Edinburgh Research Explorer

Tactics, Objectives, and Choices: Building a Fire Risk Index

Citation for published version:
Koutsomarkos, V, Rush, D, Jomaas, G & Law, A 2020, 'Tactics, Objectives, and Choices: Building a Fire Risk Index', *Fire Safety Journal*. https://doi.org/10.1016/j.firesaf.2020.103241

Digital Object Identifier (DOI):
10.1016/j.firesaf.2020.103241

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Fire Safety Journal

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 23. Feb. 2022

Tactics, Objectives, and Choices: Building a Fire Risk Index

Vasileios Koutsomarkos, David Rush, Grunde Jomaas, Angus Law

Abstract:

This work summarises the key points that can be drawn from the extensive body of literature associated with fire risk indexing methods. A comprehensive definition of fire risk indexing is provided and the sometimes opaque mechanics of indexing are described in detail. Issues arising from fire risk indexing methods are explored, and the variety of terminology associated with this method is clarified. It is also explored how the development and operation of indexing methods are entangled with issues of reliable expert elicitation and professional competence of the end user. It emerges that the greater the complexity of a method, the more the workings of the method become obfuscated. This creates an inherent tension between the simplicity of the method and its transparency to the users – an issue the developers of fire risk indices ought to address from an early point.

Keywords: fire risk indexing; indexing methodologies; risk assessment; hazard evaluation

1 Introduction

Fire risk indexing (FRI) methods are heuristic models of fire safety. Heuristics are procedures that, in the absence of a formal underlying physical theory, provide a practical approach to solving problems [1], and are typically defined as efficient rules or procedures for converting complex problems into simpler ones [2]. Heuristic methods refer to problem solving that employs a practical method that is not guaranteed to be optimal or perfect, but is instead considered (by the method's designers) sufficient for reaching an immediate goal. Heuristics are defined as "relatively simple rules of thumb which can be applied to complex decisions where not all information is known, that can result in a suitable response" [3]. Heuristic methods are therefore an attempt to facilitate the process of making the best decision about how to respond to a problem – while simultaneously acknowledging that the problem has not been perfectly solved [4]. An example of such a heuristic method is the Glasgow Coma Scale [5], which is widely used by medical professionals for the evaluation of head injuries. It assigns points to three tests: eye, verbal, and motor responses. Based on the final score, inferences can be made on the patient's state of consciousness.

In the context of fire safety in buildings, the objective of a heuristic method is typically to make decisions about the fire safety measures that should be included within a building – often with the aim of deploying limited resources to maximum effect. These fire safety evaluation systems have been referred to by various names such as risk ranking, index systems, scoring, point schemes, and numerical grading [6]. In building design there are many parameters that may affect the overall safety; different environments may pose different risks, and different fire safety precautions may be deployed to mitigate these. Fire safety in buildings is therefore a problem that requires multiple attributes to be evaluated.

A risk index [6] is where a multi-attribute evaluation [7] is used to develop risk assessments and the results are aggregated into a single number. The process of creating a fire risk index must include a process of scoring the causal, and mitigating, fire safety attributes – with the result being a rapid and relatively simple fire safety evaluation [1]. The scoring process is typically undertaken by allocation of points to each attribute. The foundation of any fire risk index is therefore a points system. These have been applied to a variety of hazards and risk assessment projects to reduce fire safety costs [8], set priorities [9], compare design alternatives [10], and facilitate the use of technical information. A recent example is the use of a points system by the local authorities to address issues of building decay in Hong-Kong high-rises by prioritising which buildings are in need of fire safety improvements [11].

Indexing methods provide an approach that, during the development of a tool, can circumvent complex scientific principles of theoretical and empirical models, in order to aid decision making on the less than perfect circumstances found in real world applications [6]. It follows that once a points system is created, then it is (relatively) easy to use. For example, the Fire Safety Evaluation System (FSES) [8] allows an individual to undertake a survey of health care facilities and whether they comply with the Life Safety Code (NFPA 101), now incorporated into NFPA 101A [12], by grading the various fire safety components and comparing them to a set benchmark.

However, an indexing method is rapid because much of the thought and judgement that would be required in a conventional engineering analysis has already been undertaken a-priori by the designer(s) of the method. Creating the points system therefore requires the method's designer to assign values (and possibly a weighting) to a selection of fire safety attributes. This can be totally arbitrary, but given the proliferation of risk indices, a series of systematic approaches have been deployed in order to assign a weighting. In most of the cases, this has included a group of 'experts' in the field – and the group have defined the weighting based on the group's collective professional judgement and experience. The attributes that the group are required to weigh can include policies, objectives, strategies, and the components that make a contribution [13]; these can represent both positive and negative fire safety features. By assigning grades on which a calculation is then performed, it is possible to arrive at a single value or index in order to obtain relative, yet comparative, levels of fire safety [5], [14].

Fire risk indexing methods have most often been developed with the purpose of simplifying the risk assessment process for a specific type of building, with their major advantage probably being their simplicity [15]. Due to the relative rapidity with which an index can be generated, FRI methods are considered to be very cost-effective tools [5]. The use of such a method can help practitioners decide when a more detailed quantitative analysis may be necessary [16]. A standardised procedure may be of particular advantage where an organisation is intending to assess a large number of similar properties [9], [17]. For example, a rating scheme was used for a systematic survey of 25 historic buildings in Portugal after a major conflagration of 18 buildings [18]. On a larger scale, risk indices have been repeatedly used for the monitoring of wildfire risk. This has been done with the integration of multiple variables (either dynamic or static) in a single system using remote sensing tools [19]. In the domain of building fires, rating forms were used for the assessment of Wildland Urban Interface (WUI) fire risk, such as “NFPA 224 - Fire Protection and Prevention for Summer Homes in Forested Areas”, which was created in 1935 and now is incorporated in NFPA 1144 [20]. Taking both into account, a combination of WUI fires and remote sensing techniques has been employed in an index in Norway [21].

To assist any assessment procedure, it is possible that the underlying calculations of a fire risk index can be programmed to produce a software. If aspects of a building’s fire safety measures are linked to costs (e.g. cost per linear meter of fire resisting construction), then the software can be programmed to rapidly iterate alternative fire safety measures that achieve the same overall index, but by different means – this process can be used to minimise the cost for a predetermined safety level [22].

This paper presents a review of the underlying mechanics of existing indexing methodologies in an attempt to build on the knowledge base from previously developed indices. To ensure clarity within the review, issues around the inconsistent use of language within the literature studied is first explored; the terminology used is defined. The connection between the mechanics of a method and expert elicitation procedures is then explored, along with its implications. Finally, the topic of the competence of the end user is brought forward and discussed in the context of how anticipated user competence can affect the development of a method.

The motivation for this work is to understand whether FRI can be used as a means to motivate stakeholders to exceed the regulatory minimum for the purposes of achieving more resilient assets. To explore if FRI can provide the supporting metrics, a review and relevant work by Watts [1], [6], [7] was used as a starting point for this study as this work led him to the formulation of specific criteria [15] that provided guidance of good practise for future developers. Building on Watts' work, this paper aims to provide a review of the issues around FRI methodologies with the aim of providing original and necessary information for the development of any future indexing tool.

2 Use of Language and Difficulties Regarding Terminology

In undertaking this review, it has been found that the terminology has fluctuated throughout the years. The authors have encountered different phrases in the literature to describe the same notion. Strongly linked with indexing are the words 'rational', 'system', and their derivatives. The first mention is found in the Fire Grading of Building reports where it is worded that "any **rational system** of fire grading should provide a combination of active and passive defence in proper balance to meet the fire hazard in each case" [23]. This terminology is also used by Watts in his PhD Thesis "A theoretical **rationalization** of a goal-oriented **systems approach** to building fire safety" [24]. Malhotra also called for "a more **rational approach** to fire safety" [25], [26].

Copping [27] has referred to indexing as a '**rationalised systematic** approach' to fire safety meaning "the use of qualitative descriptions of events, techniques and processes to which are attached numerical values assigned by a group of experts", listed as one option of analytical approaches to fire risk assessment. Shields [28] attributes the first mention of this term to Marchant, while highlighting some confusion in the terminology across the literature. Idris [29] however has simply referred to 'the **Systematic** Approach', with the authors' estimate on its origin being the phrasing of the criteria for fire risk ranking as "to elicit subjective values systematically" [15].

