Closing the energy performance gap in zero carbon homes - pro-active identification, prioritisation and mitigation of causes using FMEA.

This item was submitted to Loughborough University's Institutional Repository by the/an author.

Citation: HOBART, C. ...et al., 2015. Closing the energy performance gap in zero carbon homes - pro-active identification, prioritisation and mitigation of causes using FMEA. In: Raiden, A. and Abogya-Nimo, E. (Eds.), Proceedings 31st Annual ARCOM Conference, 7-9 September 2015, Lincoln, UK. Association of Researchers in Construction Management, pp. 307-316.

Additional Information:

- This paper is presented with kind permission of Arcom, the definitive version can be found at: http://www.arcom.ac.uk/-docs/proceedings/3f85e692df147656ee6e2b09f2d83ded1.pdf

Metadata Record: https://dspace.lboro.ac.uk/2134/20001

Version: Published

Publisher: © Arcom

Rights: This work is made available according to the conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence. Full details of this licence are available at: https://creativecommons.org/licenses/by-nc-nd/4.0/

Please cite the published version.
The energy consumption of new homes in the UK routinely exceeds design predictions, in some cases by more than 100%. Work to date suggests multiple causes for this so-called “performance gap”, ranging from technical/design issues through to procurement and behavioural influences. These varied causes are often difficult to detect and may be viewed as trivial or inevitable by the parties responsible for them. Addressing these issues not only requires concurrent technical and organisational solutions, but also a means of predicting which issues are likely to be significant for a given project. In the manufacturing industry this scenario is often addressed using a methodology called Failure Mode and Effects Analysis (FMEA). Although some building component manufacturers make use of FMEA, there is little documented evidence of this technique being applied at the whole building level. In this paper we argue that FMEA is potentially well suited to addressing the energy performance gap for dwellings, but that the approach must be carefully tailored to achieve this task. The Bicester Ecotown Process Improvement Toolkit (BEPIT) research project provides a unique opportunity to develop and test this technique, by means of an iterative application of FMEA on a development of 393 true zero carbon homes. The first iteration is described in terms of both the methodological development, and the performance effect produced at project level. This learning in turn informs a discussion of the wider potential for the use of FMEA to close the energy performance gap. It is argued that the method and approach might be applicable to other building types where similar performance concerns exist.

Keywords: energy performance, failure mode effect analysis, FMEA, energy performance, performance gap, zero carbon homes.

INTRODUCTION: THE ENERGY PERFORMANCE GAP

The built environment is widely accepted as being both a major contributor to global carbon emissions, and an arena in which significant improvements are possible (Urge-Vorsatz and Novicova 2008, Oreszczyn and Lowe 2010). A particular area in which emission reductions have been targeted is through improved energy efficiency for new dwellings, and specifically the promotion of an agenda focused on achieving zero carbon homes (Pan and Ning 2015). In the UK incremental reductions in carbon emissions have been sought using design standards applied through the Building Regulations since 2006, and in 2016 a zero carbon standard is due, requiring all new homes to achieve net zero carbon emissions in respect of energy use for heating, hot water, lighting and ventilation. Whilst this standard is expected to include scope to offset a proportion of emissions using off-site “allowable solutions”, the existing minimum standards for building energy design performance are also expected to be
retained and tightened (Zero Carbon Hub 2013; Department for Communities and Local Government 2013).

