Maturity Models and Safety Culture: A Critical Review

Anastacio Pinto Goncalves Filho[1*] and Patrick Waterson[2]

[1]Ministry of Labour,
Federal University of Bahia,
Salvador/Bahia,
Brazil.

[2]Human Factors and Complex Systems Group,
Loughborough University Design School,
Loughborough University,
Loughborough, LE11 3TU,
United Kingdom.

[*]**Address for correspondence:**
Anastacio Pinto Goncalves Filho,
Ministério do Trabalho,
Rua Ewerton Visco, 190,
Caminho das Árvores
Salvador – Bahia,
CEP 41.820.022,
Brazil.

---

[1] Email: anastaciofilho@ufba.br

Maturity Models and Safety Culture: A Critical Review

Abstract

The available evidence suggests that maturity models are a popular means of assessing safety culture in organisations. The aim of the present study was to review their conceptual underpinnings and roots, as well as provide details of how they have been used to assess safety culture (e.g., types of methods used, coverage of safety domains). A total of 41 publications were reviewed based on a set of selection criteria (e.g., studies which explicitly reported data or a case study which used a maturity model). The findings indicate steady growth in the use of maturity models to assess safety culture particularly within domains such as construction, the oil and gas industries and healthcare. We also found that most studies focus on providing a descriptive account of safety culture using maturity models and make limited attempts to assess the reliability/validity of outcomes from their use. We discuss the strengths and weaknesses of maturity models in the light of our findings, alongside identifying a number of new directions for future work of relevance to safety researchers and practitioners (e.g., the need for more detailed case studies of the use of maturity models to assess safety, as well as more attention to the underlying theory guiding use of maturity models).

Keywords: Safety culture, maturity model, safety culture assessment.

1.   Introduction

Some of the most compelling arguments that culture and safety might contribute to accidents and disasters were made in the late 1970s by Barry Turner in his pioneering work '*Man-Made Disasters*' (Turner, 1978; Pidgeon, 1988). Following the 1986 Chernobyl nuclear disaster the term 'safety culture' started to be regularly used amongst a broad community of safety scientists, psychologists and other groups (Silbey, 2009). There are a number of different explanations for the rise in interest in the construct of safety culture including increasing recognition of the importance of cultural aspects of health and safety management (Cooper, 2000; Cox and Cheyne, 2000; Flin et al., 2000; Reason, 1998) and the shift in the last few decades towards a focus on organisational factors governing risk and safety (Borys et al., 2009; Waterson et al., 2015; Robertson et al., 2016). As a result, many contemporary organisations strive to understand and improve their safety culture in order to deliver effective health and safety management and enhance their safety performance (Antonsen, 2009a; Reason, 1998, 2016).

At the same time, amongst researchers and academics, there have been a number of criticisms levelled at the construct of safety culture. Henriqson et al. (2014) for example, argue that the study safety culture encourages the view that safety is a widely shared norm, value or set of beliefs within organisations which masks important conflicts and disagreements which may exist amongst employees and managers. Others (e.g., Reiman and Rollenhagen, 2014; Dekker, 2018) suggest that a preoccupation with safety culture has shifted the focus away from more systemic accounts of the causes of accidents and encouraged a rather ore superficial account of how safety is related to system levels and other organisational dynamics (e.g., how safety culture changes over time). Finally, Antonsen (2009b) compared qualitative and quantitative descriptions of the safety culture in the same organisation (a Norwegian oil and gas platform) and found them to be dramatically different, leading him to cast doubt on the predictive validity of safety culture assessments. In the present paper we focus on a review of one popular tool or approach which is used to assess safety culture, namely maturity models. A later section of the paper discusses the findings from

our review in the light of contemporary criticisms of the safety culture construct, alongside a consideration of how maturity models fit within debates centred on research-practice gaps within safety science and human factors (Chung and Shorrock, 2011; Waterson, 2016).

2. Safety culture: some current challenges

2.1 Defining 'safety culture'

Despite the considerable literature covering theoretical and empirical aspects of safety culture (Antonsen, 2009; Cox and Flin, 1998; Diaz-Cabrera et al., 2007; Flin et al., 2000; Hopkins, 2006; Lee and Harrison, 2000; Mearns et al., 2009) there is still a lack of consensus and agreement about how to define the construct as well as assessment methods and on the overall structure of safety culture assessment (Mkrtchyan and Turcanu, 2012). The difference between safety climate and safety culture, for example, has been debated over decades by a number of safety researchers (Flin et al., 2000; Griffin and Curcuruto, 2016). Safety culture typically refers to the underlying assumptions and values that guide behaviour in organizations rather than the direct perceptions of individuals (Griffin and Curcuruto, 2016). Safety climate, by contrast, is sometimes regarded as the surface features of the safety culture discerned from the workforce's attitudes and perceptions at a given point in time (Flin et al., 2000). Andrew Hale (2000) refers to these and a range of other discussions centred on safety culture (e.g., the relation of culture to other aspects of safety management and behaviour) as examples of 'culture's confusions'. More recently, Hale stated: *"... safety culture is problematic in many of the same ways that 'accident proneness' was in the last century; in terms of its attributional consequences, the difficulties of defining it and the difficulties of deciding what you should measure as the outcome of its presence or absence; either accidents or other intermediate measures of safety"* (Waterson, 2017).

2.2 The theoretical status of safety culture

A number of authors have attempted to characterise the various theoretical approaches and methods which have been used to assess safety culture. Silbey (2009) for example, describes three dominant 'lenses' which characterise what she terms as 'talk about safety and culture'. The first 'lens', *'culture as causal*

attitude', view safety culture as something that is measureable and comprises the values, competencies, attitudes and behaviours about safety which exist within organisations. From this point of view culture "determine[s] the commitment to, and the style and proficiency of, an organisations' health and safety programs" (Silbey, 2009, p. 350 quoting Reason, 1997, p. 194). By contrast, the second 'lens', *'culture as engineered organisation'* whilst similarly focusing on the importance of cultural factors on safety outcomes, places more emphasis on how an organisation configures its processes and practices in order to improve safety, reliability and resilience. Proponents of the High Reliability Organisations (HROs) approach towards safety are viewed by Silbey (2009) as examples of the 'culture as engineered organisation' approach to safety culture (e.g., Eisenhardt, 1993; La Porte and Rochlin, 1994; Weick, 1987). A third 'lens' refers to *'culture as emergent and indeterminate'*. From this point of view, safety culture is understood to be socially constructed and mediated by artefacts and material, both mental and representational (Gherardi and Nicolini, 2000).

