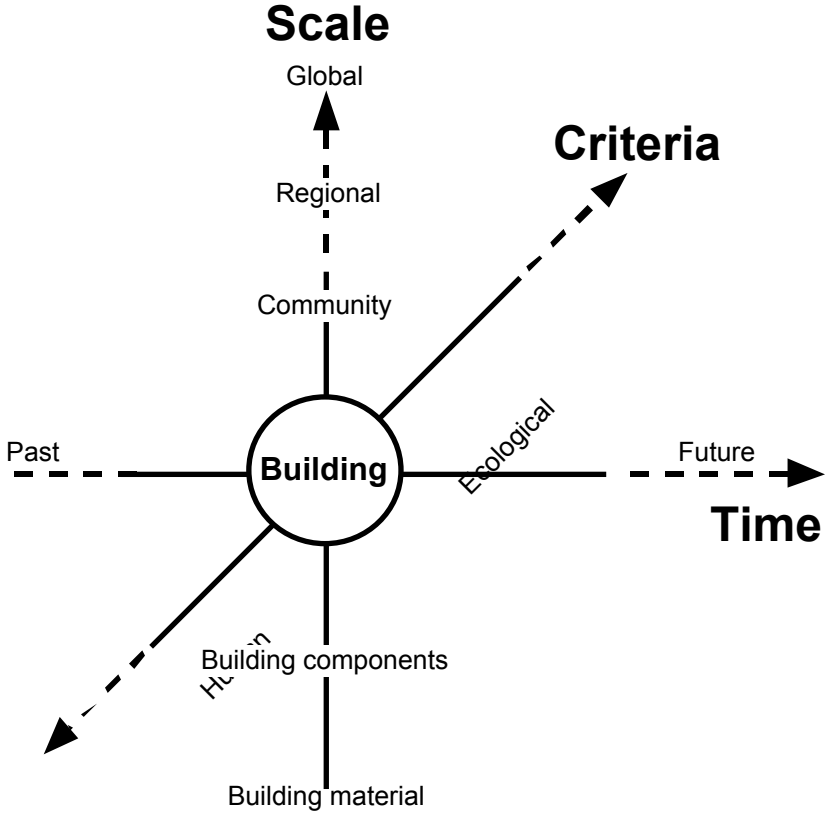


Figure 3: Three dimensions of environmental assessment – Scale, Time and Criteria

**8.2 TIME**

The *Time* dimension is that explicitly covered with Life-Cycle Assessment methodologies. The notion of *Life-Cycle Assessment* has been generally accepted within the environmental research community as the only legitimate basis on which to compare alternative materials, components and services and is, therefore, a logical basis on which to formulate building environmental assessment methods. Adopting *Life-Cycle Assessment* approaches would seem an appropriate basis for structuring *Performance Criteria* within building environmental assessment tools but may not be possible for all criteria. [Beenstra, 1997].

In *Figure 3*, both the *distant* past and long-term future are less clearly known and certain than the *immediate* past and future. As such they are distinguished by □ periods of relative confidence (shown as a solid line) and speculation (shown as a broken line) respectively. □

### 8.3 SCALE

Whereas a considerable strides have been made in the environmental performance and life-cycle assessment of individual materials and components as well as their aggregation to whole building performance, the links between building and community and regional scale are less well developed. As with other assessment methods, *GBTTool* expanded the criteria to include *contextual* issues that relate to site selection, building location and proximity to amenities. Baldwin [Baldwin, 1998] presents the life-cycle energy profiles of two UK office buildings – *Building 1*, a 1970s prestige air-conditioned office building in the centre of London well served by public transit and *Building 2*, a late 1980s atrium building near to the centre of Manchester with good car parking facilities. The critical issues are that the magnitude of the staff travel energy is similar to that of the building construction and operating energy and that there are clearly marked differences associated with the mode of transport (See *Figure 4*).

Though building location and other contextual issues are important, whether or not they can be *controlled* by the design has created significant discussion regarding their legitimacy for inclusion in either a building assessment or design tool. This debate reveals the current gaps between modeling and assessing *building* environmental impacts and *community* environmental impacts and, more generally, between the disciplines of architecture and urban planning.