With demand for new homes in the UK remaining significant despite the recent economic downturn, this policy has considerable potential to assist in managing future carbon emissions. To be effective however it is crucial that design standards translate effectively to performance in use, and a growing body of research suggests that this may not be the case (Wingfield et al 2008, Bell et al 2010, Gorse et al 2012, Pan and Garmston 2012, Levinson 2015, Pan and Ning 2015). As such, closing this "performance gap" is essential to delivering the carbon management strategy described above, however despite growing academic understanding of its causes, limited progress appears to have been made in improving performance outcomes in mainstream housebuilding. It would seem that a wholly technical focus on carbon reduction is unlikely to address performance concerns, and efforts must be made to better understand the full range of failure modes that can occur within the socio-technical system that comprises zero carbon building (Pan and Ning 2015). Crucially whilst research data has typically been obtained through forensic examination of build processes and post completion testing, these techniques may have rather limited application in a commercial context. In a project focused industry such as construction, social pressures dictate that improvements be considered alongside their time and cost implications. To achieve mainstream engagement with the performance gap, a set of tools are therefore required which will allow organisations to predict performance issues in real time on a project by project basis, introduce appropriate mitigation measures, and evaluate their effectiveness in flight.

**Performance gap components**

The availability of quantitative data to validate design predictions relating to the energy performance of buildings is currently somewhat limited, perhaps due a combination of the relative immaturity of available test methods, limited statutory testing requirements, and commercial and/or reputational sensitivities. Currently utilised as-built test methods include whole house measurement of energy use for unoccupied buildings by means of a co-heating test, along with elemental testing of the thermal envelope by means of heat flux testing and thermal imaging (Wingfield et al 2008, Bell et al 2010, Gaze 2010). Whole house as-built airtightness testing is additionally widely used in the UK, however its application to validation of design energy use predictions is somewhat limited, with the relationship between air permeability under depressurisation and its effect on energy use under atmospheric conditions being difficult to calculate and highly context specific (Sherman and Wanyu 2006).

Information regarding the performance gap therefore currently exists largely in case study format, with evaluation typically based upon various combinations of design and modelling evaluation, site observation, pre-completion testing and post occupancy evaluation. These case studies consistently demonstrate significant shortfalls in fabric energy performance and the energy efficiency of building systems, both pre and post occupancy. Numerous and multiple causes are proposed for these shortfalls, with each project citing a specific suite of problems based upon the organisational context and technical design features (Wingfield et al 2008, Bell et al 2010, Gaze 2010). Notwithstanding this, some clear cross cutting themes can be identified, and these range across a full range of construction activity, and have implications for an equally
wide range of stakeholders (Gorse et al 2012, Zero Carbon Hub 2014, Pan and Ning 2015). Potential problems begin at initial design and modelling stage, with issues being identified in relation to maintaining the continuity and clarity of design aspiration, and to the appropriate and competent use of modelling to support this. The assumptions adopted at this stage may then face both commercial pressure and lack of technical understanding during detailed design and procurement, resulting in potential for significant unrecognised changes to the design intent. At construction stage scope exists for further uncontrolled design development and product substitution, as well as ad-hoc development of construction processes, and variations in standards of workmanship, commissioning and installed product performance. Ultimately, occupation brings further potential for misunderstanding or lack of engagement with systems, whilst the relative novelty of the test methodologies themselves means that testing error may itself be difficult to quantify.

FMEA

In the manufacturing industry, holistic analysis of systems is commonly conducted using a Failure Mode and Effect Analysis (FMEA). The purpose of FMEA is to identify, prioritise and mitigate potential "failure modes" before they reach the end user. A failure mode is simply a description of the manner in which a failure occurs, and whilst it may be directly linked to a root cause, it may also exist as part of a chain of related failure modes (Mecca and Massera 1999, Hage 2002, Stamatis 2003, McDermott et al 2009, Bahrami et al 2013). FMEA typically considers failure modes at system, component and part level, and may be conducted at concept, design and/or process stage. The FMEA process consists of three essential elements:

- Identification of potential failure modes, ideally using a whole team approach.
- Prioritisation of failure modes, usually by means of a risk priority number (RPN) generated by ranking failure modes in terms of severity, likelihood of occurrence, and likelihood of detection.
- Mitigation of failure modes by preventing them or controlling their effects.

FMEA is perhaps best viewed as a methodology for enacting these elements, rather than a self-contained tool. It provides a framework to promote rigorous prioritisation and tracking of potential problems, however it is crucial that appropriate additional tools are also employed for identification and mitigation of failures (Mecca and Massera 1999, Stamatis 2003, McDermott et al 2009).