An alternative characterisation of safety culture and ways in which it is conceptualised and assessed which draws partly on Burrell and Morgan's (1979) analysis of sociological paradigms is provided by Guldenmund (2010, 2016). Guldenmund describes three approaches: (1) *interpretative or anthropological approaches* – these often treat culture as a system of meanings and symbols shared between groups of individuals who participate in this social process. Culture cannot be changed easily and cannot be assessed easily using scientific methods (Geertz, 1973; Alvesson, 2007; Martins, 1992). Qualitative methods, such as a narrative study, phenomenology, grounded theory, ethnography or case studies (Antonsen, 2009), or various combinations of these approaches, are methods used by an interpretative approach. Data collection such interviews, observational studies, document analysis are typically used to provide clues to underlying cultural assumptions (e.g., Scott et al., 2003); (2) *analytical or psychological approaches* – this is similar to Silbey's notion of *'culture as causal attitude'*, however, specifically relates to the use of questionnaires and to assess safety culture and the analysis of dimensions, factors and other statistical and psychometric properties of the survey instrument being used; (3) *pragmatic or experience-based approaches* – this approach focuses on the structure and

interactions within an organisation and the dynamic interplay between these which shapes and influences culture. The pragmatic approach also places emphasis on the types of processes that an organisation should have in order to achieve a mature or advanced status with regard to safety culture. These processes are reflected in Geller's approach towards Total Safety Culture (Geller, 1994) and safety culture maturity models such as the Shell Hearts and Minds programme (Hudson and Willekes, 2000; Hudson, 2007).

2.3 Aims, objectives and organisation of the current study

The focus of the current study is to outline the results of carrying a literature review on one particular approach towards safety culture, namely the use of maturity models for safety culture assessment. There is some evidence to suggest that maturity models are increasing in popularity (e.g., Fleming, 2001, 2017; Goncalves Filho et al., 2010; Parker et al., 2006; Health and Safety Technology and Management, 2017; Office of Rail and Road, 2017). Previous reviews have been carried out on the subject of maturity models and their use within domains such as software, management, business process management, information management and information technology management (Backer et al., 2009; Maier et al., 2012; Wendler, 2012). Little work however, has assessed the extent to which maturity models have been used to assess safety culture, as well as their scope and coverage. With this in mind, the specific aims of the review are:

1. To provide a better understanding of how maturity models to assess safety culture have been developed; their conceptual underpinnings and roots; the range of safety domains in which they have been applied; and, characteristics of their use;
2. To examine the methodological properties of maturity models to assess safety culture and the extent to which the outputs from using maturity models are evaluated (e.g., assessment of validity and reliability);
3. To use the outcomes from the review to offer some reflections on the theoretical status of the use of maturity models to assess safety culture and suggest new directions for future research and practice.

3. Maturity models and safety culture

3.1. Definition and scope

Maturity models involve defining maturity stages or levels which assess the completeness of the analysed objects, usually organizations or processes, via different sets of multi-dimensional criteria (Wendler, 2012; Becker et al., 2009). Hudson (2007) defines the use of maturity models in safety culture in terms of a continuum ranging from organisations that have unsafe cultures ('pathological' organisations) through to those who manage safety proactively ('generative' organisations) and those who are an intermediate stage of development ('bureaucratic' organisations). Organisations are seen as progress sequentially through the stages, by building on the strengths and removing the weaknesses of the previous levels (Fleming, 2001). A maturity model is a descriptive model in the sense that it describes essential, or key, attributes that would be expected to characterize an organization at a particular level.

The application of this concept is not limited to any particular domain (Wendler, 2012) and maturity models can be used both as an assessment tool and as an improvement tool (Maier et al., 2012). Focus groups, interviews, audits and checklists support maturity models in safety culture as well as questionnaires. Assessment can be also structured around a matrix or grid, where levels of maturity are allocated against key aspects of performance or key activities, thereby creating a series of cells. An important feature of this maturity matrix approach is that the cells contain descriptive text for the characteristic traits of performance at each level. One stated advantage of the use of a maturity matrix is that it is less complex than diagnostic and improvement tools (Maier et al., 2012). Based on the results of the analysis, recommendations for improvement measures can be derived and prioritized in order to reach higher maturity levels (Becker et al., 2009).

3.2 Origins

The use of maturity models as a tool to assess safety culture can be traced back to two main roots, namely, previous work on the 'Quality Management Maturity Grid' and Westrum's 'Typology of Organisations'.

3.2.1 Quality Management Maturity Grid

The Quality Management Maturity Grid (QMMG) was first proposed by Philip Crosby in 1979 (Wendler, 2012). In the QMMG, Crosby advocated that organizations go through five successive stages of quality maturity as they approach the maximum level of quality in all phases of organizational activity, these are: uncertainty, awakening, enlightenment, wisdom and certainty. In the first stage, management has no comprehension of quality as a management tool. The intermediate stages are characterized by a transformation in management understanding and attitude towards quality, how quality appears within an organization, how organizational problems are handled, the cost of quality as a percentage of sales, quality improvement actions taken by management, and how management summarizes the organization's quality problems. In the final stage of the QMMG, Total Quality Management (TQM) is viewed as an essential part of the company system (Calingo, 1996; Fraser et al., 2002). QMMG was adapted for process of building by the Carnegie-Mellon Software Engineering Institute (SEI) in 1986, where the concept of the Capability Maturity Model (CMM) was added and developed the foundation for its current use throughout industry (Paulk et al., 1993). The CMM was designed to guide software organizations in selecting process improvement strategies by determining current process maturity and identifying the few issues most critical to software quality and process improvement. It has five levels:

Initial: The software process is characterized as ad hoc, and occasionally even chaotic. Few processes are defined, and success depends on individual effort;
Repeatable: Basic project management processes are established to track cost, schedule, and functionality. The necessary process discipline is in place to repeat earlier successes on projects with similar applications;
Defined: The software process for both management and engineering activities is documented, standardized, and integrated into a standard software process for the

organization. All projects use an approved, tailored version of the organization's standard software process for developing and maintaining software;

Managed: Detailed measures of the software process and product quality are collected. Both the software process and products are quantitatively understood and controlled;

Optimising: Continuous process improvement is enabled by quantitative feedback from the process and from piloting innovative ideas and technologies. These five levels provide a framework for organizing these evolutionary steps into five maturity levels which are put forward as laying successive foundations for continuous process improvement (Paulk et al., 1993, figure 1).