*Scale* is clearly the critical dimension necessary to fully discuss building environmental performance in a comprehensive manner and, as has been emphasised earlier, is a prerequisite within the context of sustainability. The individual building, though useful in the “green” building debate, is an inappropriate scale to define and discuss optimal environmental performance within a sustainability model.

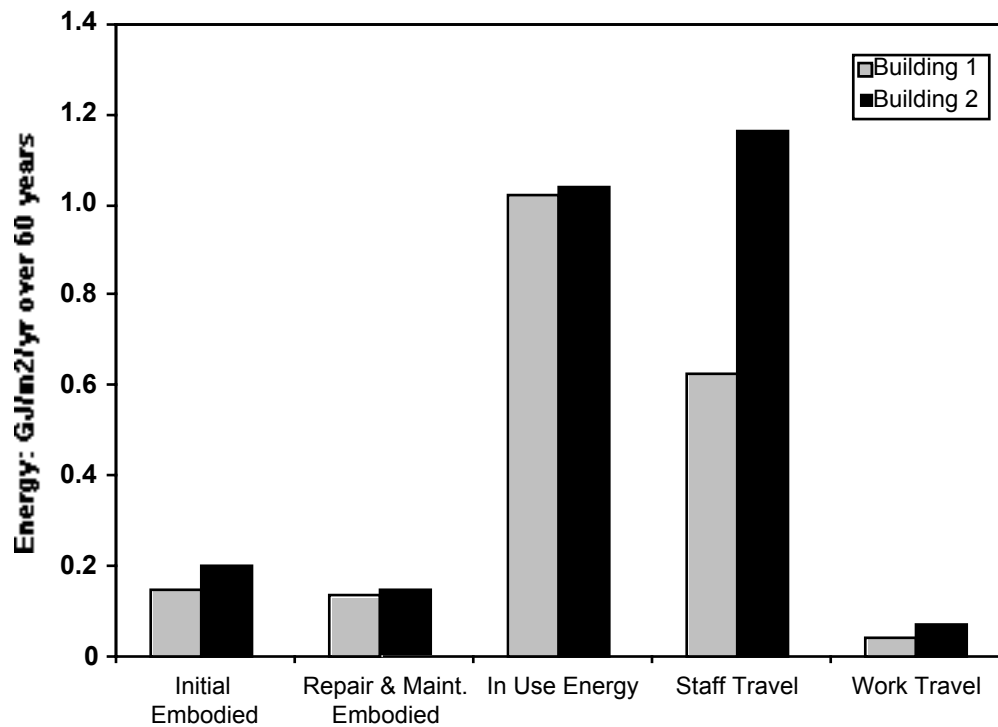


Figure 4: Annualised Life-Cycle Energy (per m<sup>2</sup> of floor area) for two UK Office Buildings (from Baldwin, 1998)

## 9. COMMUNICATING THE RESULTS

Although it is generally accepted that environmental criteria must be organized in ways that facilitate meaningful dialogue and application, the structuring of criteria within the assessment method is most important during the output of the performance evaluation. It is at this stage that the complete performance profile of the building is evident and when the ‘story’ of the performance must be told in a coherent and informative way.

### 9.1 OUPUT FORMAT

Based on arguments presented earlier in this paper, an effective output profile from a building environmental assessment should acknowledge the following:

- Comprehensive view of a building’s environmental performance.
- Selective scrutiny of various performance areas
- Enable comparisons

#### 9.1.1 Comprehensive view

Since the primary strength of building environmental assessment methods is their comprehensiveness, the output must provide an overall picture of the performance. Clearly there are practical and cost implications associated with data collection and assessment – the

more criteria the greater the difficulties. There are also limits to what can reasonably be comprehended from an output file.