Beard [30] was the first to introduce the concept of a **systemic** approach in fire safety by calling for a guiding structure (methodology) of a dynamic nature to support analyses. However, a systemic approach is not the same as a systematic approach as "the word 'systematic' may be thought of as implying 'methodical' or 'tidy'"; in this context, 'systemic' implies the capability to see the 'dynamic wholeness' in a situation. This description was later used by Shields [31] who described a points scheme as a "product of a framework which constitutes a **systemic approach** to fire safety evaluation". Watts [32] has also used the term when he stated that "fire-risk indexing is a **systemic approach** to code equivalency". Copping [27] has used the term 'systemic' as well stating that his work "promulgates a **systemic approach** to fire safety, in which a holistic philosophy is adopted".

This phenomenon has been observed in another publication for the terms 'model' and 'scheme' by Shields and Silcock [28], who made a review of the terminology which they considered "necessary since the current interest of the authors is in the development of a method of evaluating the provision of fire safety in buildings". Similarly, the present authors thought it useful to touch upon the same issue, since it was found that confusion remains around terminology relating to FRI.

In undertaking this review, differences in terminology made it difficult to compare methods because the literature occasionally uses different words used to describe the same notion. Therefore, the following sections are used to determine and define the terms that will be used in this work, the same way it was done by some developers to present their works [7], [33], [34].

According to Watts [7], fire safety attributes “provide a means of evaluating goal achievements” and practically identify the ingredients of fire safety. Depending on the conceptualisation of the method’s structure, these attributes can be grouped according to appropriate hierarchical levels or purpose groups, which are then named correspondingly.

3 The Core of Indexing

The practical necessity of trying to assess multifaceted fire risks in a variety of building types has led to the creation of several FRI methods with Watts [6] having referred to a generalised procedure in the ranking of fire safety as follows:

1. Identify hierarchical levels of fire safety specification
2. Specify items comprising each level
3. Construct and assign values to matrices of each sequential pair of levels
4. Combine (multiply) matrices to yield importance ranking of items
5. Verify the results

Based on this procedure, there are a number of fundamental concepts which must underpin any such scheme. These can be consolidated into three key decisions or judgements that the developer of any FRI must make in order to create their scheme and details about each are presented in the corresponding following subsections. The judgements are as follows:

1. Attribute identification – a decision must be made about which fire safety attributes are going to be evaluated by the FRI.
2. Attribute weighting – a decision must be made about the use of relative weights for each attribute or the group in which they belong, along with which weighting method is used.
3. Index calculation – a decision must be made about the mathematical functions (or calculation style) used to calculate the final index based on the attributes chosen, and each attribute's relative weighting produced.

The development of a scheme in these three stages requires some form of communication between the developers and the group of people that is being advised upon, usually called 'expert group' or 'expert panel'. There are different protocols and forms for that communication, and they can alter throughout the different phases of development. The prevalent approach used is the Delphi method, and the expert group is referred to as a 'Delphi panel' [8], [9], [31], [34]. These issues will be discussed in detail in section 4. Irrespective of the decision making process, the categories of decision that must be made are universal and are described in the following subsections.

3.1 Attribute Identification

Three possible ways of choosing the attributes to be evaluated have been reported in the literature: 1) An arbitrary definition and choice of the attributes by the developers [35]; 2) the attributes are derived based on components from the corresponding prescriptive guidance [36]; and 3) some combination of both of these – whereby an initial list of attributes is populated based on rule based guidance, and this is subsequently refined by asking an expert group whether attributes should be included [9].

3.2 Attribute Weighting

Once the attributes have been identified, it is necessary to assign weightings to these. It can be possible that each attribute is given an equal weighting. However, it is also common for some attributes to be assigned a greater weighting (and thus more importance in the overall index). The arrangement of attributes and the assignment of weightings requires a structure known as a ‘hierarchy’. This originated in the work by Marchant [9], when the expert panel was asked to consider an ‘orbital hierarchy of levels of fire safety’. Soja [37] also presented the hierarchy in a tree form. The levels, in order of importance, were policy, objectives, tactics, components, and sub-components and these are shown in Table 1. Together, Shields [31] characterised these hierarchies as a “finite ordered family of collections”.

When the structure of the method is chosen to be a hierarchy, then the terminology of those different levels needs to be defined. Throughout this work, for consistency reasons, the attributes will be termed based on the level they belong in accordance with Table 1.

Table 1. Attribute terminology depending on the hierarchy levels (adapted from Soja [37] and Marchant [9]).

Level	**Term | *Variants***
1	Policy
2	Objectives
3	Tactics | *Strategies*

4	Components | *Parameters* | *Factors* | *Variables* | *Elements*
5	Sub-components | *Sub-parameters* | *Sub-factors* | *Sub-variables* | *Sub-elements*
6	Survey items

Within a hierarchy, there are four weighting methods used in fire safety engineering (FSE). Some have been presented by Donegan [38] and are summarised herein. The four methods are: Edinburgh Cross-Impact Analysis (the Edinburgh model); Hierarchical Cross-Impact Analysis (HCIA) Methodology; The Analytical Hierarchy Process (AHP); and Reliability Interval Method.

Each of these weighting methods have some common concepts. In each case, to achieve a mathematical formulation of the hierarchy, each parameter is represented by a matrix, known as an effectiveness matrix.

Based on the five-level hierarchy conceptualised already, the basic matrix types needed are the following: the ‘components-to-tactics’ matrix (C/T); the ‘tactics-to-objectives matrix’ (T/O); and the ‘objectives-to-policy matrix’ (O/P). The matrix product C/T × T/O × O/P yields a ‘components-to-policy’ vector that, when normalised, gives the desired prioritisation weights in percentage form. This is what Watts [6] considers a ‘mathematical manipulation’ and includes the use of a weighting method. To populate these matrices, judgements must be made by the expert panel in order to assign values of importance. This expert elicitation process requires the choice of a communication protocol, which must then be integrated into the method.

These weights, or values of importance, are assigned on a scale. There are several different ways in which attributes can be scaled [38]. Those of interest here are as follows:

- **Ordinal Scale**: The ordinal scale ranks each attribute or orders them.

- **Interval Scale**: More mathematically tractable and of greater importance in quantitative assessments, the interval scale is a continuous scale between two points. Relative difference is maintained; that is, equal intervals of the scale have the same meaning.
- **Ratio Scale**: The ratio scale is an interval scale with the absolute zero property; in other words, one end is fixed so that the values on it are absolute rather than relative.

3.2.1 Example of Hierarchical Weighting

To further elucidate this hierarchical approach [13], an example is illustrated with a conceptual hierarchical three-level tree as in Figure 1.

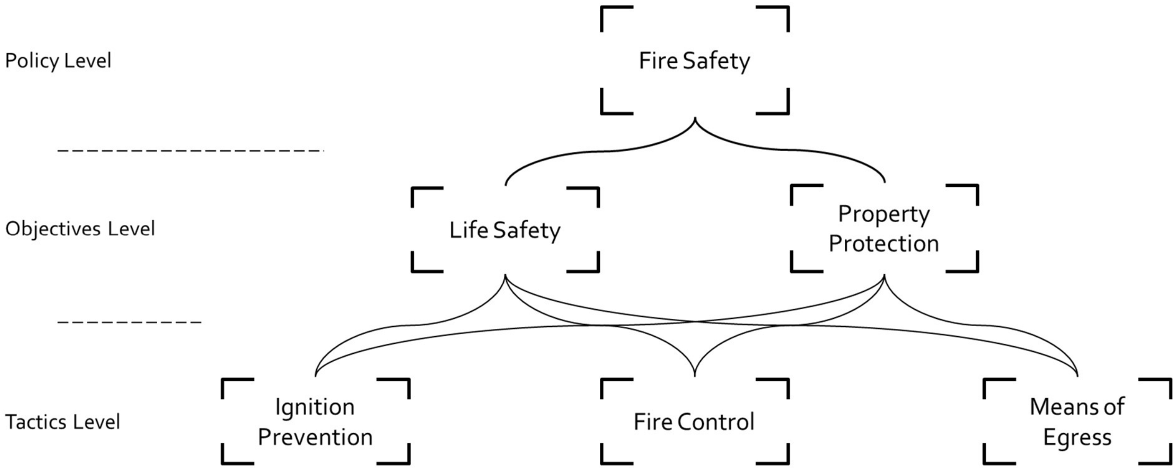

Figure 1. A simple Fire Safety Hierarchy

The numbers used to populate the matrices in this example are illustrative, arbitrary, and have been based on the ordinal scale. First, the tactics-to-objectives matrix is formulated by employing the expert group and asking them to answer the question (on an ordinal scale of 0 to 5) "how important *x*-tactic is to *y*-objective?". Its relative contribution is also calculated in another equivalent matrix – the one that will be used in the mathematical calculations – by dividing the attribute's number of importance with the maximum number it could have been allocated. Rows represent the tactics, columns the objectives, so that (for example):

$$T/O = \begin{bmatrix} 5 & 2 \\ 2 & 5 \\ 5 & 1 \end{bmatrix} \asymp \begin{bmatrix} 1.0 & 0.4 \\ 0.4 & 1.0 \\ 1.0 & 0.2 \end{bmatrix}$$