Figure 1 about here

One of the first maturity model developed to assess safety was Dupont Bradley Curve which was examined by Fleming (2001) to develop a safety culture maturity model. The four stages in this model are: (i) reactive; (ii) dependent; (iii) independent; and, (iv) interdependent (Foster and Houst, 2013). In the first "reactive" stage people do not take responsibility. They believe that safety is more a matter of luck than management, and that "accidents will happen." In a "dependent culture" there is an emphasis on management and supervisory control, with a heavy focus on written rules and procedures. An "independent culture" where the focus is on a personal commitment to and responsibility for safety. The final stage is "interdependent" where there is a team commitment to safety with everyone having a sense of responsibility for safety beyond their own work and by caring for the safety of others.

3.2.2 Westrum's typology of organisational cultures

Typology of organisational was proposed by Ron Westrum (1993, 2004). Westrum's model distinguished three types of organisation – pathological, bureaucratic and generative. Table 1 explains these categories and provides an example how their characteristics apply within organisations. This model was extended from three to five stages in sequence, replacing the label 'bureaucratic' with 'calculative' and introducing two extra stages, the reactive and the proactive stages (Hudson, 2001, 2007 - figure 2). This was done in order to allow for more

subtle and accurate classification, and at the same time increasing the accessibility of the framework to industry employees by including terms they would be familiar with (Parker et al., 2006).

Table 1 and Figure 2 about here

4. Method

4.1 Search strategy and sources

A literature search, including publications in academic journals, conference papers and grey literature covering the period January 2000 - January 2017 was carried out using the following electronic databases: Google Scholar, Science Direct, UK Health and Safety Executive (HSE), Web of Science, British Library, Open Grey (information on grey literature in Europe), ProQuest (dissertations and theses within the UK and Ireland), Canadian Centre for Occupational Health and Safety (CCOHS, include OSHLINE and NIOSHTIC), NIOSHTIC-2 and Occupational Safety and Health Resource Database.

4.2 Search terms and exclusion/inclusion criteria

The following search terms were used: ''safety culture maturity'', ''safety culture maturity model'', ''safety culture AND maturity'', ''safety AND culture AND maturity'', ''framework and safety culture maturity'', ''safety culture maturity AND model''. The reference sections and bibliographies of articles were also searched. When publications were not found directly from databases, the authors were contacted and copies of their articles were requested. The publications were filtered using a set of inclusion and exclusion criteria. Publications where the development of a maturity model was described as a tool to assess safety culture or presented an application of a maturity model were included within the review. Similarly, publications where the dimensions used to describe and assess safety culture were presented were included. Publications where actual development and/or application of a maturity model were not described were rejected, as were those written in languages other than English.

4.3 Categorization framework

In order to provide a structure for the review, we categorised publications using a framework made up of the following nine elements:

(1) *Publication title, author(s) and year of publication*;

(2) *Source* (e.g., journal, thesis);

(3) *Country of origin*;

(4) *Application domain*;

(5) *Methods used to develop, evaluate and apply the model* – this involved identification different methods used to develop, evaluate and apply the maturity model (e.g., questionnaire, interview, literature review and focus group);

(6) *Type of publication* - in this case we made use of the distinction drawn by Wendler (2012) between: 'maturity model development' (defined as 'publications where the main objective was developing or constructing a new maturity model or adaptation of an existing maturity model'); 'maturity model application' (defined as 'publications where the main aim is the application of maturity models in several contexts or specific domains'); 'maturity model validation' (defined as 'publications where the main purpose is validating existing maturity models and includes empirical as well as conceptual validation, comparisons of maturity models and other approach and simulations'); and, 'maturity model reliability' (defined as 'publications where the main purpose is reliability testing of existing maturity models');

(7) Structure – this involved checking the structure of maturity models used in this study (e.g., number of level, description or summary of the characteristics of each level).

(8) *Level descriptor* – this involved examining the different types of descriptors for the maturity levels used in the study sample (e.g., 'pathological', 'uncommitted' and 'amoral calculators').

(9) *Aims and objectives of using the maturity model* – this category was included in order to capture the main drivers or intentions for using the maturity model (e.g., assessing overall safety management or risk perceptions; assessing safety leadership). In order to generate a set of codes which could be used to categorise the reasons for using the model, each author independently compiled a list of safety culture related keywords based on our individual reading of the study sample. The authors then met and agreed upon a final set of 16 keywords;

(10) *Maturity model reliability and validity* – this involved checking which of the study sample had carried out an assessment of mature model component reliability (e.g., by calculating values of internal reliability using Cronbach's).

The different types of validity checked as well as validity testing are described in table 2.

Table 2 about here

5. Findings

A total of 2026 publications were retrieved from the databases described in section 4.1. Duplicate publications were removed and a total of 135 (67 publications in academic journals, 68 reports and other publications in grey literature) were selected by reading the title, abstract and key words. The remaining 135 publications were filtered out using the set of criteria, described in section 4.2. A total of 41 publications were selected using the criteria for detailed review. Twenty-two publications were in academic journals and the remaining 19 were grey literature, including 5 reports. Ten publications were conference papers. In addition, the search yielded 4 theses (one PhD and three Masters Theses). Appendix A lists the complete set of review publications. Table 3 is summary of the results of applying the framework to the 41 studies. In subsequent Sections of the paper, we refer to the studies with reference to their number in Table 3 (e.g., article 6 is Gordon et al., 2007).

Table 3 about here

5.1 Year of publication, country of origin and application domain

In order to provide an overview of trends (e.g., growth, country of origin) in maturity models, the publications were analysed by their year of publication. The distribution over the last two decades (January 2000 - January 2017) is shown in Figure 3. Between 2000 and 2005, the number remained stable with at least one publication a year. Beginning in 2006, a steady rise is noticeable, reaching its peak with 7 publications in 2011. The country which published the most on maturity models was the United Kingdom with 8 publications (1, 8, 9, 12, 16, 24, 27 and 31) followed by Canada with 7 publications (3, 5, 7, 15, 18, 34 and 39). One maturity model was applied in two countries, Canada and United Kingdom (Kirk et al., 2007, 8).

Figure 3 about here

There is fairly uniform application across a range of domains, but the trends indicate areas of recent growth in the healthcare (3, 4, 7, 8, 9, 18, 20, 38, 39 and 41), oil and gas (1, 12, 13, 24, 26, 27, 28, 34 and 35) and the construction industry (10, 11, 25, 30, 31 and 37).

5.2 Methods used to develop, evaluate and apply the maturity model

A wide variety of methods were used in order to develop, evaluate and apply the maturity models in the sample (table 4). Accordingly, the most used method (30 in total) was where the developer sought information from previous literature in order to develop the maturity models. Some studies (1, 4, 5, 8, 12, 14, 15, 19, 31, 32 and 34) used literature mixed with other development methods, e.g., Parker et al., (2006, 1) used literature and interview and Kirk at al., (2007, 8) used literature review, interview and consultation with experts (where safety and risk management experts are consulted for comments and opinion on maturity model by developer).

Table 4 about here

Questionnaires were the most common method used to evaluate and apply the maturity model. Four publications used comparison to evaluate the maturity model, e.g., Gordon et al., (2007, 6) presented 2 maturity models, one maturity model to assess a safety culture by questionnaire and one to assess the safety culture by interview and both maturity models were evaluated their results by comparison. A combination of two or more methods (e.g., interviews and questionnaires), was used to apply 7 maturity models. Tappin et al, (2015, 2) for example, used semi-structured interviews combined with documentary analysis of occupational safety and health (OSH) systems in order to assess OSH maturity levels.

5.3 Type of publication

Table 5 shows the distribution of the four types of publications among the sample. Maturity model development is the most common type of publication

(33 out of 41), whereas the number of publication of maturity model application (4 out of 41), validation (2 out of 41) and reliability (2 out of 41) was smaller.

Table 5 about here

5.4 Structure

The maturity models have a conceptual overlap, such as aspect at discrete stages or maturity levels, with a description of characteristic performance at various levels of maturity. The following structure is common in maturity models to assess safety culture: (a) a number of levels (typically 5); (b) a descriptor for each level which are describe following section 5.5; (c) a generic description or summary of the characteristics of each level as a whole; (d) multi-dimensional; and, (f) a description of each activity as it might be performed at each maturity level.

5.5 Level descriptor

More than two-thirds of maturity models (68%, n=28) were developed based on Westrum's model and adopted pathological, reactive, calculative or bureaucratic, proactive, and generative or sustainable levels (table 6). Filho et al., (2010, 13) relabelled 'generative' as 'sustainable' in order to make it more understandable in their country. The 'bureaucratic' level was also retained in maturity models developed by Law et al., (2010, 3) and Filho et al., (2010, 13). Six maturity models were developed based on the CMM and adopted 'emerging', 'managing', 'involving', 'cooperating' and 'continually improving' levels. Eight maturity models adopted different levels from Westrum's model and CMM. Wright et al., (2012, 16) adopted the levels 'amoral calculators', 'dependent', 'doubters', 'proactive compliers' and 'leader' because these labels were seen as more suitable for the domain of food safety. Similarly, Jespersen (2016, 5) adopted the levels 'doubt', 'react', 'know', 'predict' and 'internalize' in their models. The maturity model developed by Lunt et al., (2011, 31) combined elements of the Westrum model and CMM in order to make it more suitable for use in the construction industry. Consequently, the names for each level were changed to 'starting blocks', 'getting going', 'walking', 'running' and 'sprinting', respectively.

Table 6 about here

Gordon et al., (2007) developed two maturity models in same publication (6), one based on CMM (6A) and other one based on the DNV (Det Norske Veritas) auditing system in which the levels adopted were 'uncertainty', 'core', 'extended', 'leading edge' and 'world class' (6B). Fleming and Meakin (2004, 15) adapted the maturity model from the model presented in the 'Changing Minds' toolkit (Step Change in Safety, 2017). Their model includes the following levels: 'documenting', 'controlling', 'engaging', 'participating' and 'institutionalising'. Mohamed and Chinda (2010,10) adopted the EFQM (European Foundation for Quality Management) Excellence model, consequently, their maturity model includes the following levels: 'uncommitted', 'drifters', 'improvers', 'award winners', and 'world-class'.

5.6 Aims and objectives of using the maturity model

The most common aim was general safety management assessment, which appeared in 33 (80%) publications followed by assessment of communication about safety (68%), management commitment to safety (68%) and safety training (66%). Assessing organisational learning accounted for more than half of the publications in our sample (56%), whilst accident/incident reporting (37%), safety prioritisation (37%), employee involvement (34%) and accident/incident analysis (34%) made up more than a third of the publications. Organisational trust (22%), safety audit (22%), periodic review (20%) represented approximately a fifth of the sample, followed by a smaller proportion of publications which focused on risk perceptions (17%), risk analysis (15%) safety leadership (15%) and work pressure (12%).

5.7 Maturity model reliability and validity

The 41 publications were specifically checked regarding evidence in terms of the assessment of reliability and validity were carried out (table 7). Eleven (27%) studies carried out some form of reliability assessment (e.g., where a maturity model had involved a questionnaire and assessed reliability using Cronbach's Alpha). Fourteen (34%) publications carried out both content validity (e.g., by assessing how close the maturity model aligned with other measures or indices of

safety such as an employee safety culture perception survey) and face validity (e.g., by asking employees to assess the extent to which the components of the model mapped on to aspects of safety in their organisation). Four publications carried out convergent validity (i.e., testing whether maturity model components supposed to be either related or unrelated, in fact, related or unrelated – Campbell and Fiske, 1959) and two carried out discriminant validity. Only one study (34) provided evidence about the procedure used to assess reliability and all four types of validity. Eighteen (44%) of the publications report no evidence that assessments of either reliability or validity were carried out.

Table 7 about here

6. Discussion

6.1. Growth and expansion of maturity models

Our findings demonstrate that over the last two decades there has been a steady growth in the use of maturity models to assess safety culture across a wide range of industries. The period 2008-2011 represents a peak of activity in using maturity models with steady use characterising the succeeding years (figure 3). Similarly, articles drawn from the Anglophone countries are the most frequent ones using maturity models, but with some evidence that their use is consistently wide in terms of other countries and global regions. In terms of application domain, construction, oil and gas and healthcare are the most common areas in which maturity models are applied. To some extent this might be explained by the fact that industries such as oil and gas have a well-established track record in using specific type of models and developing tools which have influenced other attempts to use maturity models (e.g., the Hearts and Minds toolkit – Hudson and Willekes, 2000). Likewise, the growth of maturity models in the healthcare sector might be attributable to the popularity of the Manchester Patient Safety Framework (MaPSaF) and its promotion up until 2012 within the UK by bodies such as the National Patient Safety Agency (NPSA), alongside other trends such as increased efforts to assess patient safety culture and drives to accelerate the integration of human factors within healthcare (Waterson, 2014; Waterson and Catchpole, 2016). Figure 4 illustrates the growth of specific types of tools associated with maturity models since the beginning of the millennium, as well

as application domains in which they are used and links back to earlier work in the 1970s, 1980s and 1990s (section 3.2).

Figure 4 about here

6.2 The theoretical status of maturity models

6.2.1 Maturity models as tools for organizational development

The range of methods used to develop, evaluate and apply maturity models, alongside the finding that most of the publications in the review sample fall into 'development' (i.e., where the main objective was to build a maturity model, rather than necessarily describe its application or evaluation - Wendler, 2012), would seem to provide some evidence maturity models illustrate what Guldenmund's (2010, 2016) characterized as a *'pragmatic or experience-based'* approach with regard to safety culture. This 'pragmatism' extends to numerous attempts which were made to change the various elements of the maturity models and the variety of level descriptors used within the study sample (table 6). Similarly, the fact that the main driver for using the model in most of our sample was assessment of general levels of safety and core elements of safety culture (e.g., communication, management commitment) as opposed to detailed aspects (e.g., levels of organizational trust, risk perceptions) might be seen to imply that maturity models are seen as ways of assessing the 'wider, bigger picture' of how safety might work in a specific organizational culture (c.f., 'total safety culture' - Geller, 1994). Finally, the low numbers of studies which report details of reliability and validity, might be seen to imply that the *'process'* of using a maturity model is more important that the actual *'outcome'*. From this point of view, maturity models might be construed as tools for 'organizational development', as compared to a means of 'measuring' safety culture, again reinforcing their 'pragmatic 'advantages as compared to survey instruments ('analytical/psychological' approaches – Guldenmund, 2010; 2016;, 'culture as causal attitude' – Silbey, 2009) and qualitative methods such as interviews or ethnographies ('interpretive/anthropological' approaches Guldenmund, 2010; 2016; 'culture as emergent/indeterminate' – Silbey, 2009).

6.2.2 Static vs. dynamic accounts of safety culture

Aside from the resource-intensive drawbacks (i.e., the time taken to organise discussion groups or workshops and analyse/evaluate results) of using maturity models, it may also be the case that maturity models may encourage a misleading picture of how safety culture operates in organisations. Maturity models imply that organisations make steady progress on a journey that takes them from low levels of safety towards an optimal state. This might in some ways be interpreted as a 'Whig-like' interpretation of cultural improvement[2], when in reality values and beliefs held by employees about safety may spontaneously wax, wane and in some cases relapse over short periods of time (Turner, 1978; Busch, 2016). Vincent and Amalberti (2016) for example, refer to safety as a 'moving target' as opposed to a fixed entity. Maturity models by their very nature may tend to overemphasise aspects of a static view of safety and may underplay the subtlety with which safety may dynamically vary according to a host of influencing factors (resources, workload, policy changes, economic and political influences – Rasmussen, 1997). Internal /external benchmarking of the outcomes of safety culture assessment may also play to hidden agendas within organisations and potential lead to manipulation and 'gaming' (e.g., the misuse of safety culture in 'league tables' - Shorrock, personal communication).

6.2.3 The absence of theory in maturity models

In common with other well-known frequently used constructs in safety science and human factors (e.g., the systems approach - Waterson, 2009), we found few attempts within our review sample to articulate what type of theory of safety culture was being used. Authors tended to offer a definition of safety culture, but made few attempts to link their work to theoretical frameworks and approaches which are common within safety science and human factors (e.g., the theory of High Reliability Organization (HRO) – Weick and Sutcliffe, 2007; Normal Accident Theory – Perrow, 1984). In many ways this criticism might be levelled at much of the literature on safety culture (Flin, 2007; Reiman and Rollenhagen, 2014), however the use of maturity models appears to imply that there is some mechanism underlying improvement and that particular components of cultural

---

[2] The term 'Whig' refers to a 17th Century political faction. A Whig view of history presents the past as an inevitable progression towards ever greater liberty and enlightenment.

development (e.g., 'pathological' or 'bureaucratic' states, stages of phases) which either predicts the nature and course of improvement or set out a steps on the road to improvement. In the absence of a theoretical basis it is difficult to see how this can be justified. Theories by their very nature are used to predict behavior and levels of safety, as well as supporting comparison with other, sometimes unrelated, phenomena (so-called 'disciplined imagination – Weick, 1989). Sutton and Shaw (1995) point to a number of misunderstandings about the use of theory in the social sciences including the view that data on its own does not constitute a contribution to theory. In many respects the findings from our review suggest that using a maturity model and reporting the outcomes from this use, does not in itself constitute a theory of safety. Studies which report the use of maturity models need to be more explicit in stating the type of theoretical stance (e.g., epistemology, ontology – Haavik, 2014) they are adopting and move away from a non-committal approach towards theory. Part of this involves providing more details of the process of change and the mechanisms which underpin safety improvement.

6.3. Maturity models –academic and practitioner perspectives

6.3.1 Maturity models as a 'cool medium'

In many respects the findings from our review of the use of maturity models to assess safety culture also touch on a number of debates which have surrounded safety culture, as well as the assessment of work-based attitudes and the wider world of safety over the past few decades. Many of these debates are framed in terms of binary distinctions and opposing positions. Wilson (1987) provides a typical example in his account of the various methods (e.g., questionnaire instruments, semi-structured discussions with employees) which are used to assess job attitudes and inform job redesign decisions. This time drawing on the work of Marshall McLuhan (1964), Wilson argues that *'formal job attitude instruments used alone are 'hot media', in the sense of being well-filled with data but low in participation. Informal discussion methods could be seen then as a 'cool medium', high in participation or completion by the audience'* (Wilson, 1987, p. 385). Wilson further argues that informal methods can encourage employees to develop 'ownership' of the outcomes from discussions and workshops (Shipley, 1987) as well as providing detailed insights into the reasons

why particular attitudes may be shared or dominant at one particular time. It might be argued that maturity models, as compared to other methods of measuring safety culture (e.g., questionnaires) fall under Wilson's category of 'cool media' and a similar set of trade-offs may be at play (e.g., promoting participation as compared to providing more scientific or 'objective' measures of safety culture). In some cases, it may be that from the safety practitioner's point of view it may be that in using a maturity the *process* may be more important that the actual *outcome*. How practitioners use maturity models, their goals and what they do with the results is something, alongside other aspects of the day-to-day work of the safety professionals, which is worthy of further attention (Hale, 2006).

6.3.2 Flexibility and 'bricolage'

An additional strength of maturity models might be said to be the fact that they are flexible and can be tailored to fit the specific requirements and objectives of organisations and their employees. Our findings as they relate to level descriptors for example (table 6), show that the types of maturity models used to assess safety culture varied a great deal. This flexibility and adaptability is in many ways similar to other methods within safety science (e.g., Rasmussen's work on the Accimap method of accident analysis and subsequent attempts to 'remix' his work – Waterson et al., 2017). This process might be seen as akin to what the anthropologist Claude Levi-Strauss (1966) called 'bricolage', that is *'the construction [e.g., an artefact, narrative, tool] or creation from a diverse range of available things'* (Concise Oxford English Dictionary, 12th Edition, 2011). The process of bricolage is also very much in keeping with the pragmatic approach towards safety culture assessment.

7. Conclusions, limitations and future work

In this paper we have attempted to review published literature which has made use of maturity models of one sort or another to assess safety culture over the last two decades or so. Our findings show on the one hand growth in terms of the use of maturity models to assess safety culture, but also significant variation in the ways in which they are used and reported within the literature. A key limitation of our work, which was difficult to avoid, is that we focused on published

research. Anecdotal evidence suggests that use of maturity models may be under-reported. Our work in a variety of sectors (e.g., oil and gas, healthcare) suggests that maturity models are widely used by organisations eager to gain some insight into their safety culture. Much of this activity might be characterised as 'experimenting' with maturity models, particularly as a result of disappointing efforts using other methods for measuring safety culture (e.g., low response rates from questionnaire studies). Many of these 'experiments' appear as internal reports or other forms of informal feedback within these organisations and hence do not appear in published form. Maturity models, we would argue, focus on the *process* aspects of safety culture and in line with other areas of safety science there is a need for more case study or 'experienced-based' accounts of how they are used and what sorts of organisational outcomes arise from their usage and how this might change and develop over time (Hale, 2006):

*'Maturity scales … represent the steps leading from the pathological state or face to the smiling, generative one, but we know little or nothing about whether it is indeed possible to mount that scale, and if so how. We have not defined whether we should be trying to shift companies all the way from one to another end of the scale, or that we would be happy if the bulk of companies made it to the halfway point and became 'calculative'. Longitudinal research studies of companies to plot such shifts and how to facilitate them are desperately needed. We might also need to expand our discipline base to do it.' (*Hale, 2006)

In terms of the practice of safety culture assessment, would also argue that the decision to use a maturity model to assess safety culture instead of other type of methods (e.g., focus groups, questionnaire surveys), should not be seen in terms of a straightforward question of either/or. Rather, we would encourage potential users of maturity models to view them pragmatically and to acknowledge the value of using them in combination with other methods (see for example Kirwan and Shorrock, 2014, for a case study description of how this applies). Finally, the strengths and weaknesses of maturity models and their application within the field of safety culture needs to be more widely acknowledged (e.g., compromises and trade-offs - Waterson et al., 2014; Amalberti, 2013).

From an academic or scientific point-of-view we would argue for greater attention be given to the theoretical aspects and justification for using maturity models. As noted in section 6.2.3 we would also suggest, in common with many

other authors, that more attention to theory is also a need for the field of safety culture more widely. The findings from our review showed that assessments of the reliability and validity of the use of maturity models to assess safety culture tend to be the exception, rather than the rule (44% of the sample did not report any evidence of reliability or validity assessments). A key weakness of the maturity model approach may be that results obtained during one point in time may not prove to be repeatable during another. The studies in our sample do not allow us to draw firm conclusions about the reliability, validity and overall robustness of using maturity models. Further investigation probing deeper into these factors is likely to be one area for future research. It seems likely however, that discussion of focus groups with different groups and/or members are unlikely to produce outcomes which are difficult to compare and hence very difficult to assess in terms of reliability.

We should also acknowledge, as mentioned in section 6.2 of the paper, that the flexibility afforded by maturity models may be a strength rather than a weakness as compared to other methods of safety culture assessment. We note in passing, however, that the reliability and validity of other types of maturity model has been the subject of some debate in the last few years (e.g., Bach's criticism of the Capability Maturity Model – Bach, 1994; Finkelstein, 1992). The question of validity (e.g., how well do maturity assessments correlate with safety outcome data) is however, possible to assess and, in common with other types of safety culture assessment (Flin et al., 2000), is something that should be pursued in the future, at least for those organisations (e.g., those who did carry out validity/reliability assessment) who wish to pursue this goal. In conclusion, we would say that our understanding of the scientific and practice-oriented aspects of maturity models remains relatively 'immature'. Much more effort needs to be given over in the future to some of the gaps in our current understanding which have been revealed by our literature review.

References

Alvesson, M., 2007. Understanding Organizational Culture. Sage, London.

Amalberti, R. 2013. Navigating Safety Necessary Compromises and Trade-Offs - Theory and Practice. Dordrecht: Springer.

Antonsen, S., 2009a. Safety Culture: Theory, Method and Improvement. Ashgate, England.

Antonsen, S., 2009b. Safety culture assessment: a mission impossible? Journal of Contingencies and Crisis Management, 17, 4, 242-254.

Bach, J. 1994. The immaturity of CMM. American Programmer, September, 7, 9, 13–18.

Becker, J., Knackstedt, R. and Pöppelbuß, J., 2009. Developing maturity models for IT management – a procedure model and its application. Business & Information System Engineering, 3, 213-222.

Borys, D., Else, D. and Leggett, S., 2009. The fifth age of safety: the adaptive age? Journal of Health and Safety, Research and Practice, 1, 19–27.

Burrell, G. and Morgan, G., 1979. Sociological Paradigms and Organizational Analysis. Aldershot: Ashgate.

Busch, C. 2016. Safety Myth 101. Mysen: Mind the Risk. (www.mindtherisk.com).

Calingo, L.M.R., 1996. The evolution of strategic quality management. International Journal of Quality & Reliability Management, 13, 9, 19 – 37.

Campbell, D.T., and Fiske, D.W. 1959. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 56, 81-105.

Chung, A.Z.Q. and Shorrock, S.T. 2011. The research-practice relationship in ergonomics and human factors – Surveying and bridging the gap. Ergonomics, 54, 413–429.

Cooper, D., 2000. Towards a model of safety culture. Safety Science, 36, 111-136.

Cox, S.J. and Flin, R., 1998. Safety culture - Philosophers' stone or man of straw? Work and Stress, 12, 3, 189-201.

Cox, S.J. and Cheyne, A. J. T., 2000. Assessing safety in offshore environments. Safety Science, 34, 215-257.

Crosby, P.B., 1979. Quality is Free. McGraw-Hill, New York.

Dekker, S.W.A., 2018. The Safety Anarchist. London: Routledge.

Díaz-Cabrera, D., Hernández-Fernaud, E. and Isla-Díaz, R., 2007. An evaluation of a new instrument to measure organisational safety culture values and practices. Accident Analysis & Prevention, 39, 6, 1202-1211.

Eisenhardt, K., 1993. High reliability organizations meet high velocity environments: common dilemma in nuclear power plants, aircraft carriers, and microcomputer firms. In K.H. Roberts (Ed.), New Challenges Understanding Organizations. New York: Macmillan.

Foster, P., Hoult, S., 2013. The safety journey: using a safety maturity model for safety planning and assurance in the uk coal mining industry. Minerals, 3, 59-72.

Finkelstein, A., 1992. Software Process Immaturity Model, SIGSOFT Software Engineering Notes, 1992.

Fleming, M., 2001. Safety culture maturity model. Offshore Technology Report 2000/049, Health and Safety Executive, UK.

Fleming, M. 2017, A practitioner's guide to creating a high performance safety culture. Keynote presentation at the 12th International Symposium on Human Factors in Organizational Design and Management (ODAM 2017), Banff, Canada, July 31st – August 3rd.

Flin R. 2007. Measuring safety culture in healthcare: A case for accurate diagnosis. Safety Science, 45, 6, 653-667.

Flin, R., Mearns, K, O'Connor, P. and Bryden, R., 2000. Measuring climate: identifying the common features. Safety Science, 34, 177-192.

Fraser, P., Moultrie, J. and Gregory, M., 2002. The use of maturity models/grids as a tool in assessing product development capability. In: IEEE International Engineering Management Conference, pp. 244-249.

Geller, E.S., 1994. Ten principles for achieving a Total Safety Culture. Professional Safety, September, 18-24.

Gherardi, S. and Nicolini, D. 2000. To transfer is to transform: the circulation of safety knowledge. Organization, 7, 329–48

Geertz, C. 1973. The Interpretation of Cultures. New York: Basic Books.

Goncalves Filho, A.P., Andrade, J.C.S. and Marinho, M.M.O., 2010. A safety culture maturity model for petrochemical companies in Brazil, Safety Science, 48, 615-624.

Griffin, M. A. and Curcuruto, M., 2016. Safety climate in organizations. Annual Review of Organizational Psychology and Organizational Behavior, 3, 191-2012.

Guldenmund, F.W., 2000.The nature of safety culture: a review of theory and research. Safety Science, 34, 193-214.

Guldenmund, F.W., 2010. Understanding and exploring safety culture. Thesis (PhD), Delft University, The Netherlands.

Guldenmund, F.W., 2016. Organizational safety culture. In S. Clarke, T.M. Probst, F. Guldenmund and J. Passmore (Eds.), The Wiley Blackwell Handbook of Occupational Safety and Workplace Health. Chichester: John Wiley.

Haavik, T.K. 2014. On the ontology of safety. Safety Science, 67, 37-43.

Hale, A.R., 2000. Culture's confusions. Safety Science, 34, 1-14.

Hale, A.R., 2006. Method in your madness: system in your safety. Rede, Delft University of Technology, 15th September. Retrieved from https://www.researchgate.net/publication/27351879_Method_in_your_madness_ system_in_your_safety (last accessed 15th March, 2017).

Health and Safety Technology and Management (Hastam). 2017. Safety Culture Benchmarking Conference, Loughborough University, September 6th 2017.

Henriqson, E., Schuler, B., van Winsen, R., Dekker, S.W.A. 2014. The constitution and effects of safety culture as an object in the discourse of accident prevention: A Foucauldian approach. Safety Science, 70, 465-476.

Hopkins, A., 2006. Study organizational cultures and their effects on safety. Safety Science, 44, 875-889.

Hudson, P. and Willekes, F.C., 2000. The Hearts and Minds Project in an Operating Company: developing tools to measure cultural factors. Paper presented at SPE International Conference on Health, Safety and Environment in Oil and Gas Exploration and Production, 26-28 June, Stavanger, Norway.

Hudson, P., 2001. Aviation safety culture. Safeskies, 1-23.

Hudson, P., 2007. Implementing a safety culture in a major multi-national. Safety Science, 45, 697–722.

Kirwan, B. and Shorrock, S., 2014. A view from elsewhere – safety culture in European air traffic management. In P.E. Waterson (Ed.), Patient Safety Culture. Farnham: Ashgate.

La Porte, T.R. and Rochlin, G., 1994. A rejoinder to Perrow. Journal of Contingencies and Crisis Management, 2: 221-227.

Lee, T. and Harrison, K., 2000. Assessing safety culture in nuclear power station. Safety Science, 34, 61-97.

Levi-Strauss, C., 1966. The Savage Mind. University of Chicago Press, Chicago, IL.

Mearns, K., Kirwan, B. and Kennedy, R.J., 2009. Developing a Safety Culture Measurement Toolkit (SCMT) for European ANSPs. Proceedings of the Eighth USA/Europe Air Traffic Management Research and Development Seminar.

Maier, A. M., Moultrie, J. and Clarkson, P. J., 2012. Assessing organizational capabilities: reviewing and guiding the development of maturity grids. IEEE Transactions on Engineering Management, 59, 1, 138-159.

Martins, J., 1992. Cultures in Organisations: Three Perspectives. Oxford University Press, New York.

McLuhan, M., 1964. <u>Understanding Media.</u> London: Routledge.

Mkrtchyan, L. and Turcanu, C., 2012. Safety culture assessment tools in nuclear and non-nuclear domains. Review of safety culture tools. <u>Report of The Belgian Nuclear Research Centre</u>. Nuclear Science and Technology Studies (NST), Belgium.

Office of Rail and Road (ORR). 2017. <u>RM^3 – The Risk Management Maturity Model</u>. Report available from: http://orr.gov.uk/rail/health-and-safety/health-and-safety-strategy/risk-management-maturity-model-rm3 (last accessed 22nd Seprmeber 2017).

Parker, D., Lawrie, M. and Hudson, P., 2006. A framework for understand the development of organizational safety culture. <u>Safety Science</u>, 44, 551-562.

Paulk, M.C., Chrissis, C. and Weber, M.B., 1993. <u>Capability Maturity Model for Software, Version 1.1.</u> <u>Software Engineering Institute</u>, Carnegie-Mellon, Pennsylvania.

Perrow, C. 1984. <u>Normal Accidents</u>. Princeton: Princeton University Press.

Pidgeon, N. (1988), Shaking the kaleidoscope of disasters research – a reply. <u>Journal of Contingencies and Crisis Management</u>, 6, 97-101.

Rasmussen, J., 1997. Risk management in a dynamic society: a modelling problem. <u>Safety Science</u>, 27, 183-213.

Reason, J., 1997. <u>Managing the Risks of Organizational Accidents</u>. Aldershot: Ashgate.

Reason, J., 1998. Achieving a safe culture: theory and practice. <u>Work & Stress</u>, 12, 3, 293-306.

Reason, J. 2016. <u>Organizational Accidents Revisited</u>. Farnham: Ashgate.

Reiman, T. and Rollenhagen, C. 2014. Does the concept of safety culture help or hinder systems thinking in safety? <u>Accident Analysis and Prevention</u>, 68, 5-15.

Robertson, K., Black, J. Grand-Clement, S. and Hall, A., 2016. Human and Organisational Factors in Major Accidents. RAND Europe Report. Available at: http://www.rand.org/randeurope/research/projects/major-accident-prevention-factors.html (last accessed February 27th 2017).

Scott, T., Mannion, R., Davies, H. and Martin M., 2003. The quantitative measurement of organizational culture in health care: A review of the available instruments. <u>Health Services Research</u>, 38, 3, 923-945.

Shipley, P., 1987. The methodology of applied ergonomics: validity and value. In J.R. Wilson, E.N. Corlett and I. Maneca (Eds.), <u>New Methods in Applied Ergonomics</u>. London: Taylor and Francis.

Silbey, S. S., 2009. Taming Prometheus: Talk about safety and culture. <u>Annual Review of Sociology</u>, 35, 341 - 369.

Step Change in Safety, 2017. Changing Minds Guidance. Available from: https://www.stepchangeinsafety.net/safety-resources/publications/changing-minds-guidance (Last accessed February 27th, 2017).

Sutton, R.I. and Staw, B.M. 1995. What theory is *not*. <u>Administrative Science Quarterly</u>, 40, 371-384.

Trochim, W. M. K., 2006. Research Methods Knowledge Base, 2nd Edition. Internet WWW page, at URL: <http://www.socialresearchmethods.net/kb/> (version current as of October 20, 2006).

Turner, B.A., 1978. <u>Man-Made Disasters</u>. London: Wykeham Science Press.

Vincent, C. and Amalberti, R., 2016. <u>Safer Healthcare</u>. New York: Springer.

Waterson, P.E. 2009. A critical review of the systems approach within patient safety research. Ergonomics, 52, 10, 1185-1195.

Waterson, P.E. 2014. Ed. Patient Safety Culture. Theory, Methods and Application, Farnham: Ashgate.

Waterson, P.E. 2016. Bridging the gap between the gap between research, policy and practice in health and safety. Policy and Practice in Health and Safety, 14, 2, 97-98

Waterson, P.E., 2017. 'I came into safety by accident' (interview with Professor Andrew Hale), The Psychologist, August, 64-66.

Waterson, P.E. and Catchpole, K. 2016. Human factors in healthcare: welcome progress, but still scratching the surface. BMJ: Quality and Safety, 25: 480-484.

Waterson, P.E., Clegg, C.W. and Robinson, M. 2014. Trade-offs between reliability, validity and utility in the development of human factors methods. In O. Broberg, N. Fallentin, P. Hasle, P.L. Jensen, A. Kabel, M.E. Larsen, T. Weller (Editors), Human Factors in Organizational Design and Management XI. Santa Monica: IEA Press.

Waterson. P.E., Robertson, M.M., Cooke, N.J. Militello, L, Roth, E. and Stanton, N.A., 2015. Defining the methodological challenges and opportunities for an effective science of sociotechnical systems and safety. Ergonomics, 58, 650-8.

Waterson, P.E., Jenkins, D.P., Salmon, P.M. and Underwood, P., 2017. 'Remixing Rasmussen': The evolution of Accimaps within systemic accident analysis. Applied Ergonomics. 59, Part B, 483-503.

Weick, K., 1987. Organizational culture as a source of high reliability. California Management Review, 29, 112-127.

Weick, K.E. 1989. Theory construction as disciplined imagination. <u>Academy of Management Review</u>, 14, 516-531.

Weick, K.E. and Sutcliffe, K.M. 2007. <u>Managing the Unexpected: Resilient Performance in and Age of Uncertainty</u>, Second Edition. San Francisco, CA: Jossey-Bass.

Wendler, R., 2012. The maturity of maturity model research: A systematic mapping study. <u>Information and Software Technology</u>, 54, 1317- 1339.

Westrum, R., 1993. Cultures with requisite imagination. In: Wise, J.A., Hopkin, V.D. and Stager, P., (Org.), <u>Verification and Validation of Complex Systems: Human Factors Issues</u>. Springer-Verlag, New York.

Westrum, R., 2004. A typology of organizational cultures. <u>Quality & Safety in Health Care</u>, 13, 22-27.

Wilson, J.R., 1987. Whose attitudes to which aspects of work? *or* Cool media for high participation. <u>Work and Stress</u>, 1, 385-395.

Appendix: Review sample

1. Parker, D., Lawrie, M. and Hudson, P., 2006. A framework for understand the development of organizational safety culture. Safety Science, 44, 551-562.

2. Tappin, D.C., Bentley, T.A. and Ashby, L.E., 2015. An implementation evaluation of a qualitative culture assessment tool. Applied Ergonomics, 47, 84-92.

3. Law, M. P., Zimmerman, R., Baker and G. R., