Title: **Redefining the impact assessment of buildings: an uncertainty-based approach to rating codes.**

Silvio Caputo and Mark Gaterell

Abstract: Discrepancies between predicted and in-use building performance are well documented in impact assessments for buildings such as rating codes. This is a consequence of uncertainties that undermine predictions, which include procedural errors as well as users' behaviour and technological change. Debate on impact assessment for buildings predominantly focuses on operational issues and does not question the deterministic model on which assessments are based as a potential, underlying cause of ineffectiveness. This article builds on a non-deterministic urban planning theory and the principles it outlines, which can help manage uncertain factors over time. A rating code model is proposed that merges its typical steps of assessment (i.e. classification, characterisation and valuation) with those principles, applied within the impact assessment of buildings. These are *experimentation* (of other criteria than those typically appraised), *exploration* (the process of identifying the long-term vulnerability of such criteria) and *inquiry* (iterating and critically evaluating the assessment over time).

Keywords: Impact assessment of buildings; Rating codes; Uncertainty; Uncertainty-based planning.

1. Introduction.

Increasingly used worldwide (Cole and Valdebenito, 2013), rating codes are perhaps the most popular assessment to measure the impact of buildings on the environment, which in Europe – as for 2012 - account for 40% of total energy use and 36% of total CO_2 emission (Zhao and Magoulès, 2012). Rating codes were designed to predict the 'whole building' performance (Fowler and Rauch, 2006) by using clusters of indicators representing several areas of sustainability in relationship to building design and use (Chandratilake and Dias, 2013), measuring such indicators and aggregating them to express a final rating. The introduction of the European Energy Performance of Buildings Directive (2002) has further contributed to promote the use of these impact assessment tools within the building industry, and national and local governments (Schweber and Hasan Haroglu, 2014).

Being mainly voluntary, rating codes are used by developers, building owners and practitioners to demonstrate the high quality of their buildings. The ratings used in many of these tools have also become benchmarks in policies and planning frameworks, thus utilised in the decision-making process leading to planning consent (Retzlaff, 2009). However, rating codes are not without problems. Their stated aim – at least in the UK – is to facilitate a holistic approach to sustainability (BREEAM, 2014). Cole (1998) links their aim to sustainable development, hence encompassing social, environmental and economic dimensions. But rating codes struggle with the difficulty of integrating multiple aspects of sustainability within their assessment's structure, in particular social sustainability (see Mateus and Bragança, 2011; Lutzkendorf and Lorenz, 2006). Furthermore, effectiveness of rating codes is questioned (see Cole, 2005), also in the light of the increasing evidence that buildings in use do not perform as initially rated (Carbon Trust, 2012; Menezes et al., 2012; Perez-Lombardi et al, 2009), because of many uncertain factors that are not considered in the assessment process such as users' behaviour (Fabi et al, 2012), which some authors claim to be the main cause of discrepancies between predictions and real performance (Galvin and Sunikka-Blank, 2012; Zachary et al., 2010; Haas and Biermayr, 2000).

Since their introduction in 1990 rating codes have been in constant evolution, attempting to improve their predictive accuracy. BREEAM – a UK rating code – issued different versions (1998, 2006, 2011 and 2014) which have progressively improved the assessment system by, for example, specialising the appraisal depending on the building type (e.g. supermarkets, education, industrial buildings, etc.). Although significant, these improvements do not address uncertain factors mainly because, this paper argues, it would require a structural shift from the current rating code's quantitative approach, which is deterministic and leads stakeholders involved in the process of design and construction to accept predictions as real, to one that is sensitive to project-specific characteristics and open to multiple outcomes. This article proposes an outline model of rating code which learns from principles elaborated in an urban planning theory that, by recognising uncertainty as a defining feature of the present urban context, identifies principles that can help manage it (Hillier, 2011).

The paper is structured as follows: in the next section, rating codes are briefly introduced and shortcomings that have been identified in relevant literature highlighted. Subsequently, different typologies of uncertainty are reviewed in order to identify one that is typically not considered in impact assessments for buildings. Principles of the urban planning theory mentioned above are subsequently discussed in order to transpose them to the rating code field and generate a new rating code model to manage uncertainty. The article also identifies lack of debate as a reason why rating codes still preserve their deterministic approach.

2. Impact assessment for buildings: advantages and shortcomings of rating codes.

There are two main systems used to assess the environmental impact of buildings, life cycle analysis and criteria-based tools (Cheng et al. 2017; Assefa et al., 2007). The former, initially designed to assess the life cycle of products or processes (Bribián et al., 2009), measures the impact of the entire building's lifecycle within some boundaries set at the beginning of the analysis (e.g. from the extraction and processing of materials to the decommissioning of the building). The latter is a quantitative assessment, measuring the performance of criteria (i.e. indicators) for resource use, social (e.g. health and wellbeing) and ecological impact. Criteria are scored, and scores weighted and aggregated in order to generate a final rating for the whole building performance. BREEAM, the first rating code launched in 1990 by the UK-based Building Research Establishment, is a criteria-based tool assessing issues such as energy, water efficiency, waste management, and land use and ecology. BREEAM was successful, and other rating codes followed (e.g. LEED in the USA, CASBEE in Japan and DGNB in Germany), with 40 rating systems established worldwide by 2008 (Pushkar and Shaviv, 2016). All rating codes are based on the same assessment system but with different weighting and selection of criteria. Such differences are sufficient to generate differences in final results when different rating codes are used to assess a building (Wallhagen and Glaumann 2011; Wallhagen et al., 2013; Cheng et al., 2017), thus showing that – despite sharing the same system of assessment - a common methodology and theoretical approach for criteria-based assessments is missing (Wallhagen et al., 2013). Nevertheless, as mentioned above, the use of these impact assessments for buildings is increasingly popular and is now embedded in many planning procedures and policies (Retzlaff, 2009) or used by financial and insurance companies utilising them as 'a basis for risk and mortgage appraisals and real estate valuations' (Cole, 2005).