FMEA was initially used in military and space exploration contexts and has subsequently become commonplace in other sectors including civil aviation, automotive, nuclear energy and electronics (Liu et al 2012). Whilst not widespread, the use of FMEA in buildings has also been previously explored to some extent, and has been documented in applications including ensuring the durability of building components, reducing risk in cladding design and installation and improving construction safety (Hage 2002, Talon et al 2006, Song et al 2007). The use of FMEA has additionally been documented in connection with retrospective evaluation of the performance gap, specifically as it relates to building services installations (Tuohy 2013, 2014). Despite the emerging popularity of FMEA as an evaluative method, its use at the whole building level in a predictive role, as would be needed to address the performance gap, is yet to be tested empirically. As noted by Mecca and Massera (1999) this pre-emptive project based application is likely to present considerable
technical and organisational challenges, particularly with regards to contractual supply chain arrangements, and the site production environment. In this paper we report on research that is specifically identifying and addressing these challenges.

**METHODOLOGY**

The Bicester Ecotown Process Improvement Toolkit (BEPIT) project aims to develop tools to assist organisations in effectively implementing the next generation of low carbon housing. The project, funded by Innovate UK, aims to derive such tools from detailed observation and analysis of the North West Bicester project, the UK's largest zero-carbon housing scheme. The initial "Exemplar" project of 393 new homes is now underway, with the first phase of 91 units currently under construction. Zero carbon for these homes is defined as including unregulated energy use and without the use of off-site offsetting, and a total of approximately 6000 units are planned on the site. The BEPIT team consists of the developer, project manager, architect, main contractor, energy services consultant, sustainability consultant and academic partner. The project therefore includes a substantial element of industry engagement, with team members both providing expertise and data to the research project, and feeding findings back to their respective organisations. The wider aim of the research is to disseminate learning relating to closing the performance gap in the form of a publicly accessible toolkit.

FMEA was identified as a potentially useful tool at an early stage of the research project, and selected for testing in an action research context. Identification of failure modes began at the design stage of Phase 1, when representatives of the BEPIT team, with the addition of the timber frame subcontractor, were asked to suggest "examples of potential problem areas", based upon their personal experience and expectations for the project. Responses were collected by means of a questionnaire which identified 18 particular categories for these problems, and which also listed 22 examples previously identified by literature review. These categories and examples were provided for guidance of participants, the majority of whom had no previous experience of using FMEA. The problem areas identified by this exercise were then analysed by a researcher and used to populate the "FMEA Schedule" with a total of 328 distinct failure modes. It should be noted however that no contribution was received from representatives of the main contractor or architect. To facilitate the next stage of prioritisation, the researcher rated each failure mode in terms of likelihood of Occurrence (O), Severity (S) and likelihood of Detection (D) using standard 1-10 FMEA rating scales sourced from the manufacturing industry. Finally, in accordance with standard practice, these scores were multiplied together to generate the overall RPN's which allow the failure modes to be prioritised for action.

To facilitate mitigation of failure modes it was decided to raise them at the main contractor's design team meetings. These meetings were the focus of the detailed design and procurement process, and appeared to offer an ideal opportunity to engage with the delivery team to pro-actively consider and address the potential causes of the performance gap. Due to the large number of failure modes identified it was felt to be impractical to attempt to mitigate them all, and indeed a core aim of FMEA is to target resources towards addressing the most significant problems (Stamatis, 2003, McDermott 2009). The FMEA Schedule was therefore used as a tool to select appropriate items for discussion, and in addition to prioritisation by RPN, further functionality was provided by categorising failure modes according to "work
package", "item/function/process", "responsibility" and "time line/project stage". This provided potential to group items by a particular task, responsibility or programme stage, according to the agenda and available expertise at each meeting.