*GBTool* is structured hierarchically in four levels, with the higher levels logically derived from the weighted aggregation of the lower ones: *Performance Areas*, *Categories*, *Criteria* and *Sub-Criteria*. The *nesting* principle of four levels is a critical feature of the framework in that it enables a building performance to be described at successively detailed levels. It was initially conceived that it would also be possible to make assessments at the various levels and thereby expedite the process for gaining a quick overview of building performance. However, it now appears doubtful that it will ever be possible to make simple and single evaluation of the efficiency of, for example, building *Resource Use* without an aggregation of the assessments of the constituent resource issues (energy, land, water and materials). *GBTool* is therefore currently only usable if and when one starts from the most detailed level of sub-criterion assessment and building and proceeds upwards through the criteria by means of a weighting process, to the category scores. This remains the only meaningful way to describe and report on building performance in the *GBTool*.

### **9.1.2 Selective scrutiny**

Different aspects of the output may hold greater interest for different users, and thus the output must allow scrutiny of more detailed areas of performance. The nesting principle discussed above provides an elegant means to view performance in detailed or general terms, and the performance areas distinguish between qualitatively different environmental issues. It is evident that greater partitioning of the performance results is necessary:

- The separation of more objective assessment criteria and scores from those that are more open to interpretation would improve the confidence given to any aggregated score.
- The partitioning of building criteria and operations and management performance issues. The recently released version of BREEAM [Baldwin *et al*, 1998] has been structured in modules to recognize differences between design and operation and the assessment of new and existing buildings.

### **9.1.3 Enable comparisons**

Whether the assessment method is a design tool, suasion mechanism or stand-alone assessment, an important requirement is that it facilitates comparisons between the performance of the case-study building performance and other known and declared references. *Figure 5* schematically shows the output profiles of five performance criteria for two buildings, and highlights four types of “comparisons” that may be expected to be made:

- For a specific building performance criterion, the requirement of assessing relative to a declared benchmark. This is a requirement of all assessment methods and the choice of benchmarks by which a criterion is measured is a defining characteristic of an assessment method.
- A comparison of the performance score of one criterion with that of others for the same building. Given that green buildings are recognized as much by the integration of systems

and strategies, revealing the individual performance scores side-by-side in an output bar diagram can highlight where trade-offs and compromises had been made.

- For a specific performance criterion, the requirement of comparing with other buildings either in the same location or internationally. This comparison raises a host of issues regarding the use of “relative” or “absolute” scoring values and how the performance values or scores are normalised.
- A comparison of the overall performance profile with that of other buildings, again either locally or internationally. Invariably this requires the reduction of the overall assessment score to a single value or label.

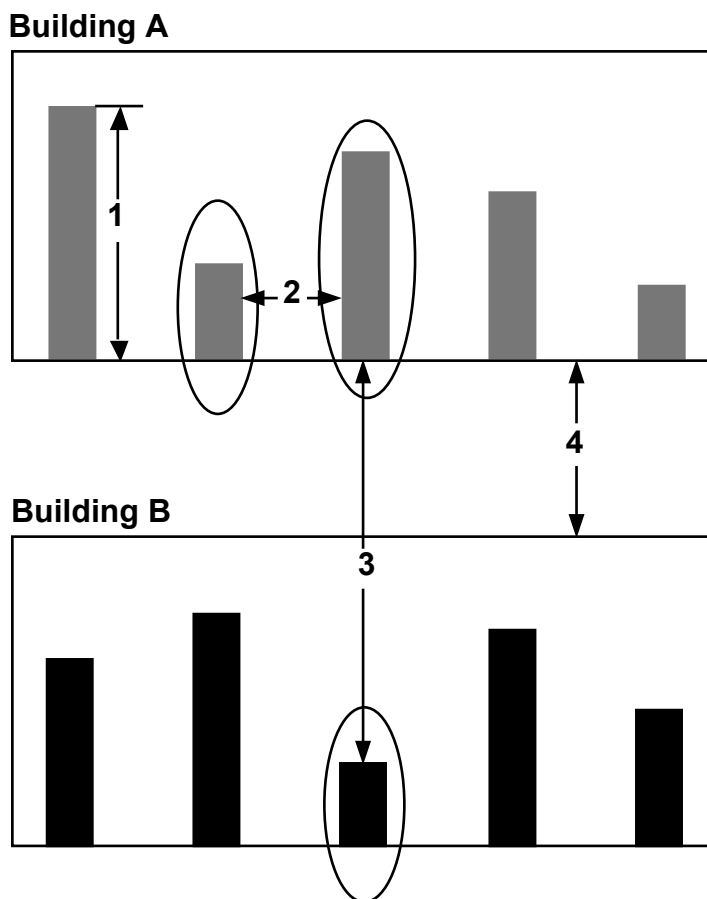


Figure 5: Types of comparisons made in assessments

## 9.2 WEIGHTING CRITERIA

Building environmental assessment methods cover a wide range of performance issues, e.g., *GBTool* assesses approximately 120 individual sub-criteria and criteria. It is necessary to reduce these assessment scores to a manageable number in the output modules. Weighting is now recognized as an essential part of building environmental assessment methods although not without contention. The two critical issues within this debate are the basis for deriving

weightings and the manner in which the weighting process affects the interpretation of the aggregated result.

Within the *GBC* documentation a series of criteria were offered as a basis for developing appropriate weightings such as: Is the effect upon the environment irreversible? Is the effect upon the environment long lasting? What number of people are affected by the issue covered within the sub-criterion or criterion? Does the practice in question have momentum that will require an extraordinary effort to counter? However, no clearly defined methodology was proposed for their application. An important criticism was that these recommendations mix the *importance* of a criterion or sub-criterion in terms of its effect on human health, wellbeing, or the environment, with the *difficulty* of achieving it. This implicitly identifies a further manifestation of a recurring theme throughout this paper – whether the assessment is attempting to give an objective environmental profile of a building, or to acknowledge practical and cost implications of attaining improved performances. Todd [1998] identifies that in deriving appropriate weightings, “the key to understanding the relative importance of environmental criteria lies in the selection of final endpoints - ones which reflect potential impacts on the environmental components of concern, not simply the changes in quality or quantity of environmental media (air, water, soil). Thus the question of importance should not be whether air pollution is more important than water pollution, for example, but instead whether air pollution or water pollution exerts a greater specific potential impact on endpoints of concern.” Although this represents the most conceptually appropriate direction for deriving weightings in environmental assessment, the development of the attendant links and relationships between buildings and impacts advocated in the approach will require considerable research and data collection before it can be fully realized.

Todd’s approach to the derivation of weightings in effect seeks the *equivalency* between the impacts of various resource use or ecological loadings. This notion is currently applied in other aspects of environmental performance, albeit in a more modest way. For example, greenhouse gases (CO, NO<sub>x</sub>, and CH<sub>4</sub>) are combined based on their *CO<sub>2</sub> equivalence*, or the *Ozone Depletion Potential* of various refrigerants is specified in terms of their *equivalence* to effects of R-11. Papamichael and Protzen (1997) challenge the validity of using weightings other than in these circumstances where the relative significance of the constituent components can be confidently stated. These concerns centre on the inability to derive relative weighting with any precision and interdependence of many performance criteria, particularly those that are by nature more qualitative.

Weighting is linked to the “nesting” principle in *GBT* and the desire to be able to present performance scores in varying degrees of detail. In this context weighing represents an explicit declaration of the importance of a criterion against others. Although weighting is used extensively within the process, the current graphic presentation in *GBT* does not make immediately apparent which environmental categories should have priority. This can lead the reader to conclude that all the issues are of equal importance. This is the disadvantage of using a uniform scale of –2 to +5 throughout, and could not be easily rectified by simply redesigning the graphic format.

Given the difficulties discussed earlier, the need for *three* levels of weighting in order to obtain a summarized score for one of the major performance areas (e.g., Resource Use), places considerable doubt on the validity of the exercise wherein the limitations of weighting are compounded at each application. Without weighting, however, aggregation of the results would not be possible and it would be necessary to present the performance scores of all the assessed criteria in a direct way. Inevitably this would significantly limit the number of criteria that could be reasonably communicated in the output profile. Considerably more work is clearly required in the applicability of weighting protocols within environmental assessment and the way in which one can meaningfully interpret aggregated scores.

### **9.3 EXPLANATION OF PERFORMANCE**

The notion of 'environmental labeling' is often used in conjunction with environmental assessment as a logical outcome. The labels currently used are typically a classification of the performance into descriptive categories such as *Fair, Good, Very Good* or *Excellent*. In BREEAM, for example, this type of categorization is based on the variation in the number of credits possible for the particular building being assessed. Similarly, the summary of performance in the US LEED program is judged as meeting a *bronze, silver, gold* or the best, *platinum* performance benchmarks through the simple addition of the various performance scores [USGBC, 1996].

In *GBTTool*, the 0 to 5 scale is applied at the sub-criterion, criterion, category and indeed whole building level. The issue emerges as to where on the scale does and on what basis does one begin to define an overall 'good' building. Only after a detailed review of the performance profiles of the case-study buildings examined in the *GBC* process will it be possible to overlay thresholds which begin to denote ranges of performance that are deemed outstanding or otherwise. Again, this would have to acknowledge the regional context of the assessment.

#### **9.3.1 Links**

A performance profile offers a graphic display of the scored criteria that signals areas of progress relative to declared benchmarks. But this is only a means to an end – the primary roles are:

- *Link to cause:* There must a means of explaining why the performance is what it is – good or bad. That is, the output must provide a link back to its cause or origin. Whereas some of the characteristics of the building that were collected to perform an assessment, additional information may be required to explain the performance.
- *Link to action:* Since the output represents the *link with action*; the output must link with information that offers a basis for improving on deficient performance.

#### **9.3.2 Non-Applicable Criteria**

The nesting principle inherent in *GBTTool* is a powerful mechanism of summarizing performance scores. However, similar to the significant influences that weightings can have on the overall score, other user-controlled features can influence the results.

To permit more customization of *GBTool*, it was intended that National Teams/design teams would designate criteria as non-applicable if not relevant in the region or case-study building, and reapportion the weightings to the remaining ones. This was a widely used approach, but highlighted one significant issue in customizing a general assessment scale to regional application. In those countries where environmental standards are more demanding, what had been designated as the best performance (i.e., 5) in the general scale, was typical practice regionally. This raised the question whether a non-applicable (since every other building in the region would meet this performance) or 5 should be assigned. Resolution of this debate raises a host of questions about the structure of *GBTool*:

- There is a distinction between environmental criteria that have been man-dated in one region and not in others, i.e., if something is legally banned, then compliance is mandatory, compared to if the criterion is voluntary.
- If the maximum performance level (5) is an absolute limit (e.g., zero CFC/HCFC emission, zero potable water consumption etc.) of whether the maximum performance level could be bettered. The issue here is constraining nature of a closed scoring system versus an open system which enable differences to be shown between buildings even in a region where the average performance was very high.

## 10. DISCUSSION AND CONCLUSIONS

The notion of a universally applicable tool that will be widely adopted in different countries is questionable. However, the notion of a collectively developed tool that exposes and addresses controversial aspects of building performance and from which the participants can selectively draw ideas for incorporating and modifying their own tools, remains a valid role for *GBC* process. The development and testing of *GBTool* exposed numerous detailed technical issues that must clearly be addressed in future versions. But the process also identified several critical concerns that can help clarify how building assessment methods should be structured in more fundamental ways.

This paper has attempted to clarify the distinction between “green” and “sustainable” agendas and their implications for the development of building environmental assessment methods. Assessing “green” performance is more fundamentally rooted in the practicalities of current building delivery and is therefore perhaps more readily acceptable to building owners and practitioners. This is clearly both an advantage and a limitation. There is an implicit assumption in existing methods that “green design,” by continually reducing resource use and ecological loadings, is charting a sustainable path. Whether or not this is the case is uncertain at this time. However, arguments presented in the paper suggest that a significant realignment toward the more holistic “sustainable” model will not be possible until the links between “building” performance and larger scales such as community are acknowledged. The “individual” building is a too constraining level to define “sustainable” practice and the next significant advances in environmental assessment methods will invariably have to explicitly extend boundaries of analysis.

Although the distinction between “green” and “sustainable” is essential to help clarify the various roles of building environmental assessment methods, the considerable practical overlap between the two agendas suggests that they can be reconciled within a single tool. *Figure 6* shows conceptually how this may be possible, and illustrates that the problem is primarily one of partitioning of the performance issues while simultaneously clarifying and making explicit links between them. The broadest distinctions are:

- “Green” performance is most usefully described in relative terms in comparison to similar buildings in the regions, while absolute energy and mass flow measures are a prerequisite for assessing progress toward sustainability. A fewer number of carefully selected performance measures are required to provide a measure of a building’s role in ecological sustainability than to describe its “green” performance.
- In both “green” and “sustainability” assessments, all performance measures would ideally be “actual” performance values rather than predicted. Although the issue here is primarily one of the practicality of data collection and quality, the use of actual performance values is more critical in assessing progress toward sustainability.
- Within “green” assessment, many performance issues must be customized to account for building or region-specific considerations. Similarly, in deriving summarized descriptions of building performance, weightings would be customized to reflect those priorities. Experience suggests that too much customization was permissible within the *GBC* process and compounded by a lack of clear guidance on how to change the default values. [see Cole and Mitchell, 1999] It would be appropriate to seek consensus among the National Teams and fix the criterion descriptions, scales and weightings for critical performance issues. These would certainly be the ones that are selected as the basis for indicators of progress toward sustainability. Other performance criteria could be open to greater customization with the proviso that they are kept distinct in the output.
- Todd [1998] suggests that providing qualitative and quantitative assessment scales for many of the green criteria would enable alternative types of judgements to be made, particularly where the data for the more desirable quantitative assessment is either not available or prohibitively expensive to acquire.

By venturing into relatively uncharted areas the *GBC* process has exposed a great many issues regarding the scope and structuring of building environmental assessment methods. It has also clarified the many roles and applications demanded of these tools. Since *GBC* process is now entering a second iteration, the expectation is that many of the issues addressed in this paper will be manifest in the restructuring of *GBT* and its application again contribute the wider debate on environmental assessment.

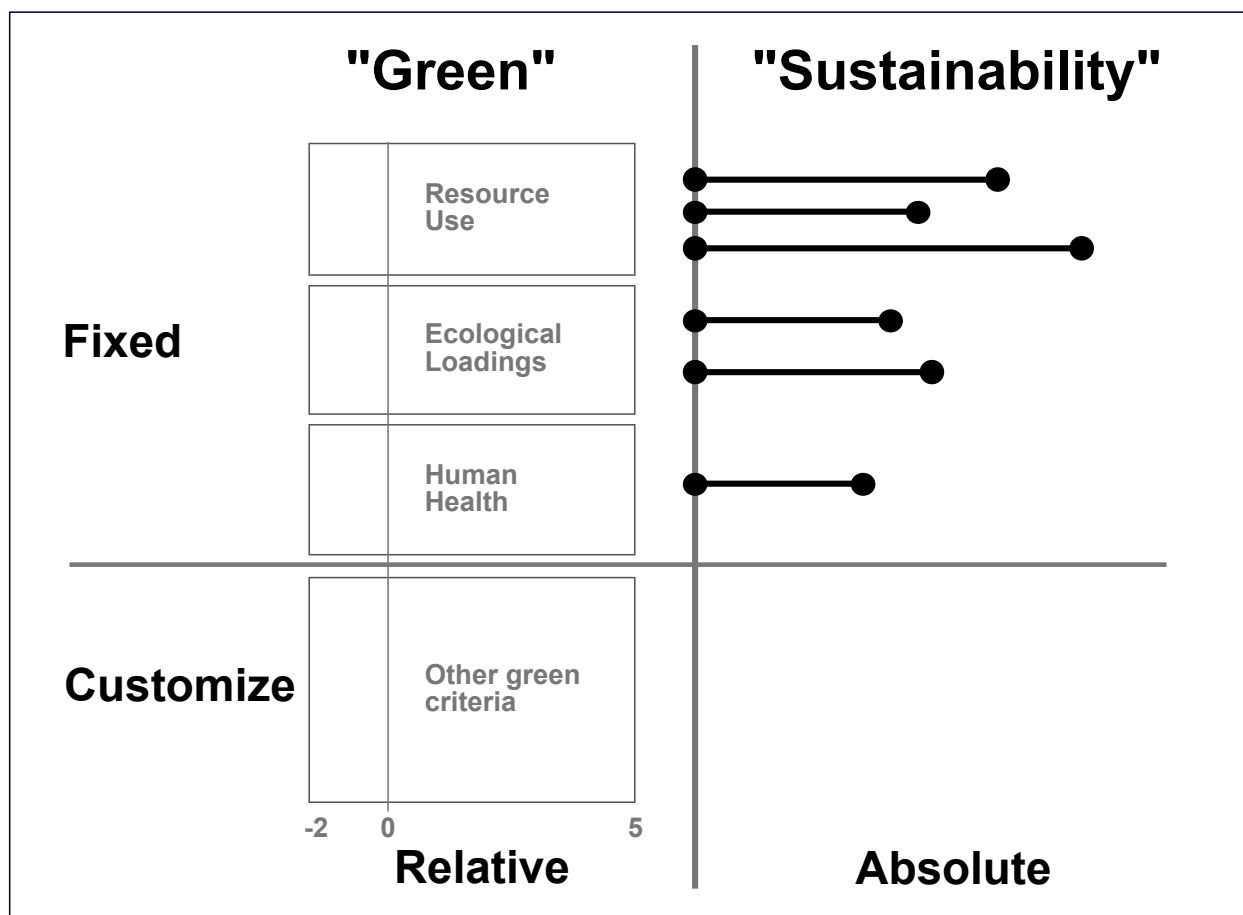


Figure 6: Reconciling “Green” and “Sustainable” agendas in environmental assessment

## 11. REFERENCES

Baldwin, R., Yates, A., Howard, N., and Rao, S., (1998) *BREEAM 98 for Offices*, Building Research Establishment, UK.

Baldwin, R., (1998) Environmental Assessment of Buildings in the UK, *Proceedings of Green Building Challenge '98* Vol. 2, pp10-16, Vancouver BC, October 26-28<sup>th</sup> 1998

Beetstra, F., (1997) Beyond LCA: Building related environmental decisions, *Proceedings of CIB TG8 Second International Conference on Buildings and the Environment*, Paris, France June 9-12, 1997, Vol. 2, pp565-573

Cole, R.J., (1998) Charting the Future: Emerging Trends in Building Environmental Assessment Methods, *Building Research and Information*, Vol. 26 (1), 3-16

- Cole, R.J., & Larsson, N.K., (1998) Preliminary Assessment of the GBC Assessment Process, *Proceedings of Green Building Challenge '98* Vol. 2, pp251-267, Vancouver BC, October 26-28<sup>th</sup> 1998
- Cole, R.J., & Mitchell, L. (1999) Customizing and Using GBTool: Two case-study projects, *Building Research and Information*, Vol 27, Number 4, pp.
- Cooper, I., (1999) Responding to challenges raised at the GBC '98 conference: moving from environmental assessment towards sustainability?, *Building Research and Information*, Vol 27, Number 4, pp.
- Kohler, N., (1999) An observer's perspective on the relevance of the Green Building Challenge, *Building Research and Information*, Vol 27, Number 4, pp.
- Lowe, R., (1996) "Energy, climate and housing" in Bell *et al*, Energy efficiency in housing. Aldershot, UK Avebury
- Lowe, R., Leeds Metropolitan University, UK., Personal communication, 27<sup>th</sup> October 1998
- Papamichael, K., and Protzen, J.P., (1993) *The Limits of Intelligence in Design*, Energy and Environment Division, Lawrence Berkeley Laboratories, University of California, Berkeley, CA, Report # LBL-31742/UC-350
- Rees, W., (1999) The Built Environment and the Ecosphere - A Global Perspective, *Building Research and Information*, Vol 27, Number 4, pp.
- Todd, J., (1998) Incorporating Life-Cycle Assessment Concepts and Other Suggestions for Enhancing the GBC 98 Assessment Framework, *Proceedings of Green Building Challenge '98* Vol. 1, pp330-336, Vancouver BC, October 26-28<sup>th</sup> 1998
- US Green Building Council, (1996) *LEED (Leadership in Energy and Environmental Design), Environmental Building Rating System Criteria*.

## **12. ACKNOWLEDGMENTS**

This paper is based in part on information and ideas from the five *International Framework Committee* meetings through the *GBC* process and the National Team reports submitted as part of the *GBC '98* conference.