Cole (1998) defined rating codes as tools enabling 'informed decisions based on the outcome of the assessment that is most critical'. This definition portrays rating codes as tools designed to provide evidence-base for decision-making. Much of the literature on this topic focuses on effectiveness in terms of precision and reliability of results (Krizmane, 2016; Yu et al., 2015; Alyamia and Rezguib,, 2012; Menezes et al, 2012; Kajikawa, 2011; Mateus and Luís Bragança, 2011; Reijnders and van Roekel, 1999) or comparability across the different rating codes (Becchio et al., 2014; Cheng et al, 2017; Adegbile, 2013; Chew and Das, 2008; Crawley and Aho, 1999) in order to increase their effectiveness within the decision-making process. But only few studies (mentioned in the following sections) discuss fundamental shortcomings, which affect the capability of the impact assessment tool to meet its broader aim and point at the danger of relying on ratings that are merely predictive when taking decisions. What follows is a brief overview of such shortcomings.

Scope and complexity - Within a criteria-based system of assessment, sustainable performance is defined by the selection of criteria, which, in rating codes, typically privileges environmental, rather than social, factors (Conte and Monno, 2012; Fenner and Ryce, 2008). But the complexity of sustainability can hardly be captured within a set of categories/criteria (Lützkendorf and Lorenz, 2006; Berardi, 2012). Moreover, there are several interpretations of social sustainability (Dempsey et al., 2009), which is understood in different ways. Generally, rating codes refer to it as a function of health and wellbeing (e.g. ventilation, view out) (Haroglu, 2013), whereas it is suggested that it should include factors such as education and awareness of sustainability (Mateus and Braganca, 2011) or even factors related to social cohesion and participation in the design process (Amasuomo et al., 2017). Such a broader understanding of social sustainability has implications not only in terms of the assessment model (e.g. how can awareness of sustainability be measured?) but also in terms of the role of the actors, who may need to be involved, for example, in a post occupancy phase of the building life, as a means to assess the impact of the educational component of the building design and process. Other authors point at the excessively general nature of categories/criteria that sometimes fail to reflect contextual conditions, e.g. water scarcity, which may necessitate local or even building-specific modifications to the weighting system as a consequence of site-specific vulnerabilities and criticalities (Chandratilake and Dias, 2013; Alyami and Rezgui, 2012). Furthermore, by excluding or including certain criteria, technologies or design strategy, rating codes can generate imbalances in the appraisal (Retzlaff, 2009).

The need to include more refined criteria for social sustainability and other aspects of buildings' sustainable performance is a symptom of a wider problem related to the scope of the assessment. Such a scope is generally confined to the building and the building site, whereas there are externalities that should be considered in order to generate an *absolute* (Cole, 1998), rather than local, impact assessment. To this end, Conte and Monno (2012) propose a rating code that links criteria typically included in the rating code assessment to a broader impact at an urban scale, with scores assigned to building-related criteria only when these generate positive impact at an urban scale. This proposal, however, exposes the complexity of an absolute assessment: in fact, the difficulty of identifying and including a sufficient number of criteria capturing the multi-dimensional, multi-scale concept of sustainability and building construction or the attempt to measure its absolute impact poses the problem of manageability: increasing complexity may lead to higher effectiveness of the assessment but at the cost of operability (Chandratilake and Dias, 2013). It would also require a shift in the impact assessment culture (Conte and Monno, 2012; Cole, 1998) which at present sees buildings as discrete entities rather than part of a wider urban system.

Assessment and educational tool – Literature on rating codes is quite limited and rarely questions the use of the impact assessment's results within the decision-making process (Haapio and Vittaniemi, 2008). However, a few studies can be found on the capability of rating codes not only to assess but to promote and raise awareness about sustainability (Haroglu, 2013). These tools are voluntary and therefore used only for a small share of the newly built. Nevertheless, the impact they generate in the process of assessment amplifies their effectiveness since it raises awareness amongst the actors involved in the design and construction process, including practitioners, building industry and decision-makers at large (Cole, 2005). Scientific analysis alone cannot elucidate the impact of human interventions on sustainability (Krizmane, 2016; Cole, 2005). It is therefore the role and utilisation of the assessment tool within the wider process of design, implementation and use that can generate real effectiveness. To this end, the potential of rating codes to direct design choices towards sustainable building design and construction could turn it into a powerful design tool. But rating codes were not originally created as a design tool (Cole, 1998). In order to do so, the rating code should provide guidelines at an initial design stage and more accurate criteria as the design and construction progress (Thuvander et al. 2013), or a more flexible selection of sustainability criteria which does not constrict design options (Cole, 1998). Effectiveness in raising awareness is also problematic for other actors such as occupants. Cheng et al. (2017) maintain that the involvement of the building users within the design process, in order to identify their needs and goals, is necessary. Without, it will be difficult to judge which one of the energy saving concepts and measures perform well and which ones do not work at all. Moreover, it could be added, the identification and engagement of representative samples of occupants can be problematic. These reflections imply not only that the post occupancy phase, in which measurement of the real resource use can be gathered and analysed, must become an essential requirement of the assessment but also that the assessment must be conceived as a flexible tool in which criteria that have proved ineffective can be exchanged for others.

Gap - Perhaps the main shortcoming debated is the difference between predicted building performance and real operational life, which often do not match for a number of reasons both technological and behavioural (Carbon Trust, 2012; Menezes et al., 2012; Perez-Lombardi et al, 2009). Performance gaps were not evidenced only in the UK but also in studies conducted in China (Zhao and Zhou, 2017) and in LEED certified buildings worldwide (Newsham et al., 2009). The majority of these studies focus on energy consumption, comparing real usage with prediction. There is a paucity of studies on other criteria such as ecology, which is probably more difficult to measure. Nevertheless, an energy performance gap points not only at operational assessment shortcomings but also at failure to raise awareness in occupants, which is one of the aspirations of the tool. The high degree of uncertainty associated with predictions formulated further confirms that ratings generated from assessments are merely hypothetical (or aspirational) performance targets (Fenner and Ryce, 2008).