The same process is followed for the next level of the hierarchy, which is "How important *x*-objective is to *y*-policy?". Rows represent the objectives, columns the policy, so that:

$O/P = \begin{bmatrix} 5 \\ 2 \end{bmatrix} \asymp \begin{bmatrix} 1.0 \\ 0.4 \end{bmatrix}$ and their product yields a tactics-to-policy vector: $T/P = \begin{bmatrix} 29 \\ 20 \\ 27 \end{bmatrix} \asymp \begin{bmatrix} 1.16 \\ 0.80 \\ 1.08 \end{bmatrix}$

This, when normalised, gives the desired prioritisation weights in percentage form. For example, the importance of Fire Control (Tactic 2) to Fire Safety (Policy) is:

$$\frac{20}{(29 + 20 + 27)} = \frac{0.80}{(1.16 + 0.80 + 1.08)} = 0.2632 = 26.32\%$$

Similarly, the percentage contribution of each tactic is calculated with Ignition Prevention at 38.16%, Fire Control at 26.32%, and Means of Egress at 35.53%. However, since the range of the original matrix grading is coarse, these are usually rounded to single figures to simplify the calculations and to represent a warranted number of significant figures [9].

This is the simplest approach for ranking attributes. In some cases, statistics, experimental results, or case studies have been provided by the developers to the panel members to be taken into account in the weighting process. This approach has been expanded with additional features that led to the formulation of individual methods with their own benefits and flaws, all of which will be presented herein.

Once the method's designers determine the relative weights, then the method is ready to be applied. The developers provide users with working sheets that contain only the lowest level attributes (i.e. the tactics). Users then assign a building specific grade to each attribute based on their own judgement if no further guidance is provided (e.g. if the ignition prevention measures are perceived to be excellent, then a score of 5 might be assigned). In the case study above, if all of the tactics receive the maximum grade of 5, then the assessment will yield a result of 500 points. This result can be normalised if the developers desire to present it in percentage form. The following table shows the process for the above example, if the user/assessor provides assessment grades for each tactic. The tactics' grades are multiplied with the corresponding percentage contribution and the final assessment is as shown in table 2. Based on this final score, inferences can be made about the fire safety provisions of the building being surveyed.

Table 2 – Calculation process and final result for the hierarchical example.

Tactic	Assessment grade	Percentage contribution	Final Result
Ignition Prevention	4	38%	152
Fire Control	5	26%	130
Means of Egress	2	36%	72
		Sum	354/500

This is how the calculation works with the relative contribution matrices. However, there is another option to facilitate this calculation; this would be to calculate the percentage contribution matrices of the attributes (the value of the attribute divided by the sum of the column's values), in every stage, and use that number in the calculation. This would eventually give slightly different results. In the Edinburgh model, the 'expert' panel decided to use the relative contribution matrix approach, but no explanation behind the thinking is provided in the original report. This is an example of a common issue with the lack of extensive and detailed documentation in the development of some indices [9].

The approach described above assumes that attributes are independent. This might not be realistic, and Magnusson and Rantatalo [39] state that “listing them as independent implies [a] redundant protection system and the concept of ‘defend in depth’ which has been applied in other areas”.

To account for the possibility of interactions between attributes in a hierarchical structure (e.g. that poor fire control measures may increase the importance of means of egress) various authors have sought and incorporated this concept into their methods using ‘parameter interaction’.

3.2.2 Parameter Interaction

The Edinburgh Cross-Impact Analysis (1982) introduced a ‘parameter interaction matrix’ on the lower (component) level. This was added in the calculation process to incorporate the “degree to which the contribution of a component to fire safety is enhanced by the interaction of other components”[38]. The process has been presented by Stollard [40] and Soja [37].

Soja [37] phrased it as “with two or more components working together, their combined value is worth more than the sum of the individual contribution, so a method was sought to express this idea in numbers”. If 1000 points were to be gained by a system with no interactions, then with attributes working together the points should eventually sum to a larger value to correspond with the conjugated impact.

During the development of the Edinburgh Cross-Impact Analysis, five different mathematical approaches for this interaction were investigated and their impact on the produced values was analysed [9], [37]. The reasoning behind the final choice of mathematical approach was based on the idea that the combined action of the attributes provided a 20% increase in the total points, which “was considered to be correct in relation to the thinking that went into formulating the interaction array" and facilitated a good balance so that the interactions would not play a more important role than the one the components already had.

To conceptually incorporate this method in the example above, a 3x3 interaction matrix is required. This uses the same scale (i.e. 0 to 5), and is used to rate how one attribute affects another on the same level. That leads to the formulation of a tactics-to-tactics matrix (T/T). When its values are normalised, the matrix expresses the relative contribution in an Attribute Interaction Array (AIA). Then, the following formula is used to quantify the interaction of attribute i with attribute j (for a total of 3 attributes):

$$V_{int}(i) = \frac{\sum_{j=1}^{3} \left[\frac{V_{or}(i) + V_{or}(j)}{2}\right] AIA(i,j)}{3} + V_{or}(i)$$

Where:

- V_{int}: the enhanced weighting after the interaction has been taken into account
- V_{or}: the original weighting of the attribute
- AIA: the attribute interaction array

The calculation of the interaction between the same attribute is omitted because it cannot enhance its own performance, so that AIA(i,j)=0 for $i=j$. For reasons of brevity, no further calculations will be presented on the interaction. Nevertheless, implementing the formula for all the tactics breeds a new parameter interaction matrix, which sums to more than 100%. In order to maintain the functionality provided by a percentage scale, the interaction matrix is then normalised.

Using this approach, the relative importance of the attributes now incorporates a more 'realistic' interaction between different attributes. It is, however, repeatedly highlighted in all the resources of the Edinburgh model [9], [37], [40] that these interactions parameters have been created on the assumption that each attribute achieved a perfect grade (i.e. five out of five). If this is not true (which is almost always the case), then the parameter interaction should be recalculated (where the $V_{or}(i)$ and $V_{or}(j)$ values are modified to reflect the survey grade) to correspond with the building-specific attribute grade.

Reflecting on this approach to interactions, Soja stated that this approach was “reasonably valid, but does not give a totally true picture of the situation. Giving simple multiplying factors to each component implies that the value of a component is not linked with the scores of the other components and so the logic of saying that [the final score matrix] is determined by the scores of the interacting components is not carried through to this calculation”.

A second, alternative approach was created whereby using a tabular method which operates on a 1000 point scale. When interactions are accounted for, they add even more points, to a maximum of approximately 1200 points. Building specific attribute grades affect the interaction points earned.

There is also a third and final approach, which is the most ‘accurate’, but it necessitates the recalculation of the final attribute weights by modifying in the calculation procedure the original percentage contribution values so that they correspond to the building-specific grades allocated. Then, the final score is calculated by multiplying the attribute grades with the new non-normalised weightings. For practical reasons, this method requires implementation within a computer algorithm.

Soja suggested the use of the simplest approach for initial assessments and, only if additional ‘accuracy’ is required, resorting to the other two options. In the final version of the Edinburgh model, the first method was employed. This could be because, as Soja stated, “the coarseness of the survey grading suggests that any intricate and detailed mathematical use of the survey grades would be inappropriate and would give only a misleading sense of accuracy” [37].

This approach to the interaction between attributes was further developed by Donegan et al. [41] and formalised to what is known as the Hierarchical Cross-Impact Analysis (HCIA) Methodology. This method addressed the interdependence of attributes in every level of the hierarchy with the introduction of sequential perturbations in the matrix multiplications. When Shields [31] compared the two methods, he found that the spread of percentages in relative importance was compressed with the introduction of sequential perturbations, meaning relative weighting values closer to the mean, creating the tendency for the components to have an equalised contribution. Additionally, rank reversal between some components was observed, but the percentage contributions remained similar when rounded, so overall no major implications were introduced in the operation of the method. Nevertheless, this new approach is considered theoretically more formalised and comprehensive by Donegan [38].