Whilst essential for prioritisation, the FMEA Schedule was felt to be inappropriate for presenting failure modes to design team meeting participants; both because participants were expected to be largely unfamiliar with the process and language of FMEA, and also because the sheer quantity of failure modes might be considered overwhelming. The failure modes raised at each meeting were therefore added to a second "Tracking Schedule", which listed selected failure modes and recorded the resulting discussions, actions and ongoing monitoring requirements for each. Failure modes were ultimately presented at 8 design team meetings, however in practice this proved to be quite a limiting means of dissemination. Due to time restrictions just 36 failure modes were raised in total, and the selection of these was further constrained by the availability of specific expertise at particular meetings. Additionally, of the items raised, just 13 were deemed to require action, and of these only 5 are recorded as having been substantially resolved as a result of this process.

To capture learning from this exercise in time to apply a second iteration to Phase 2, a review was carried out part way through the Phase 1 construction process. To maximise objectivity whilst retaining access to the experiential learning this evaluation was carried out in discussion with the original researcher, but was designed and directed by a second researcher not involved in the initial exercise. Evaluation consisted of qualitative discussion of the researcher's experience of conducting the FMEA, combined with quantitative analysis of the output. To provide context for the assessment, functional analysis was additionally carried out jointly by the two researchers, allowing the functions which support achievement of design energy performance in buildings to be mapped and their interdependencies and associations more clearly understood.

RESULTS

Evaluation of the Phase 1 exercise has produced a number of key observations. These are presented below, as they relate to each core element of FMEA.

Identification of potential failure modes

Review of the FMEA Schedule indicates that significant expertise and experience was mobilised, with a minimum of 18 failure modes contributed by each respondent. Identified failure modes included a wide range of technical, organisational and behaviour related issues, ranging from highly specific component failure to general systemic shortfalls. Strong support for more extensive involvement of sub-contractors was provided by the timber frame sub-contractor, who contributed 43% of all failure modes. This numerical imbalance does however suggest potential for bias within the schedule, and lends weight to the general assumption that failure modes are most usefully elicited in a group setting (Stamatis 2003, McDermott et al 2009). The lack of contribution from the main contractor and architect is additionally considered a significant methodological shortfall, and suggests a need for clearer incentivisation to be provided to encourage participation by key stakeholders. A further significant shortfall was identified in the method of elicitation of failure modes, with participants being asked to identify "problem areas", but without the nature of the problem being
defined in the questionnaire. In this regard it is noted that whilst functional analysis may serve to focus attention on critical areas (Stamatis 2003), this was not applied in this instance. As a result the questionnaire headings provided as a guide for participants do not align closely with the functional analysis carried out as part of the evaluation, and whilst some such as "thermal performance of building fabric" might be expected to channel responses towards relevant function areas, others such as "design management" may have been too general to be helpful in this respect.

**Prioritisation of failure modes**

It is noted that the usual consensus based scoring of failure modes (Stamatis 2003, McDermott 2009) was not employed in the FMEA. Ranking failure modes based upon the opinion of a single individual as took place in this case is considered significantly inferior, as it fails to make full use of the range of expertise available. Additionally, no attempt was made to anchor or calibrate his responses beyond the generic, manufacturing focussed linguistic descriptions provided in the rating scales, and this individual reported a number of fundamental difficulties in respect to the ranking procedure. In particular he felt there was a lack of clarity regarding the metric for assessing "severity" with, for example, some failure modes relating purely to energy use, whilst others related fully or partly to occupant health and comfort. This resulted in one particular failure mode being ranked 11 out of 328, despite having no clear potential to influence energy use. The linguistic descriptions provided with the rating scale were of little assistance in this regard, being generated for the automotive industry and therefore expressed in terms of safety, statutory compliance and customer satisfaction. There was also confusion over whether the stated numerical probabilities for "occurrence" related to the chance of a failure mode occurring on an individual property, on the current Phase of 91 dwellings, or on the whole Exemplar site of 393 homes. The probabilities used in the scales were in any case rather low overall, with a rating of 5 for example equating to a 1 in 500 chance of occurrence. Finally a lack of clarity was reported as to whether the probability of "detection" related to detection of failure modes prior to the start of construction, prior to occupation, or over the life of the building. This is highly significant, as whilst failures routinely detected and remedied prior to occupation may be undesirable, they would not contribute to the performance gap. Overall, the original researcher reported finding a 1-10 scale unwieldy, and suggested that they would have been more confident using a 5 or 6 point scale. In terms of the results produced, the RPN values produced a relatively weak level of prioritisation across failure modes, with for example, the top rated 10% of failure modes accounted for just 14% of the total cumulative RPN score. This is perhaps to be expected however, given that the performance gap is widely understood to be comprised of multiple factors.

**Mitigation of failure modes**

As previously noted, only 36 failure modes (11%) were ultimately presented for mitigation at detailed design stage. This suggests strongly that a supplementary or alternative means of dissemination should be considered. Of these only 13 were considered worthy of action, and this perhaps suggests a degree of complacency on the part of these delivery focused stakeholders. This a view is supported by site observation conducted as part of the wider BEPIT project which suggest that, at the time of writing, 49 predicted failure modes had subsequently been observed as
occurring on site, whilst the fact that a separate "Tracking Schedule" was required to make the method more acceptable to these parties suggests a further lack of engagement. In practical terms it is noted that referencing of failure modes was not consistent between the two schedules, and that some failure modes were additionally altered when transferred to the Tracking Schedule. The result of this is that although it is possible to see which failure modes have been raised and actioned, it is less clear which have not. Effective filtering of failure modes within the "FMEA Schedule" was also found to be rather problematic, as failure modes were often defined as applying across multiple stages and work packages, and having multiple owners. The result of this was that an unmanageable number of combinations were typically generated; for example whilst only seven individual responsible parties were identified, these were presented in 59 different combinations within the schedule. It is expected that this issue could be resolved simply by ensuring a primary result is identified in each case.

**DISCUSSION**

As described above, previous research highlights some significant barriers to addressing the performance gap in mainstream housing. Firstly the performance shortfall appears to be composed of multiple components, which vary between projects depending upon numerous factors including the form of construction, management regime, heating and ventilation strategy and tenure. Secondly, contractual responsibility for resolving the multiple factors referred to above may be distributed across a number of parties. Thirdly, the test methods available to test as-built performance are underdeveloped, whilst the lack of design repetition between projects concurrently limits their usefulness.

It is proposed that an organisational approach based upon FMEA could address the first difficulty directly, allowing relevant factors to be predicted in a structured manner, on a project by project basis. Such an approach also has potential to mitigate the second problem by allowing issues ranging across different build stages and work areas to be managed using a single, relatively simple process. Finally, FMEA is able to operate using limited quantitative data where necessary, by instead making use of the qualitative expertise and experience of individuals within the project team (Stamatis 2003, McDermott et al 2009). However, despite this apparently excellent fit, it has been suggested that FMEA may require substantial adaptation to make it fit for purpose in a construction context (Mecca and Massera 1999). In particular there is a need to adjust the scope of the analysis and the metrics by which failure is defined. Where this scope extends to site work there is additionally a need to make allowance for significantly reduced levels of production process and performance related information.

Experience of applying FMEA on the BEPIT project has underlined the potential usefulness of the method, whilst also suggesting that considerable adaptation and learning may be needed to achieve impact in terms of identifying and mitigating elements of the performance gap. In terms of generating potential failure modes the exercise was successful, producing over 300 issues for consideration. The proportion of these which prove to be relevant will not be determined for some time, however site observations to date indicate that a significant number of predicted problems have already manifested themselves as predicted. The absence of key stakeholders in this process is however something which should be addressed, and consideration should be given to incentivisation to achieve full team, and ideally group-based, participation.
Sub-contractor input was also limited to a single organisation, and a practical means of integrating further trades would also appear to be highly desirable, albeit that this may need to take place at different points in the procurement process. In terms of organising the identified failure modes a rigorous approach has been identified as being beneficial. Stamatis (2003) suggests that function analysis can be useful in generating high level categories from which to generate failure modes. Categorisation by work package, work stage and responsibility have also been identified as needing refinement, particularly by listing a single primary result for each in lieu of multiple categories.

Prioritisation of failure modes through the use of rating scales to generate RPNs was more problematic. There appears to have been a fundamental lack of definition of the metric by which the failure modes were to be assessed, the scale of the system being assessed, and the timescales within which fault detection was required. It is considered that these issues could be addressed by providing bespoke rating scales, with descriptions relevant to construction, and to the performance gap in particular. These could additionally be tailored to reflect the expected nature of the failure modes based upon available research; that is failures which are common, non-catastrophic, and not routinely detected. Improved categorisation by function area might additionally be beneficial in this context, by allowing particular areas of concern to be separated into more manageable sized sections. Lastly, it is noted that scoring of scales based on group consensus is recommended (Stamatis 2003), and this should be considered to reduce any bias inherent in a single individual scoring all failure modes, and to increase ownership of RPN values; consideration might also be given to anchoring of rating scales by, for example, agreeing scoring for cases representing high and low values and of which all parties involved have a good understanding.

The final stage of the FMEA - mitigating failure modes - proved to be the most problematic of all. Overall, mitigation was only positively recorded for 5 failure modes, representing just 1.5% of the number initially identified. This success rate does not suggest an effective path to addressing an issue expected to be comprised of multiple factors, particularly given that 49 failure modes have so far been observed as occurring on site and which have not been prevented by this process. Early participation by key stakeholders has been identified as a factor which might go on to improve the understanding and engagement of individuals at mitigation stage. A more time efficient means of presenting failure modes must also be considered, and more effective categorisation of failure modes may assist in this regard by allowing whole “families” of failure modes to be identified and addressed in relation to particular responsibilities or work packages. Finally, as recommended by Stamatis (2003), monitoring of mitigation of failure modes might be usefully added to the main "FMEA Schedule", to make it clear which failure modes have been addressed and in what way.

Opportunity exists within the BEPIT project for iterative development of the FMEA methodology. It is therefore proposed to carry out a design stage FMEA relating to Phase 2 of the exemplar site, which consists of a further 69 dwellings. It is intended that this analysis will be adapted wherever possible to incorporate the potential improvements identified above. Following this a further evaluation will be undertaken to assess the success of these changes, and to identify further potential refinements. It is expected that it will also be possible to repeat this process for Phase 3 of the exemplar development, following which the accumulated learning will be incorporated into the BEPIT toolkit for public dissemination.
CONCLUSIONS

Based upon an analysis of the problem, the structured, cross discipline, multi data source analysis facilitated by FMEA appears to have great potential to assist in understanding and addressing the performance gap in construction. Experience of carrying out an FMEA on the BEPIT project has however highlighted the considerable challenges associated with utilising the methodology in a new application, and in an environment in which stakeholders have little or no familiarity with it. In particular the first iteration has identified the following specific needs:

- Incentivisation to ensure whole team involvement.
- Group based elicitation and ranking of failure modes.
- Tailored rating scales for prioritisation of failure modes in the context of the performance gap
- The use of function analysis to generate a framework for identification of failure modes relevant to the performance gap
- A more efficient means of presenting and mitigating failure modes at detailed design and procurement stage
- An FMEA Schedule format covering the full scope of the analysis.

By repeating the action research approach described in this paper iteratively across Phases 2 and 3 of the NW Bicester Exemplar Development, it is expected that significant progress can be made towards developing a commercially viable methodological framework for addressing the performance gap in new housing. Although challenging, the development of such a proactive technique applicable at development level is likely to have wide ranging value, both in terms of improving outcomes, and in reducing project risk.

REFERENCES


Department for Communities and Local Government (2013) "Next steps to zero carbon homes - allowable solutions". London: Department for Communities and Local Government.