It is worth stressing that the majority of a limited literature on rating codes focuses on procedural issues. This may have limited the role that debate in literature has played in the evolution of this impact assessment. As a term of comparison, we note that literature on another model of assessment, Environmental Impact Assessment (EIA), has played an important role in its evolution. EIAs were introduced in the 1970s to assess the impact of human interventions, following the US National Environmental Policy Act (NEPA) and in response to environmental concerns that were later on captured in the definition of sustainable development (Cashmore, 2004). EIAs were subsequently introduced in the UK in the 1980s and since then, they have been evolving in response to three modifications of the European Directive 85/337/EEC, and they are likely to change in response to the latest 2014 Directive (Jha-Thakur and Fischer, 2017). One of the main issues highlighted soon after its introduction in the UK is the risk for this assessment to be used as scientific evidence on which choices can be made by decision-makers (Cashmore, 2004), which was subsequently debated in other studies (Cashmore et al, 2010; Morgan, 2012; Lobos and Partidario, 2014). The role of the assessment within the process of decision-making and the factors at play within it (i.e. political, economic, etc.) are such that this process is neither linear nor rational (Pope et al., 2013; Weston, 2000). Within such debate, the review of theories on decision making (Weston, 2000; Fischer et al., 2010) led, amongst other things, to understand the assessment as one that must be adapted to the context. Fischer et al (2010), for example, suggest that an appropriate selection of context-sensitive indicators (i.e. understood and valued by the stakeholders who will take a decision) can lead to higher effectiveness of the assessment in terms of impact on the planning decisions taken.

Another much debated issue is uncertainty, which is directly addressed in the latest EU Directive, requiring that a list of uncertainties involved in a project be included in EIA reports (Fischer et al., 2016). Uncertainty as an element impeding the effectiveness of the assessment is debated from many standpoints, including a conceptual perspective focusing on the aims of the assessment and how their correct definition impacts effectiveness (Cashmore et al, 2010), the precautionary measures that should be formulated in connection with uncertainties (Weston, 2000) and more. Jalava et al. (2013) argue that EIAs are meant to reduce risks and uncertainties of human interventions but at the same time they may not express all the uncertainties that remain unresolved with sufficient clarity. In a review of follow-up (ex-post) assessments of transport infrastructure projects in England and Norway, Nicolaisen and Driscoll (2016) too note a lack of communication of the uncertainties related to the reliability of internal and external factors of projects. In fact, a follow-up to an assessment is not only instrumental to measuring its effectiveness but also a way to learn from previous failures (Jones and Fischer, 2016), thus possibly mitigating uncertainties in subsequent projects and assessments.

As mentioned above, the richness and depth of the issues debated in this abundant stream of literature stimulate change by pointing to new directions, whereas, in comparison, literature on rating codes is not so active. In fact, the overview presented in this section shows that the impact assessment model of rating codes, in particular its deterministic, path-dependent nature, limits their potential to be effective at several levels (assessing real impact, educating, and linking the assessment of the building to the wider scope of sustainable development). The predictive character of ratings is acknowledged within the BREEAM manual (2014) and – although only optional - post-occupancy evaluation is offered as part of the assessment. Although important, such an option does not address the fact that predictions are in reality the evidence-base on which planning consent and design choices are made. We propose an uncertainty-based approach to address such limits and, in the following section, we give a brief overview of the concept of uncertainty and the way this has been defined in different fields of impact assessment.

3. Typologies of uncertainty and uncertainty management in an urban planning theory.

Uncertainty has been defined not only as the mere absence of information but also its incompleteness. New information can resolve uncertainty or generate further uncertainty at a deeper level (Walker et al, 2003). Uncertainties in predicting the environmental impact of planned interventions can refer to inaccuracy of baseline information, changes operated within the project assessed and incorrect understanding of causal effects (Tullos, 2009; Perdicoúlis and Glasson, 2006). They can also refer to collection of data (Booth and Choudhary, 2013; Garcia Sanchez et al., 2014) and users' behaviour, which are inherent to any environmental assessment process (Weston, 2000; Leung et al., 2015). A useful categorisation of uncertainties is provided by Rotmans and van Asselt (2001). They point out that there are two recurrent typologies of uncertainty which in turn characterise several common types. These are *lack of knowledge* and *variability*. The former includes inexactness and immeasurability, the latter includes human behaviour, technological surprise and societal randomness. A brief review of uncertainty according to different discipline-specific perspectives shows similar understandings of uncertainty as defined by these two typologies (see Table 1).

TABLE 1

Uncertainty associated with *lack of knowledge* is generally modelled though ever-more sophisticated mathematical and statistical methods such as Bayesian, fuzzy-rule based methods and model divergence corrections (see Ascough et al., 2008). *Variability* is arguably more difficult to quantify and is perhaps better captured through tools for qualitative assessments such as scenario analysis. Duiker and Greig, (2007) point out that scenario analysis is particularly useful for EIAs, especially for the development of risk management strategies. Scenario analysis is a systemic investigation which can be used to broaden the scope of analysis to include factors exogenous to the system considered both in space and time, which may have significant impact on performance. A case in point is given by a study documenting an assessment on a local ecological system that, by looking at the effect of climate change on the migration of species exogenous to the system, surmises the impact of such a migration on the local fauna (Duinker and Greig, 2007). Such a migration is hypothetical but plausible and, when considered as a concrete threat, can generate different strategies than those with a conventional appraisal procedure.

Examples of applications of scenario analysis to the impact assessment of buildings can also be found. For example, Hunt et al. (2012) merge a rating code (the Code for Sustainable Homes) with a scenario based exploration of domestic water efficient technologies. This leads to the identification of the technology that is likely to be more efficient under different scenarios of water consumption. Caputo et al. (2012) assess the long-term conformity to several levels of energy efficiency within the Code for Sustainable Homes of a development in Birmingham, using scenario analysis. In all these experimental studies, quantitative and qualitative assessments are not generated deterministically. Instead, *variability* is taken into account using several methods of scenario analysis (e.g. horizon scanning, scenarios and visioning) in order to identify a number of possible outcomes. Inevitably, the process is holistic and also discursive, in that it does not only offer quantifications but also reasoning, which is in turn instrumental to the identification of causes behind uncertainty and ways to address them. For rating code models, moving away from determinism would therefore entail embracing a very different approach that recognises the impossibility of reaching precise results and the advantage of working flexibly with multiple options.

Scenario analysis, however, is only a tool that can be helpful if used within a structured approach in which results from the analysis can be meaningfully utilised. It is difficult to imagine how this technique can be integrated into the path-dependent model of rating codes. In fact, a conceptualisation provided by Wallahagen (2013) depicts such a model as follows:

- Structure (hierarchical structure, components, complexity);
- Content (labels, scoring, categories, parameters);

- Aggregation (method, weighting) and Scope (functional equivalent, spatial boundaries, temporal boundaries, impacts).

Another conceptualisation that is less prescriptive and attempts to capture the underlying principles of the impact assessment model is provided by Fenner and Ryce (2008):

- *classification* (i.e the identification of inputs and categories),
- *characterisation* (i.e. definition of the contribution of each input to the assessment); and
- *valuation* (i.e. scores and rankings).

We use this conceptualisation as a stepping stone allowing to include *variability* in the assessment. To this end, we turn to a theory developed in urban planning which directly addresses *variability* in order to learn and apply the learning to rating codes.

A non-deterministic approach to urban planning to manage uncertainty.

In reaction to an approach to planning relying excessively on trends and forecasts to determine patterns of urban development, Myers and Kitsuse (2000) call for qualitative approaches integrating data analysis, which can help make sense of past events and the present, and construct a line of continuity to better anticipate future challenges. Prescriptive targets such as housing units and commercial floor space risk to be meaningless and unattained in a world with high uncertainty (see Balducci, 2011). Hillier (2011) proposes a theoretical approach to deal with 'virtualities unseen in the present' (Balducci, 2011). She introduces the concept of different 'trajectories or visions of the longer term future' as opposed to a future envisioned in continuity with the present, or as a path-dependent repetition of the past. She argues for a 'cartographic method' to develop planning, in which potentialities are traced and maps of the interplay of critical factors and phenomena are drawn up. Myers and Kitsuse (2000) reach the same conclusion when they say that scenarios have the power to demystify the future by 'reducing complexity while bringing multiple perspectives into consideration'. *Variability* as a form of uncertainty can be addressed by charting future possible events with the aim of generating a *possibility space* (see Duinker and Greig, 2007), within which options for urban development can be examined and their performance evaluated under a number of variables.

Hillier is aware of the difficulties of applying theoretical insights into practice (2005; 2011). Hillier is not alone; other scholars have developed work and provided insights on the difficulties of moving from strictly normative ways to envisage and implement urban development to new approaches focusing on process (i.e. a dynamic understanding of phenomena) (see Fainstein, 2005; Galloway and Mahayni, 1977). Nevertheless, Hillier attempts the formulation of three guiding principles that recognise the dynamic rather than static nature of urban transformation, which can have an impact on the way planning is understood in practice:

- the investigation of ‘virtualities’ unseen in the present;
- the experimentation with what may yet happen; and
- the temporary inquiry into what at a given time and place we might yet think or do.

What follows is a brief elaboration of these principles and an attempt to transpose them to the rating code field.

The first principle can be associated with a permanent exercise of horizon scanning ensuring that, when planning, what is possible is identified and not ignored. This exercise, for example, can give a voice to those urban stakeholders (e.g. local communities, associations and small enterprises) who are part of (and informally involved in) any urban transformational process, and with their actions elicit surfacing needs and wants or influence the success or failure of top-down plans. The principle can thus be seen as a call to planning intended as an exploratory practice, attentive to how bottom-up processes can steer transformation in cities in ways that are not intentionally and centrally planned. Harnessing these processes becomes a way to turn uncertainties into opportunities and can lead to a planning strategy highly adaptive to emergent phenomena and therefore endowing resilience. With regards to rating codes, it is this exploratory dimension that can be useful to transform them into effective design tools. This dimension requires systemic inquiry into the possible vulnerabilities of design options. For example, buildings designed with open spaces and to perform efficiently through natural ventilation may be, shortly after their delivery, renovated with cellular spaces, thus compromising their passive cooling strategy (Montazami et al, 2015). Passive design principles are currently strongly promoted, although it is unsure whether they will perform effectively against a medium-to-long term scenario of higher mean temperatures (Sameni et al, 2015). Exploration, in other words, can also help identify technical solutions and connected criteria that are appropriate for particular contexts, which is another shortcoming of rating codes highlighted above.

The second principle suggests experimentation as an approach to ascertain benefits and advantages of emerging trends in urban transformation. Herein, the eventualities are not only perceived as adverse events to be managed but also as occasions to test new arrangements and take advantage of their positive aspects. In planning, this entails a shift of attitude to governance allowing emergent phenomena to influence the planning agenda and be tested for their effectiveness in addressing societal issues. Eventualities are place-specific and experimentations are thus responses to specificities of local conditions. This can be linked to another characteristic of rating codes, which offer a generalised, universal set of requirements for compliance, thus leaving no space for options that are not included within the rating frameworks or for any other alternative that departs from an understanding of sustainable building performance and its scope as defined within such frameworks.

The last principle promotes a permanent attitude to inquiry and reflection on the state of things at any time. It suggests critical and self-critical analysis as an approach to verify the effectiveness of directions undertaken and also preparedness to change when analysis points to the need for different directions. It is a principle that brings together the first two, recognising that exploration and experimentation necessitate critical reflection to evaluate effectiveness of all options. This requires openness to change and flexibility in decision-making for urban development. By extension, it can be an invitation to understand rating codes differently, not only as a quantitative and/or qualitative evaluation of buildings' performance but also as instruments enabling inquiry, therefore dialectical exchange between stakeholders, leading to awareness of substantive objectives for sustainable performance and solutions that are robust over time.

4. An outline of an uncertainty-based approach to rating codes for buildings.

In the sections above, shortcomings of the rating codes have been outlined together with principles of a non-deterministic planning theory, suitable to deal with *variability*. Factors of uncertainty for rating codes such as limited scope of the assessment, educational impact and gap between predicted and in-use performance, which limit their effectiveness can be revisited using the concepts of exploration, investigation and inquiry. We bring together these insights and propose a new model of rating codes, starting from the conceptualisation of Fenner and Ryce (2008) introduced above. A diagram of a new rating code merging the two is represented in Figure 1.

FIGURE 1

In the diagram, the stages of *classification and characterisation*, which are currently fixed components in all rating codes, are complemented with an *experimentation* stage, in which new technologies or strategies that are not captured in the existing *classification* and *characterisation* stages can be identified and proposed. For example, a study shows how, in some of the most common rating codes (e.g. LEED, BREEAM and GBRT), passive design features are penalised if compared to conventional energy saving strategies (Chen et al., 2015). In an amended rating code model it would be possible to propose and include passive solar design criteria under the *energy* category, thus superseding some of the existing criteria for energy efficiency. Different weighting and scores can be proposed to encourage higher efficiency in water usage, renewable energy generation or ecology, in response to particular contextual conditions and stresses. Other categories could be introduced, focusing on, for example, users' behaviour, household waste and food production, whenever relevant to the particular site, ambition of the development proposed and social profile of the users. To this end, a site and building specific investigation must be developed, which can lead to the identification of alternative strategies to sustainable performance that are more likely to be successful in the long term, within a particular socio-economic and environmental context. Furthermore, the identification of optimal strategies that need to be captured with appropriate criteria within the rating system requires dialogue with planning departments, thus encouraging dialectic debate and active participation in shaping the assessment.

In the *exploration* stage, a scenario analysis can be developed, in which the lifetime of the proposed building is specified and vulnerable factors that may undermine buildings' performance are identified. For example, as mentioned above, ventilation strategies can be impacted by changes in layout over the lifetime of buildings (Montazami et al., 2015). The perceived economic value of office buildings can be strictly related to its flexibility of spaces and systems upgrading (Vimpari and Junnila, 2016). Similar to the aforementioned need for EIA to make internal and external factors of uncertainty explicit within the EIA assessment, rating codes too can increase their effectiveness by eliciting uncertainties and use this process to generate solutions mitigating future risks. A way to implement this in practice implies the use of scenario-based techniques that can lead to broaden the scope of assessment and elicit relationships between actors, policies and diverse factors (e.g. 'what ifs' inquiring consequences of changes of use, layout, external conditions, number and profile of users, etc.), which cannot be captured in checklists for sustainable performance (Hacking and Guthrie, 2008). At its most basic, this type of quantitative evaluation could take the format of a risk analysis such as those required for large development or infrastructural projects. Other frameworks for this stage of the assessment that can be used are however available and in use. For example, BREEAM Renovation, organises the lifecycle of buildings in sub-cycles such as structural, systems and components, each one with a particular life cycle (e.g. 60 years for the structural cycle). A similar framework could be used to identify points of vulnerability across each cycle and demonstrate that such points have been addressed within the project.

In the final stage, *valuation* must be formulated that can capture both the performance forecasted, and vulnerabilities possibly undermining such performance and connected causes. For example, quantifications can be expressed with performance ranges, rather than discrete figures, and qualitative evaluations explaining the reasons for each particular performance within the range. Valuation should not be limited to the building as modelled during the design stage but extended to the in-use performance. Hillier envisions planning as a practice in which 'outcomes are volatile; where problems are not 'solved' once and for all but are rather constantly recast, reformulated in new perspectives' (Hillier, 2005:278). This is a dynamic vision of urban planning that suggests, by extension, an assessment iterated over time, following a reflective phase in which solutions are revisited and lessons are learned. Stakeholders involved in the design, construction and use of a building are therefore participating in a long-term design and monitoring process of the building, learning form this process and applying lessons to periodically improve performance. Conceptually, this principle seems distant from the linearity of the rating code model of classification-characterisation-valuation. Here again, the parallel with the EIAs debate mentioned above regarding the advantages of a follow-up assessment, can offer a useful term of comparison. Extending the timescale of the assessment can be functional both to establishing the level of exactitude of predictions and using this knowledge to improve future assessments, and to modifying, whenever technically and economically viable, anything that does not function as predicted. To this end, Soft Landings (www.bsria.co.uk/services/design/soft-landings) offers a framework which could be valid also for a new type of uncertainty-based assessment. A protocol rather than a conventional appraisal, Soft Landings expands the temporal limits of the assessment to the post-occupancy phase, at the same time modifying relationships and obligations of the actors involved in the building process (i.e. clients, designers and constructors collaborating beyond completion to ensure the correct use of the building). This, in turn, requires the redefinition of stakeholders' remits and responsibilities (within the design, construction and management process), which can no longer be limited to the delivery of buildings but also include their maintenance.

A further reflection is necessary about the issue of effectiveness. In reviewing literature on effectiveness and EIAs, Chanchitpricha and Bond (2013) identify four categories contributing to its conceptualisation: procedural (i.e. complying with standards and principles), substantive (i.e. attaining intended objectives), transactive (i.e. cost and time effective) and normative. In particular, normative effectiveness (that is: the potential of assessments to influence positively attitudes towards sustainable development of stakeholders involved in any development process), suggests a role for impact assessments that transcends the mere provision of scientific evidence and somehow stimulate a process of change. Transferring this to rating codes entails that these tools can be used to (and designed in a way that) help embed sustainability in urban policies. However, such a normative change risks to be static because of the rating codes' path dependent model, which reduce sustainable performance to a number of possible options universally applied and considers performance as predicted rather than in use. A normative change that is more dynamic can only be achieved through progressive learning and models of assessment are needed that can facilitate this process. The uncertainty-based model of assessment for buildings proposed here is an initial attempt to emphasise the potential for dynamic normative change.

5. Conclusions.

As a contribution of this paper to the debate on rating codes for buildings, a new model based on uncertainty has been outlined in the section above. The new rating code model requires a shift of focus from an effectiveness understood as reliability and robustness of the assessment results to one that is based on an identification of a *possibility space,* in which buildings can be examined during their lifetime, vulnerabilities impacting predicted performance values identified and fluctuations of such values determined, thus making uncertainties explicit. The resulting model is an evolution of the three-stage model that typically characterises rating codes (i.e. *categorisation*, *classification* and *valuation*), which are reformulated in accordance to the principles of *experimentation* (of other options of sustainable performance that transcend those typically appraised in rating codes) and *exploration* (the process of identifying the long-term vulnerability of such options), thus enabling to address *variability* (i.e. uncertainty related to randomness of nature, human behaviour and technological surprises). *Inquiry* is also used to ensure that the resulting assessment is iterated over time, with strategies initially formulated adjusted if needed. *Variability* is addressed in three ways: firstly by identifying approaches that are in line with site-specific conditions (with site boundaries that can vary from local to city-wide depending on the ambition and nature of the project); secondly, by ensuring that such approaches are implemented effectively over the life-cycle of the building; and thirdly, by providing a form of scoring that encourages this exploration. This, in turn, can improve effectiveness of the building's impact assessment by addressing issues of scope, educational impact and performance gap that are indicated in literature as ineffectively dealt with in the current rating code model.

References.

Adegbile, M. B. O. 2013. Assessment and Adaptation of an Appropriate Green Building Rating System for Nigeria. *Journal of Environment and Earth Science* 3(1): 1-10.

Alyami, G. S. H. and Rezgui, Y. 2012. Sustainable building assessment tool development approach. *Sustainable Cities and Society* 5: 52–62.

Amasuomo, T. T., Atanda, J. and Baird, G. 2017. Development of a building performance assessment and design tool for residential buildings in Nigeria. Procedia Engineering 180: 221 – 230.

Ascough II, J.C., Maier, H.R. Ravalico, J.K. and Strudley, M.W. 2008. Future research challenges for incorporation of uncertainty in environmental and ecological decision-making. *Ecological Modelling* 219: 383–399.

Assefa, G., Glaumann, M., Malmqvist, T., Kindembe, B., Hult, M., Myhr, U. and Eriksson, O. 2007. Environmental assessment of building properties—Where natural and social sciences meet: The case of EcoEffect. *Building and Environment* 42: 1458–1464.

Balducci, A. 2011. Strategic planning as exploration. *Town Planning Review*, 82(5): 529-546.
Becchio, C., Corgnati, S. P., Fabrizio, E., Monetti, V. and Seguro, F. 2014. Application of the LEED PRM to an Italian existing building. *Energy Procedia* 62: 141 – 149.

Berardi, U. 2012. Sustainability Assessment in the Construction Sector: Rating Systems and Rated Buildings. *Sustainable Development* 20(6): 411–424.

Booth, A.T. and Choudhary, R. 2013. Decision making under uncertainty in the retrofit analysis of the UK housing stock: Implications for the Green Deal. *Energy and Buildings* 64: 292–308.

Bribián, I. Z., Usón, A. A, and Scarpellini, S. 2009. Life cycle assessment in buildings: State-of-the-art and simplified LCA methodology as a complement for building certification. *Building and Environment* 44: 2510–2520.

Caputo, S., Caserio, M., Coles, R., Jancovic, L. and Gaterell, M. R. 2012. A scenario-based analysis of building energy performance. Proceedings of the ICE - Engineering Sustainability 165 (1): 69-80.

Carbon Trust 2012. Closing the gap – Lessons learned on realising the potential of low carbon building design. The Carbon Trust.

Cashmore, M. 2004. The role of science in environmental impact assessment: process and procedure versus purpose in the development of theory. *Environmental Impact Assessment Review 24:* 403–426.

Cashmore, M., Richardson, T., Hilding-Ryedvik, T. and Emmelin, L. 2010. Evaluating the effectiveness of impact assessment instruments: Theorising the nature and implications of their political constitution. *Environmental Impact Assessment Review* 30: 371–379.

Chanchitpricha, C. and Bond, A. 2013. Conceptualising the effectiveness of impact assessment processes. *Environmental Impact Assessment Review* 43: 65–72.

Chandratilake, S.R. and Dias W.P.S. 2013. Sustainability rating systems for buildings: Comparisons and Correlations. *Energy* 59: 22-28.

Chen, X., Yang, H. and Lu. L. (2015) A comprehensive review on passive design approaches in green building rating tools. *Renewable and Sustainable Energy Reviews* 50: 1425–1436.

Cheng, W., Behzadzodagar and Feifesun 2017. Comparative analysis of environmental performance of an office building using BREEAM and GBL. *International Journal of Sustainable Development and Planning* 12(3) 528–540.

Chew, M. Y. L. and Das, S. 2008. Building Grading Systems: A Review of the State-of-the-Art. *Architectural Science Review* 51(1): .3-13

Cole, R. J. 1998. Emerging trends in building environmental assessment methods. *Building Research & Information* 26(1): 3–16.

Cole, R. J. 2005. Building environmental assessment methods: redefining intentions and roles. *Building Research & Information* 35(5): 455–467.

Cole, R. J. and Valdebenito, M. J. 2013. The importation of building environmental certification systems: international usages of BREEAM and LEED. *Building Research & Information* 41(6): 662-676.

Conte, E. and Valeria Monno, V. 2012. Beyond the buildingcentric approach: A vision for an integrated evaluation of sustainable buildings. *Environmental Impact Assessment Review* 34: 31–40.

Crawley, D. and and Aho, I. 1999. Building environmental assessment methods: applications and development trends. *Building Research & Information* 27(4/5): 300–308.

Dempsey, N., Bramley, G., Power, S. and Brown, C. 2009. The Social Dimension of Sustainable Development: Defining Urban Social Sustainability. *Sustainable Development* 19(5):289-300.

Directive 2002/91/EC of the European Parliament and of the Council of 16 December 2002 on the energy performance of buildings.

Duinker, P. N. and Greig, L. A. 2007. Scenario analysis in environmental impact assessment: Improving explorations of the future. *Environmental Impact Assessment Review* 27: 206–219.

Fabi, V., Andersen, R. V., Corgnati, S. and Olesen, B. W. 2012. Occupants' window opening behaviour: A literature review of factors influencing occupant behaviour and models. *Building and Environment* 58: 188-198.

Fainstein, S. 2005. Planning Theory and the City. *Journal of Planning Education and Research* 25: 121-130.

Fenner, R. A. and Ryce, T. 2008. A comparative analysis of two building rating systems. Part 1: evaluation. *Proceedings of the Institution of Civil Engineers Engineering Sustainability* 161(1): 55–63.

Fowler, K. M. and Rauch, E. M. 2006. Sustainable Buildings Rating Systems – Summary. Pacific Northwest National Laboratory. Available at http://www.pnl.gov/main/publications/external/technical_reports/PNNL-15858.pdf. Accessed on 05.10.2017.

Galloway, T. D. and Mahayni, R. G. 1977. Planning Theory in Retrospect: The process of Paradigm Change. *Journal of the American Planning Association* 43(1): 62-71.

Galvin, R. and Sunikka-Blank, M. 2012. Economic viability in thermal retrofit policies: Learning from ten years of experience in Germany. *Building and Environment* 58: 188-198.

Garcia Sanchez, D.. Lacarrière, B., Musy, M. and Bourges, B. 2014. Application of sensitivity analysis in building energy simulations: combining first- and second-order elementary effects methods. *Energy and Buildings* 68: 741–750.

Gill, Z. M., Tierney, M. J., Pegg, I. M. and Allan, N. 2010. Low energy dwellings: the contribution of behaviours to actual performance, Building Research & Information, 38(5): 491-508.

Haapio, A. and Viitaniemi, P. 2008. A critical review of building environmental assessment tools. *Environmental Impact Assessment Review* 28: 469–482.

Haas, R. and Biermayr, P., 2000. The rebound effect for space heating Empirical evidence from Austria. *Energy policy* 28(6-7):403-410.

Hacking, T and Guthrie, P. 2008. A framework for clarifying the meaning of Triple Bottom-Line, Integrated, and Sustainability Assessment. *Environmental Impact Assessment Review* 28: 73–89.

Haroglu, H. 2013. The impact of Breeam on the design of buildings. *Proceedings of the Institution of Civil Engineers: Engineering Sustainability* 166(1): 11-19.

Hillier, J. 2005. Straddling the post-structuralist abyss: between transcendence and immanence? *Planning Theory* 4(3): 271-299.

Hillier, J. 2011. Strategic navigation across multiple planes -Towards a Deleuzean-inspired methodology for strategic spatial planning. *Town Planning Review*, 82 (5): 503-527.

Hopfe, C.J. and Hensen, J.L.M. 2011. Uncertainty analysis in building performance simulation for design support. *Energy and Buildings* 43 (10): 2798-2805.

Hunt, D.V.L., Lombardi, Farmani, R., Jefferson, I., D.R., Memon, F.A., Butler, D., and Rogers, C.D.F. 2012. Urban Futures and the code for sustainable homes. *Proceedings of the Institution of Civil Engineers – Engineering Sustainability* 165(1): 37-58.

Jha-Thakur, U. and Fischer, T. B. (2016) 25 years of the UK EIA System: Strengths, weaknesses, opportunities and threats. *Environmental Impact Assessment Review* 61: 19–26.

Jones, R. and Fischer, T. B. (2016) EIA Follow-Up in the UK — A 2015 Update. *Journal of Environmental Assessment Policy and Management* 18(1): 1650006.

Kajikawa, Y., Inoue, T. and Goh, T. N. 2011. Analysis of building environment assessment frameworks and their implications for sustainability indicators. *Sustainability Science* 6: 233–246.

Krizmane, M., Slihte, S. and Borodinecs, A. 2016. Key criteria across existing sustainable building rating tools. *Energy Procedia* 96: 94 – 99.

Leung, W., Noble, B., Gunn, J. and Jaeger, J. A. G. 2015. A review of uncertainty research in impact assessment. *Environmental Impact Assessment Review* 50: 116–123.

Lützkendorf, T. and Lorenz, D. P. 2006. Using an integrated performance approach in building assessment tools. *Building Research & Information* 34(4): 334-356.

Mateus, R. and Bragança, L. 2011. Sustainability assessment and rating of buildings: Developing the methodology SBToolPT-H. *Building and Environment* 46: 1962-1971.

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

Mirakyan, A. and De Guio, R, 2015. Modelling and uncertainties in integrated energy planning. *Renewable and Sustainable Energy Reviews* 46: 62–69.

Montazami, A., Gaterell, M. and Nicol, F. 2015. A comprehensive review of environmental design in UK schools: History, conflicts and solutions. *Renewable and Sustainable Energy Reviews* 46: 249–264.

Myers, D. and Kitsuse, A. 2000. Constructing the Future in Planning: A Survey of Theories and Tools. *Journal of Planning Education and Research* 19(3): 221-231.

Newsham et al., 2009 Newsham GR, Mancini S, Birt BJ (2009) Do LEED-certified buildings save energy? Yes, but. Energy and Buildings 41:897–905.

Nicolaisen, M. S., & Driscoll, P. A. (2016). An international review of ex-post project evaluation schemes in the transport sector. *Journal of Environmental Assessment Policy and Management*, *18*(01): 1650008.

Pérez-Lombard, L., Ortoiz, J., Gonzáles, R. and Maestre, I. R. 2009. A review of benchmarking, rating and labelling concepts within the framework of building energy certification schemes. *Energy and Buildings* 41, pp.272-278.

Perdicoúlis, A. and Glasson, J. 2009. The causality premise of EIA in practice. *Impact Assessment and Project Appraisal* 27(3): 247-250.

Pushkar, S. and Shaviv, E. 2016. Using shearing layer concept to evaluate green rating systems. *Architectural Science Review* 59(2) 114-125.

Ragas, A. M. J., Huijbregts, M. A. J., Henning-de Jong, I. and Leuven, R. S. 2009. Uncertainty in Environmental Risk Assessment: Implications for Risk-Based Management of River Basins. *Integrated Environmental Assessment and Management* 5(1): 27 – 37.

Regan, H. M., Colyvan, M. and Borgman, M. A. 2002. A taxonomy and treatment of uncertainty for ecology and conservation biology. *Ecological Applications* 12(2): 618–628.

Reijnders, L. and van Roekel, A. 1999. Comprehensiveness and adequacy of tools for the environmental improvement of buildings. *Journal of Cleaner Production* 7: 221–225.

Retzlaff, R. 2009. Green Buildings and Building Assessment Systems: A New Area of Interest for Planners. *Journal of Planning Literature* 2009 24(1): 3-21.

Rotmans, J. and van Asselt, M. B. A. 2001. Uncertainty management in integrated assessment modelling: towards a pluralistic approach. *Environmental Monitoring and Assessment* 69: 101–130.

Sameni, S. M. T., Gaterell, M., Montazami, A. and Ahmed, A. 2015. Overheating investigation in UK social housing flats built to the Passivhaus standard. *Building and Environment* 92: 222-235.

Schweber, L. and Hasan Haroglu, H. 2014. Comparing the fit between BREEAM assessment and design processes. *Building Research & Information* 42 (3), pp. 300–317.

Thuvander, L., Femenías, P., Mjörnell, K. 2 and Meiling, P. 2012. Unveiling the Process of Sustainable Renovation. *Sustainability* (4): 1188-1213.

Tullos, D. 2009. Assessing the influence of environmental impact assessments on science and policy: An analysis of the Three Gorges Project. *Journal of Environmental Management* 90: 208–223.

Vimpari, J. and Junnila, S. 2016. Theory of valuing building life-cycle investments. *Building Research & Information* 44(4): 345-357.

Wallhagen, M. and Glaumann, M. 2011. Design consequences of di¡erences in building assessment tools: a case study. Building Research & Information 39(1): 16–33.

Wallhagen, M., Glaumann, M., Eriksson, O. and Westerberg, U. 2013. Framework for Detailed Comparison of Building Environmental Assessment Tools. *Buildings* 3: 39-60.

Walker, W. E., Harremöes, P., Rotmans, J., van der Sluijs, J. P., van Asselt, M. B. A., Janssen, P. and Krayer von Krauss, M. P. 2003. Defining Uncertainty A Conceptual Basis for Uncertainty Management in Model-Based Decision Support. *Integrated Assessment* 4(1): 5–17.

Weston, J. 2000. EIA, Decision-making Theory and Screening and Scoping in UK Practice. *Journal of Environmental Planning and Management* 43(2): 185-203.

Yu, W., Li, B., Yang, X and Wang, Q. 2015. A development of a rating method and weighting system for green store buildings in China. *Renewable Energy* 73: 123-129.

Zhao, H. and and Magoulès, F. 2012. A review on the prediction of building energy consumption. *Renewable and Sustainable Energy Reviews* 16: 3586– 3592.

Zhao, L. and Zhou, Z. 2017. Developing a Rating System for Building Energy Efficiency Based on In Situ Measurement in China. *Sustainability* 9: 208.

Table 1. Categories of uncertainties (right hand side column) identified in literature

Authors	**Categories of uncertainty**
Walker et al, 2003	• Level of uncertainty (statistical; scenario and ignorance, total ignorance) • Nature of uncertainty (epistemic and variability)
Rotmans and van Asselt, 2001	• Variability (randomness of nature, human behaviour and technological surprises) • Lack of knowledge (lack of measurements, conflicting evidence and ignorance)
Leung et al, 2015	• Incomplete information, and the prediction and management of those outcomes • Communication (decision-making under uncertain conditions) • Avoidance (behaviour).
Hopfe and Hensen, 2011	• Physical (materials properties), • Design (geometry), • Scenario uncertainties (internal gains and climate change).
Mirakyan and DeGuio, 2015	• Linguistic (vagueness and ambiguity) • Knowledge (context; model and technical) • Variability (natural; human; institutional and technological) • Decision (objectives; criteria and strategies) • Procedural (available time, resources and imperfect communication)
Regan et al, 2002	• Epistemic (imperfect measurement devices, insufficient data, extrapolations and interpolations, and variability over time or space.) • Linguistic (scientific vocabulary or theoretical indeterminacies.)
Ragas et al., 2009	• problem definition uncertainty • true uncertainty (lack of knowledge) • variability (phenomenon of the real world)

Figure 1. The three stages of the rating code model (Fenner and Ryce, 2008) are represented in black. Intermediate stages – mediated from the uncertainty-based planning principles – are added in order to form an uncertainty-based model of assessment.