3.2.3 Criticism and Other Interaction Methods

One limitation that was identified with the Edinburgh model and the HCIA is that the experts may not apply the same level of rigour to their assignment of rankings across each attribute. This is a serious drawback of the methods as there is no way to check the consistency of decision making across all the attributes that have been ranked. To address this shortcoming, Shields and Silcock [13] explored the utility of the Analytical Hierarchy Process (AHP) in FSE, a process created by Saaty [42], because it introduced a consistency check. The aim of the AHP is to introduce order and objectivity into the largely subjective process of attaching weights to a set of decision criteria [31]. The AHP relies on the pairwise comparisons of components to define their relative importance using an interval scale. Donegan [38] has described AHP by stating that "the procedure entails the comparison of all the pairs of individual attributes at each level relative to each attribute in the superior level. The intrinsic complexity of the process prohibits a simple description". It has been presented and explained extensively in other publications [13], [14], [31], [38], [41], [42].

The mathematical background of the method is based on the solution of the characteristic equation of the effectiveness matrices to find the dominant eigenvalue, which then facilitates both the calculation of the parameter relative importance and of a consistency index, which is suggested to be lower than 10% for a valid assessment [14], [41]. The method becomes unstable when more than 7±2 parameters are ranked [6], [7], [39]. This was originally investigated by Saaty [43] because Miller's [44] conjecture in psychological theory states that the number of seven parameters is the limit for the amount of information that an observer can give about an object from an absolute judgement. That spread emerged in various psychological experiments, but in fear of this being just a 'pernicious coincidence' Miller proposed to withhold judgement. However, Saaty and Ozdemir [43] investigated further, and (serendipitously) found that, due to the underlying mechanics of the method, meaningful conclusions on the results' consistency cannot be drawn when the number of parameters exceeds seven. This is also supported by an observation from Shields and Silcock, who "experienced some considerable difficulty in carrying out consistent pairwise comparisons when more than five components were under consideration at any one time" [13]. They suggested that in order to use the AHP with more parameters, "components will have to be grouped into clusters according to their relative importance and pairwise comparisons performed on the clusters".

Most recently, the fourth distinct method was developed in 2005. After assessing previous experience in the development of point schemes, the option of using fuzzy aspects [45] from probability theory was explored, detaching the weighting process from the AHP and employing the Reliability Interval Method to calculate the relative importance of attributes by Lo et al. [46]. With such fuzzy assessment of weights (the 'expert' grades in a range of possible values, not a single number), a statistical analysis of the results is possible by calculating the parameters of reliability, centre variance, and interval variance. The benefit is that these three factors can facilitate a consistency check of the responses, but the methodology does not account for any interaction between the parameters.

3.3 Index Calculation

The calculation of an index can occur once the weighting of each attribute (w_i) has been determined by the method's designer, and the score (x_i) has been chosen by the user – where subscript i represents each individual attribute. For the building fire safety indexing, there are several methods of calculation. These were described by Sugahara [47], who categorised them into different styles as follows:

1. Additive style. Component grades (x_i) are multiplied by their weight (w_i) and summed to produce a score. The calculation of the index in the Edinburgh model (as per the example above) is of this nature. This is notated as $\sum w_i \times x_i$;
2. Multiplicative style. Component values are multiplied in order to reach a final score. Multiplying probabilities in event tree analyses falls in that category. Notated as $\prod x_i^{w_i}$;

3. Divisional style. Where a ratio of values calculates an index. The basic formula of the Gretener method operates in that fashion, to assess the efficacy of countermeasures against possible fire hazards. Notated as $\prod x_i^{w_i} / \prod x_j^{w_j}$;
4. Vector style. Where the absolute impact of the grade is accounted for. However this method is of limited use in FSE, only found in Sugahara [47]. Notated as $\left(\sum (w_i \times x_i)^2\right)^{1/2}$;
5. Mixed style. Any combination between the basic four styles.

Each of these styles have different implications for the indexing method. Watts [6] noted that "the implication of addition is that there is no interaction among the fire safety parameters" and that "multiplication implicitly suggests these factors are interdependent". It then follows that the choice about how to treat attribute interactions is not limited to aspects of the weighting process, but is interdependent on the choice of the above styles too, so that eventually influences the shaping of the overall method structure as well.

3.4 Discussion

The previous subsections described the conceptual and mathematical structure for any indexing method – and each of the key methods that are used in the field of FSE. What emerges is, at its core, a simple concept. However, layers of complexity are overlaid onto this perceived simplicity. In each case the complexity is introduced in order to correct for a perceived failing that is inherent to the more simple method. For example, the idea that fire safety attributes are not independent is a simple concept to grasp, but the approach by which a method's designer might choose to incorporate this can become unwieldy as they are scaled across many attributes. Furthermore, any attempt to address obvious problems with a simple model frequently introduce new, less self-evident problems of their own. For example, while a method's designers might be unanimous that two parameters should be considered to interact, the relative importance of this interaction (when compared to other parameters) may be the subject of disagreement. Thus, the addition of complexity has the tendency to create the appearance of 'solving' problems where, in fact, it may simply bury a new set of issues so deeply within the workings of a method that it becomes difficult (even for a reviewer) to fully untangle these. On reviewing these methods, it emerges that the more complexity is added to the method, the more the successful application of this method become obfuscated.

Similarly, it should be noted that all of these concepts are usually hidden from the user of the method. Indeed, as discussed, one of the major advantages of FRI is the simplicity of the method for the practitioner. Thus, by design, it is seldom that the user is exposed to the underlying mechanisms of the method they use, and even more rare that they understand their implications.

The utility of any method is closely linked to the perception of the final index. Some may regard an indexing method as a well organised structuring of the risk assessment process; others may see the final index as a panacea that a building is 'adequate' when it achieves an acceptable score – if an acceptable score limit is set by the designers. There are inherent dangers in such views – a lesson to be learnt by the fate of the Edinburgh model, which was superseded because "the arbitrary interaction between factors could lead to an acceptable risk score and an inadequate fire strategy" [48].

Conversely, there are documented cases in the development of the NFPA's Fire Safety Evaluation System where a design could be compliant with the prescriptive Life Safety Code, but fail to meet the minimum score of the rating schedule [8]. This is an inherent challenge with any method that presents a single pre-defined acceptability criterion. According to a report that reviews risk assessment methods [49], “not all ranking methods include a basic level for a satisfactory protection, but give only a relative position as situation A is better/ worse/ equivalent to situation B. This can be an advantage for the user [who] can define [their] own level of protection, but in practice, most inexperienced users want that an expert system gives them a clue on ‘what is good enough’”. Similarly, Hultquist and Karlsson [50] have found that “it is quite possible [to] achieve a good index rating by giving some parameters a very bad rating and other parameters extremely good rating. In spite of the good index rating, the resulting building design may be totally unacceptable or absurd from a fire safety point of view”. This is linked with issues of ethics and competence of the method’s user, which will be partly explored in the following section. Still, it would be suitable to draw a corollary with the quote by Bullock and Monaghan, which states that “the ethical imperative on a competent professional fire engineer is to ensure that anything that he or she is submitting for said approval passes his or her own test of adequacy” [51].

The review presented in this section therefore shows that there is a tension inherent to the mechanistic operation and deployment of any FRI method. A simple method has many flaws and these can lead users to inadvertently produce a ‘good’ rating for a ‘bad’ building. Correcting these flaws requires complexity to be added to the method, and can prevent inexperienced users from inadvertently producing an ‘incorrect’ result. However, added complexity has two consequences – 1) that the workings of the method becomes so complicated that additional flaws are added that are difficult to understand and evaluate; 2) the elimination of the ‘incorrect’ result leads users to regard to method as a fire safety panacea – rather than a structured risk assessment.

4 Expert Elicitation in Fire Risk Indexing

It has been shown that every indexing method requires a series of judgements to be made – not just about the method design, but also about the weights that are allocated to specific attributes and (interactions considered) how interactions between different attributes are quantified. The previous section focused on the mechanistic aspects of the method's design, but circumvented the question of how data are generated by the method's designers to determine the weighting. This section presents a review of expert elicitation methods in FRI in the context of the decision making process required in order to populate the various matrices described in the previous section.

4.1 Delphi Use in Fire Safety Engineering

The available literature suggests that experts are typically assembled in the form of a Delphi panel [6], [31], [34], [38], [52]. The Delphi technique is thoroughly presented by Linstone and Turroff [53], who define that "Delphi may be characterized as a method for structuring a group communication process so that the process is effective in allowing a group of individuals, as a whole, to deal with a complex problem".

The core structure of the (classical) Delphi method is a series of questionnaires that are designed by the experimenters, and are answered anonymously in numerical or literal format by a group of responders (usually considered 'experts'). Between every round, controlled feedback is provided, usually in a result summary form so that the group members can comment on the responses or the summarised results, with the aim of fostering convergence. That is why the responders have the option to alter their answers and the process is iterated until consensus – or a level of response stability – has been achieved [54].

The Delphi method was developed with the use of questionnaires to formalise a process that deals with subjective values and to assist decision-makers to structure, quantify, and evaluate a problem.

The Delphi method was first employed in fire safety engineering by Nelson and Shibe [8] in the 1980s. Similarly, it was used by Marchant for the development of the Edinburgh Model [9], [40] and by Shields et al. [55] for the evaluation of dwellings. Later Stollard et al. [56] used it for the development of a points scheme to calculate running costs in hospitals. A well-documented use of the method can be found in Karlsson and Larsson [34].

When reviewing the employment of the method in FSE, Shields [31] identified cases where anonymity was not ensured and face-to-face discussion between the ‘experts’ occurred. Shields noted that in FSE there had been many deviations from the classical approach, mostly due to the fact that “Delphi administrators participate as Delphi Group members, act as members of advisory and vetting committees and engage in face-to-face meetings”. He noted that “all such actions are in conflict with the precept of anonymity”, but that these deviations were “apparently acceptable [to the method’s designers] on the basis of procedural expediency”. In another publication, Shields et al. stated that “the ‘modified’ Delphi process maintained anonymity of response, i.e. ‘secret ballot’ but introduced confrontation to expedite convergence” [52]. He finally judged that “the validity of the conclusions and output information of any studies obtained using variations of the Delphi procedures must therefore be called into question”.

At the time of their review, Shields et al. [52] recognised some methodological issues to be addressed with the use of a Delphi Panel. Problems that had been encountered in FSE were associated with choice of experts, group attrition, anonymity, scale use, and the objective evaluation of consensus. Marchant [57] addressed some of these issues following the rationale that “it might be accepted generally that FSE is too complex for the opinion of one expert to be valid”, so multiple experts of different background are needed, yet that leads to different levels of confidence when answering the questionnaires. He noted that “if the experts’ common knowledge base is [the reason] for the anonymity, then anonymity should not be preserved when a group of heterogeneous expertise is employed, in order to conduct meetings which would facilitate a ‘group education to achieve a degree of common understanding’”.

Shields [31] expanded on this discussion and explored the use of ‘conviction weightings’ in the matrix population process. That is, every panel member would also submit a grade from 0 to 5 that would signify their conviction – or certainty regarding their expertise – of their evaluations on each whole matrix. This could then facilitate the calculation of more ‘accurate’, weighted relative weightings that were a ‘truer’ representation of the members’ varying expertise. Tangential to this concept and a theoretical improvement is Cooke’s method, which has been used outside FRI, but still in a FSE context [58]. Following this method, the expert group members are ‘calibrated’ before the elicitation exercise by answering to questions to which the answers are known. Depending on the level of accuracy of their answers, their expertise - or predictive capability - is evaluated and used to weigh their answers. In regards to Delphi use outside FRI but in FSE, Harmathy [59] considered the Delphi method to be a complement to research and believed that the provision of fire safety is an ideal area for application of the method; he subsequently used the Delphi method for the generation of supplementary data in the development of a tool to make decisions about how to ‘trade-off’ fire safety measures [60]. A comparison with other fire risk analysis methodologies found that systems developed with the Delphi approach were more comprehensible in addressing an analysis throughout the various stages of a building fire [47].

4.2 Selection of Experts

The general simplistic term is that “an expert is regarded as one practiced or skilful within the area of consideration” [52]. When forming steering or ‘expert’ groups, such subjective definitions may easily lead to a selection bias, thus increasing the chances that a group of ‘experts’ will be set up that is not representative of all ‘experts’ in the field of the problem involved [61]. This becomes even more complicated, if one takes into account the fact that in the first Delphi exercise where ‘experts’ were consulted in a group “it was found that confidence in prediction does not necessarily show a correlation with success in prediction” [62], which could be interpreted as a foretelling of the Dunning-Krueger effect [63].

As a result, it was observed that developers of indexing schemes choose ‘experts’ with a pattern, even though expertise is not clearly defined. This pattern of selecting ‘experts’ was mentioned by Shields et al. [52] and originally presented by Rowe et al. [64] as:

1. Persons who are involved in the general area of study and possess some minimum formal criteria, e.g. membership of a professional body.
2. Persons who are known by the researcher.
3. Persons who by reputation are informally known by the researcher.
4. Persons who are readily available or can be pressed into service.

Ironically, this observed pattern was used as a guide when selecting members of an expert panel by Idris [29] “for these studies, the group members involved were selected from the following:

- Government officials,
- Persons who are known to the researcher,
- Persons who are readily available for service,

- Professional associates.”

The developers of various methods have over the years discussed the key issues associated with the choice of experts. Shields [31], perhaps suggesting some cynicism about the expertness of his expert panel, employed a “randomly selected group of final year honours students from the Faculty of Science and Technology in the University of Ulster […] to make similar value judgements as the current Delphi group”. He evaluated the reproducibility of the result between the ‘experts’ and the students. The comparison of the results between the two groups showed minor deviations, so the conclusion was that the results could “tentatively support the hypothesis that, given a collection of issues, bounded informational sources and well defined procedures, the expertness of group members is not an issue”. Shields suggested that group membership may be determined on the basis of “an awareness of the various issues being considered rather than any particular expertise”.

However, the literature is contradictory as there are two recorded cases (to the authors’ knowledge) where the opposite happened. One case was briefly reported by Lo [65] who compared the results of two ‘expert’ groups of different backgrounds and found that “the background of the experts may affect the weightings. Another set of weightings was obtained from a panel of building services engineers […]. Accordingly, further studies on the selection of experts should be carried out”. Lo et al. have presented in detail the differences of the expert input in another publication [66]. Similarly, Ibrahim et al. [67] compared three different groups on how they perceived the importance of four attributes and the results differed significantly.

Both cases highlight the need to evaluate the consistency of an expert group and robust procedures in place for the selection of ‘experts’ – though it remains undefined what these procedures should be.

4.3 User Expertise

The design of any indexing method requires judgements and decisions to be made by a group of ‘experts’, while also the user (i.e. the person undertaking the grading on a real building) plays a critical role in the final score generated by the method. The expertise of the user has therefore also been of concern to the developers of indices.

Kinsey et al. [3] believe that if “a fire engineer is lacking in-depth technical expertise or experience about a given subject, they may rely on fire codes for guidance, potentially without a complete awareness of any underlying basis or assumptions for the guidance”. However, Stollard [68] stated that for a reliable analysis to be conducted “it will be necessary to know not only the contents of the legislation and guidance, but also the basis on which these documents were developed”. This is relevant to the statement that “Fire engineers need to understand what the codes are effectively saying in engineering terms and, equally to understand the flaws in the codes or those parts that are not based on scientific or engineering principles” [69]. It would be ideal if such knowledge could be fostered from the education level, as the phenomenon is touched upon in a proposal for a model curriculum in FSE where it is mentioned that “standards should never be introduced without an evaluation of the engineering and scientific background” [70]. These quotes are of relevance, because when Hultquist and Karlsson [50] were evaluating an indexing method against another more sophisticated one, they reported that “during the work it has also become clear that the method can be misused, if an engineer consciously wishes to misuse it”. While developing a ranking scheme, Purt [71] stated that “we aimed for simplicity, because not everyone working in fire protection has the time to do exercises in higher mathematics”. Nonetheless, when Law [72] was considering what constitutes a fire safety engineer, she brought up the point that a fire safety engineer “must understand how to measure and quantify fire phenomena and fire safety”.

This discussion is of interest, because indexing methods tend to remove most of the responsibility for defining adequacy (or otherwise) from the user to the developer of the method [73]. This raises a question of liability with regard to the responsibility for the use of a method. Does liability rest with the user, the developer, or some combination of both? This question is significant, because even when using prescriptive guidance there is the “perception in the construction community that achieving fire safety is more about achieving ‘ticks in boxes’ rather than a clear motivation to ensure a coherent and balanced engineered design” [69].

One could argue that an indexing method is a (possible) heuristic solution to that lack of competence. Nevertheless, it is vital to acknowledge that “heuristics may not be appropriate if used outside of the intended set of scenarios, placing greater importance on the pattern matching capability of the individual and the quality of the information available” otherwise systematic errors can unintentionally be produced [3].

From the review of the existing methods [74], it was found that competence is removed from the individual in some cases, since virtually anyone could conduct an evaluation by assigning grades (correct or not) and come up with a score. It is therefore up to the developer to either carry responsibility for, or disclaim, the results of the method’s application.

Most developers of indexing methods made the assumption that the evaluators will be capable and competent people, in the hope that they will use the indexing method as a tool to guide an assessment, and not as proof that a random (and possibly irrational) design passes a predetermined score. For example, the FSES for health care facilities [8] had four different parameter checks, because the developers accounted for the possibility of disguising a critical element's absence when rewarding redundant systems.

Similarly, in another, hierarchical model, Shields [31] formulated the concept of component basal norm scores, in order to identify components that fall to an unacceptable minimum that cannot be compensated by other better graded features, thus defining a more strict domain of equivalency which limited the freedom of the evaluator to misuse the method.

In each case, there appears to be an inherent tension within the development of any method. Complexity is often added to methods in an attempt to prevent the misuse of the index. However, the introduction of such complexity risks enabling use of the method by those who may not fit the developer's definition of a 'competent' person – thereby increasing the potential for any method to be misused or regarded as a fire safety 'panacea'.

4.4 Remarks on Expertise

Ramachandran and Charters [48] elaborated extensively on the merits and demerits of point schemes, focusing on demerits. They criticised the reliance on expert judgement, because, amongst other reasons, "there is room for argument and serious disagreement between people in the determination of points or ratings for different factors enhancing or reducing fire risk".

This, however, is the case for rule-based guidance as well since "it is recognised that much of the technical content of fire regulations has been introduced on the advice of committees or groups of experts. [Expert judgement] is the principle vehicle for the production of fire regulations, codes and standards, based on the consensus of committees" [31]. Similarly, Magnusson et al. [70] stated that for prescribed standards and regulations "in most cases, the methodology does not arise from scientific principles, but from a consensus process based on technical judgement and experience”. However, this common aspect with indexing has not been a cause for such criticism on prescriptive guidance in the literature.

Quintiere [75], in the preface of his book “Fundamentals of Fire Phenomena” stated that “standards have been generally established by committees under public consensus, albeit with special interests, and their shortcomings are not understood by the general public at large” and that “those that have expertise in fire standards readily know that the standards have little, if any, technical bases”. However, Shields [31] argued that if regulations have so far been successful, when properly implemented, then the opinions of experts are vindicated and somehow valid. In this respect, it is important with the strong warning provided by, Spinardi et al. [76], namely that “the fact that major fire disasters are rare does not mean that buildings are inherently safe from fire; it may rather mean that latent weaknesses can lie dormant for many years until a particular chain of events occurs”. So no definitive answer can be formulated on the issue of expert judgement.

5 Conclusions

Researchers and practitioners have developed dozens of fire risk indexing schemes since the 1980s. These schemes have proved to be a potentially useful approach under certain circumstances. Although each index is different, there are common components that every method must have, and common decisions that the developer(s) of each method has to make. In summary, the fire safety attributes to be evaluated must be recognised and chosen, their relative weighting must be defined, and a final index calculation procedure has to be orchestrated.

In reviewing the literature on this topic key tensions emerged about how, in order to tackle the perceived failings of simplified methods, additional layers of complexity are added that can address those failings. These layers increase the sophistication of the methods' workings and it has been found that this increased sophistication makes it harder for developers (or users) to assess the efficacy of a method. Sophistication also distances the user from the mechanics of the method they employ, sometimes leading users to perceive a fire risk index as a panacea with an 'infallible' result.

The competence of the users of fire risk indices is not adequately defined in the field, yet it has shown that this can have a significant impact on the employment and results of a method. The first impression is that FRI can allow people with limited knowledge to conduct evaluations, however this raises troubling questions about who is responsible for the output of the method. Do the method's designers retain some level of responsibility for how their methods are applied, or can they shift all responsibility on to the user?

Finally, it has been found that during the design of a method there is a similar lack of definition about what an 'expert' is, how to select a group of 'experts', and how to guarantee objectivity in procedures of expert elicitation.

It is concluded that at each stage of the design and implementation of a fire risk index there are a series of decisions to be made by the method's developers. It is unavoidable that these decisions will have an impact on the results of a method – and it seems impossible for the developer to remove themselves from the process in order to achieve a method that is 'independent' from its creating mind. Method developers must choose how they will select the attributes to be evaluated, and if those attributes should have a different relative importance. They must then decide how this relative importance will be quantified, and how they will treat any interaction between the attributes they chose. They must decide whether 'experts' are to be employed and decide how to select these experts. Who can be ‘pressed into service’? Finally, they must make a statement about the necessary competence of the people using the method.

Through reviewing the implementation of existing methods, it is considered feasible that FRI can provide the metrics to build a method that assesses aspects of fire resilience in building design. The corresponding attributes could be recognised, their relative weightings could be quantified if the developers find utility in them, and then a final score could be calculated. However, each of these decisions requires some level of compromise to be made: a complex method lacks transparency, can be perceived as a panacea, but can compensate for competency deficiency in the user; a simple method is transparent, can easily give an erroneous and non-repeatable result, but is less likely to be mistaken for an infallible multi-objective fire safety assessment method. If the method can be used by an unregulated user group, it is highly beneficial for it to remain simple. If there is some level of control (and trust) on the competence (and ethics) of the end user, additional layers of complexity to improve the accuracy could be a sensible step. In summary, though, it is impossible for the developers to cater for all scenarios and cases – thus the value to be found in a method’s transparency should not be underestimated. A relatively simple index can allow users to comprehend the mechanics of the method and thus recognise its shortcomings – and avoid using it as a panacea. If a method’s shortcomings are obvious to an engaged user then this serves to demonstrate that the individual has thought critically about the various fire safety attributes and their potential interactions within a building. In making such a critique, a user would have exhibited a level of competence that might justify their use of a more complex (and accurate) method. Ironically, if a user does not recognise the shortcomings of a method, then this is, perhaps, an indication that they are not competent to use it. Ultimately, it is up to the method’s developers to decide which aspects of a method they value most – and knowingly accept the compromise of this choice.

6 Acknowledgements

The authors gratefully acknowledge the support of BRE Trust for this project.

7 References

[1] J. M. Watts, ‘Fire Risk Rating Schedules’, in *Fire Hazard and Fire Risk Assessment*, Philadelphia: American Society for Testing and Materials, 1992, pp. 24–34.

[2] S. P. Hey, ‘Heuristics and Meta-heuristics in Scientific Judgement’, *Br. J. Philos. Sci.*, vol. 67, no. 2, pp. 471–495, Jun. 2016, doi: 10.1093/bjps/axu045.

[3] M. Kinsey, M. Kinateder, and S. Gwynne, ‘Burning Biases: Mitigating Cognitive Biases In Fire Engineering’, in *Interlam 2019*, Royal Holloway, University of London, UK, Jul. 2019, vol. 2, pp. 1882–1890.

[4] J. M. Watts, ‘Index Approach to Quantifying Fire Risk’, in *SFPE Symposium on Risk, Uncertainty, and Reliability in Fire Protection Engineering*, Bethesda, MD, May 1999, pp. 39–41.

[5] D. Rasbash, G. Ramachandran, B. Kandola, J. Watts, and M. Law, ‘Chapter 13 - Point Systems - A Single Index’, in *Evaluation of fire safety*, John Wiley & Sons, Ltd, 2004, pp. 363–390.

[6] J. M. Watts, ‘Chapter 82 - Fire Risk Indexing’, in *SFPE Handbook of Fire Protection Engineering*, Fifth Edition, Ed. Morgan J. Hurley., Springer, New York, NY, 2016, pp. 3158–3185.

[7] J. M. Watts, ‘Fire Risk Assessment Using Multiattribute Evaluation’, in *IAFFS 5th International Symposium*, 1997, vol. 5, pp. 679–690, doi: 10.3801/IAFSS.FSS.5-679.

[8] H. E. Nelson and A. J. Shibe, ‘NBSIR 78-1555: A system for fire safety evaluation of Health Care Facilities’, Center for Fire Research, National Engineering Laboratory, National Bureau of Standards, U.S. Department of Commerce Washington, D.C. 20234, Washington, D.C., Internal Report, Nov. 1978.

[9] E. W. Marchant, ‘Fire Safety Evaluation (Points) Scheme for Patient Areas within Hospitals: A Report on its Origins and Development Sponsored by the DHSS’, University of Edinburgh, Department of Fire Safety Engineering, Jun. 1982.

[10] B. Karlsson, B. Tomasson, and V. Stenstad, ‘Using a Fire Risk Index Method for Multi-Storey Apartment Buildings in the Design Process.’, in *Reaching New Horizons. Performance-Based Codes and Fire Safety Design Methods, 6th International Conference.*, Tokyo, Japan, Jun. 2006, pp. 1–12.

[11] L. T. Wong and S. W. Lau, ‘A Fire Safety Evaluation System for Prioritizing Fire Improvements in Old High-rise Buildings in Hong Kong’, *Fire Technol.*, vol. 43, no. 3, pp. 233–249, Sep. 2007, doi: 10.1007/s10694-007-0014-8.

[12] National Fire Protection Association, ‘NFPA 101A - Guide on Alternative Approaches to Life Safety’. National Fire Protection Association, 2019, [Online]. Available: https://www.nfpa.org/codes-and-standards/all-codes-and-standards/list-of-codes-and-standards/detail?code=101A.

[13] J. Shields and G. Silcock, ‘An application of the hierarchical approach to fire safety’, *Fire Saf. J.*, vol. 11, no. 3, pp. 235–242, Dec. 1986, doi: 10.1016/0379-7112(86)90066-4.

[14] T. J. Shields and G. W. H. Silcock, ‘Chapter 13 - Fire Safety Evaluation’, in *Buildings and Fire*, 1st ed., Longman Scientific & Technical, 1987.

[15] J. M. Watts, ‘Criteria For Fire Risk Ranking’, *Fire Saf. Sci.*, vol. 3, pp. 457–466, 1991, doi: 10.3801/IAFSS.FSS.3-457.

[16] D. Rasbash, G. Ramachandran, B. Kandola, J. Watts, and M. Law, *Evaluation of fire safety*. 2004.

[17] J. M. Watts and M. E. Kaplan, ‘Fire Risk Index for Historic Buildings’, *Fire Technol.*, vol. 37, no. 2, pp. 167–180, Apr. 2001, doi: 10.1023/A:1011649802894.

[18] J. C. Valente, ‘The Evaluation of the Fire Safety Level in a Historic Centre in Lisbon’, in *International Conference on Fire Protection of Cultural Heritage*, Thessaloniki, Greece, Jun. 2000, pp. 225–234.

[19] P. A. Hernandez-Leal, M. Arbelo, and A. Gonzalez-Calvo, ‘Fire risk assessment using satellite data’, *Adv. Space Res.*, vol. 37, no. 4, pp. 741–746, Jan. 2006, doi: 10.1016/j.asr.2004.12.053.

[20] National Fire Protection Association, ‘NFPA 1144 - Standard for Reducing Structure Ignition Hazards from Wildland Fire’. National Fire Protection Association, 2018, [Online]. Available: https://www.nfpa.org/codes-and-standards/all-codes-and-standards/list-of-codes-and-standards/detail?code=1144.

[21] T. Jarnskjold, G. Jensen, and A. Knudsen, ‘Performance profiles of exterior fire protective building envelopes’, *MATEC Web Conf.*, vol. 46, pp. 04005–10, May 2016, doi: 10.1051/matecconf/20164604005.

[22] R. E. Chapman and W. G. Hall, ‘Code compliance at lower costs: A mathematical programming approach’, *Fire Technol.*, vol. 18, no. 1, pp. 77–89, Feb. 1982, doi: 10.1007/BF02993490.

[23] Joint Committee of the building research board of the department of scientific & industrial research and The fire office’s committee, ‘Fire Grading of Buildings - Part I - General Principles and structural precautions’, HMSO - Ministry of Works, London, Post-war building studies 20, 1946.

[24] J. M. Watts, ‘A theoretical rationalization of a goal-oriented systems approach to building fire safety’, PhD Thesis, University of Massachusetts, 1978.

[25] H. L. Malhotra, *Fire safety in buildings: report of a study for the Fire Research Station on behalf of the Department of the Environment.* Borehamwood, England: Building Research Establishment, 1987.

[26] H. L. Malhotra, ‘Fire safety in assembly buildings’, *Fire Saf. J.*, vol. 7, no. 3, pp. 285–291, Jan. 1984, doi: 10.1016/0379-7112(84)90026-2.

[27] A. G. Copping, ‘Fire safety evaluation of ecclesiastical estate: the development and application of a fire safety evaluation procedure for the property protection of parish churches’, PhD Thesis, De Montfort University, Leicester, 2000.

[28] T. J. Shields and G. W. Silcock, ‘Comments on “The confusion existing in the use of the terms ‘model’ and ‘scheme’ in fire safety modelling”’, *Fire Saf. J.*, vol. 10, no. 3, pp. 239–240, May 1986, doi: 10.1016/0379-7112(86)90020-2.

[29] M. F. M. Idris, ‘The development of a fire safety evaluation procedure for the educational establishment’, PhD Thesis, The University of Edinburgh, Edinburgh, Scotland, 1997.

[30] A. N. Beard, ‘Towards A Systemic Approach To Fire Safety’, *Fire Saf. Sci.*, pp. 943–952, 1986, doi: 10.3801/IAFSS.FSS.1-943.

[31] T. J. Shields, ‘A fire safety evaluation points scheme for dwellings’, PhD Thesis, University of Ulster, Jordanstown, 1991.

[32] J. M. Watts, ‘Fire-Risk Indexing: A Systemic Approach to Building-Code “Equivalency” for Historic Buildings’, *APT Bull.*, vol. 34, no. 4, pp. 23–28, 2003, doi: 10.2307/1504867.

[33] L. L. Parks, B. D. Kushler, M. J. Serapiglia, L. A. McKenna, E. K. Budnick, and J. M. Watts, ‘Fire Risk Assessment for Telecommunications Central Offices’, *Fire Technol.*, vol. 34, no. 2, pp. 156–176, May 1998, doi: 10.1023/A:1015329519575.

[34] B. Karlsson and D. Larsson, ‘Using a Delphi Panel for Developing a Fire Risk Index Method for Multistorey Apartment Buildings’, Department of Fire Safety Engineering and Systems Safety, Lund University, Lund, Sweden, 3114, 2000.

[35] M. Gretener, ‘Attempt to calculate the fire risk of industrial and other objects (Versuch zur Rechnerischen Bestimmung Der Brandgefährdung von Industrie- und Anderen Objekten)’, in *3e Internationaal Brandweer Symposium = 3. Internationales Brandschutzseminar = 3rd International Fire Protection Seminar*, Eindhoven, The Netherlands, Oct. 1968, pp. 56–88.

[36] Chow W. K., ‘Proposed Fire Safety Ranking System EB-FSRS for Existing High-Rise Nonresidential Buildings in Hong Kong’, *J. Archit. Eng.*, vol. 8, no. 4, pp. 116–124, Dec. 2002, doi: 10.1061/(ASCE)1076-0431(2002)8:4(116).

[37] E. Soja, ‘A scheme for the evaluation of fire safety components involving the interactions between components, for hospital fire safety. Appendix 5 of the report “Fire Safety

Evaluation (Points) Scheme for Patient Areas within Hospitals" by E.W. Marchant', University of Edinburgh, Department of Fire Safety Engineering, Jun. 1982.

[38] H. A. Donegan, 'Chapter 77 - Decision Analysis', in *SFPE Handbook of Fire Protection Engineering*, Fifth Edition, Ed. Morgan J. Hurley., Springer, New York, NY, 2016, pp. 3048–3072.

[39] S. E. Magnusson and T. Rantatalo, 'Risk Assessment of Timberframe Multistorey Apartment Buildings. Proposal for a Comprehensive Fire Safety Evaluation Procedure', Lund University, Lund, Sweden, Internal Report 7004, 1998.

[40] P. Stollard, 'The development of a points scheme to assess fire safety in hospitals', *Fire Saf. J.*, vol. 7, no. 2, pp. 145–153, Jan. 1984, doi: 10.1016/0379-7112(84)90034-1.

[41] F. J. Dodd and H. A. Donegan, 'Prioritisation methodologies in fire safety evaluation', *Fire Technol.*, vol. 30, no. 2, pp. 232–249, May 1994, doi: 10.1007/BF01040004.

[42] R. W. Saaty, 'The analytic hierarchy process—what it is and how it is used', *Math. Model.*, vol. 9, no. 3, pp. 161–176, Jan. 1987, doi: 10.1016/0270-0255(87)90473-8.

[43] T. L. Saaty and M. S. Ozdemir, 'Why the magic number seven plus or minus two', *Math. Comput. Model.*, vol. 38, no. 3, pp. 233–244, Aug. 2003, doi: 10.1016/S0895-7177(03)90083-5.

[44] G. A. Miller, 'The magical number seven, plus or minus two: some limits on our capacity for processing information', *Psychol. Rev.*, vol. 63, no. 2, pp. 81–97, 1956, doi: 10.1037/h0043158.

[45] J. M. Watts, 'Fuzzy fire safety', *Fire Technol.*, vol. 31, no. 3, pp. 193–194, Aug. 1995, doi: 10.1007/BF01039190.

[46] S. M. Lo, B. Q. Hu, M. Liu, and K. K. Yuen, ‘On the Use of Reliability Interval Method and Grey Relational Model for Fire Safety Ranking of Existing Buildings’, *Fire Technol.*, vol. 41, no. 4, pp. 255–270, Oct. 2005, doi: 10.1007/s10694-005-3732-9.

[47] S. Sugahara, ‘Basic Concepts into Methodologies on Systems Approach to Building Fire Safety’, *Fire Sci. Technol.*, vol. 7, no. 2, pp. 35–41, 1987, doi: 10.3210/fst.7.2_35.

[48] G. Ramachandran and D. A. Charters, ‘Chapter 2 - Qualitative and semi-quantitative risk assessment techniques’, in *Quantitative risk assessment in fire safety*, Milton Park, Abingdon, Oxon ; New York : Spon Press, 2011, pp. 34–69.

[49] ‘WG6 - Fire Risk Assessment Methods’, European Commission, Final Report FiRE-TECH/WG6/01/iBMB/30/10/2003, 2003. [Online]. Available: http://www.framemethod.net/indexen_html_files/wg6finalreport.pdf.

[50] H. Hultquist and B. Karlsson, ‘Evaluation of a Fire Risk Index Method for Multistorey Apartment Buildings’, Department of Fire Safety Engineering and Systems Safety, Lund University, Lund, Sweden, 3088, 2000. [Online]. Available: https://lucris.lub.lu.se/ws/files/4725595/1266574.

[51] M. Bullock and A. Monaghan, ‘Shouldering the responsibility’, *Int. Fire Prof.*, no. 9, pp. 29–32, Jul. 2014.

[52] T. J. Shields, G. W. H. Silcock, H. A. Donegan, and Y. A. Bell, ‘Methodological problems associated with the Use of the Delphi technique’, *Fire Technol.*, vol. 23, no. 3, pp. 175–185, Aug. 1987, doi: 10.1007/BF01036934.

[53] H. A. Linstone and M. Turoff, Eds., *The Delphi method: techniques and applications*. Reading, Mass: Addison-Wesley Pub. Co., Advanced Book Program, 2002.

[54] N. Dalkey and O. Helmer, ‘An Experimental Application of the Delphi Method to the Use of Experts’, *Manag. Sci.*, vol. 9, no. 3, pp. 458–467, 1963.

[55] T. J. Shields, G. W. Silcock, and Y. Bell, ‘Fire safety evaluation of dwellings’, *Fire Saf. J.*, vol. 10, no. 1, pp. 29–36, Jan. 1986, doi: 10.1016/0379-7112(86)90029-9.

[56] P. Stollard, K. Croucher, and T. Clark, ‘The Development of a Points System to Compare the Potential Running Costs of Different Development Options for New Hospitals’:, *Health Serv. Manage. Res.*, vol. 3, no. 1, pp. 49–58, Mar. 1990, doi: 10.1177/095148489000300106.

[57] E. W. Marchant, ‘Methodological problems associated with the use of the Delphi technique — some comments’, *Fire Technol.*, vol. 24, no. 1, pp. 59–62, Feb. 1988, doi: 10.1007/BF01039641.

[58] I. Ioannou, W. Aspinall, D. Rush, L. Bisby, and T. Rossetto, ‘Expert judgment-based fragility assessment of reinforced concrete buildings exposed to fire’, *Reliab. Eng. Syst. Saf.*, vol. 167, pp. 105–127, Nov. 2017, doi: 10.1016/j.ress.2017.05.011.

[59] T. Z. Harmathy, ‘The Delphi method—a complement to research’, *Fire Mater.*, vol. 6, no. 2, pp. 76–79, Jun. 1982, doi: 10.1002/fam.810060208.

[60] T. Z. Harmathy *et al.*, ‘A Decision logic for trading between fire safety measures’, *Fire Mater.*, vol. 14, no. 1, pp. 1–10, Mar. 1989, doi: 10.1002/fam.810140102.

[61] S. J. van Zolingen and C. A. Klaassen, ‘Selection processes in a Delphi study about key qualifications in Senior Secondary Vocational Education’, *Technol. Forecast. Soc. Change*, vol. 70, no. 4, pp. 317–340, May 2003, doi: 10.1016/S0040-1625(02)00202-0.

[62] A. Kaplan, A. L. Skogstad, and M. A. Girshick, ‘The Prediction of Social and Technological Events’, *Public Opin. Q.*, vol. 14, no. 1, pp. 93–110, 1950, doi: 10.1086/266153.

[63] J. Kruger and D. Dunning, ‘Unskilled and unaware of it: how difficulties in recognizing one’s own incompetence lead to inflated self-assessments’, *J. Pers. Soc. Psychol.*, vol. 77, no. 6, pp. 1121–1134, Dec. 1999, doi: 10.1037/0022-3514.77.6.1121.

[64] G. Rowe, G. Wright, and F. Bolger, ‘Delphi: A reevaluation of research and theory’, *Technol. Forecast. Soc. Change*, vol. 39, no. 3, pp. 235–251, May 1991, doi: 10.1016/0040-1625(91)90039-I.

[65] S. M. Lo, ‘A Fire Safety Assessment System for Existing Buildings’, *Fire Technol.*, vol. 35, no. 2, pp. 131–152, May 1999, doi: 10.1023/A:1015463821818.

[66] S. M. Lo, K. C. Lam, and R. K. K. Yuen, ‘Views of building surveyors and building services engineers on priority setting of fire safety attributes for building maintenance’, *Facilities*, Dec. 2000, doi: 10.1108/02632770010328153.

[67] M. N. Ibrahim, M. S. Ibrahim, A. Mohd-Din, K. Abdul-Hamid, R. M. Yunus, and M. R. Yahya, ‘Fire Risk Assessment of Heritage Building – Perspectives of Regulatory Authority, Restorer and Building Stakeholder’, *Procedia Eng.*, vol. 20, pp. 325–328, Jan. 2011, doi: 10.1016/j.proeng.2011.11.173.

[68] J. Abrahams and P. Stollard, *Fire from First Principles : A Design Guide to Building Fire Safety*, 3rd ed. Routledge, 2003.

[69] M. Bullock and A. Monaghan, ‘Code compliance or fire engineering for life safety design - have we moved on?’, *Int. Fire Prof.*, no. 8, pp. 25–28, Apr. 2014.

[70] S. E. Magnusson *et al.*, ‘A proposal for a model curriculum in fire safety engineering’, *Fire Saf. J.*, vol. 25, no. 1, pp. 1–88, Jul. 1995, doi: 10.1016/S0379-7112(95)00038-0.

[71] G. A. Purt, ‘The evaluation of the fire risk as a basis for planning automatic fire protection systems’, *Fire Technol.*, vol. 8, no. 4, pp. 291–300, Nov. 1972, doi: 10.1007/BF02590534.

[72] M. Law, ‘What is a fire engineer?’, *J. Appl. Fire Sci.*, vol. 1, no. 1, pp. 3–6, 1991 1990.

[73] D. Larsson, ‘Developing the Structure of a Fire Risk Index Method for Timber-frame Multistorey Apartment Buildings.’, Lund University, Lund, Sweden, 2000.

[74] V. Koutsomarkos, D. Rush, G. Jomaas, and A. Law, ‘Comparative Analysis Of Fire Indexing Methodologies’, in *Interlam 2019*, Royal Holloway, University of London, UK, Jul. 2019, vol. 2, pp. 1647–1659.

[75] J. G. Quintiere, *Fundamentals of fire phenomena*. Chichester, England: John Wiley, 2006.

[76] G. Spinardi, J. Baker, and L. Bisby, ‘Post construction fire safety regulation in England: shutting the door before the horse has bolted’, *Policy Pract. Health Saf.*, vol. 17, no. 2, pp. 133–145, Apr. 2019, doi: 10.1080/14773996.2019.1591040.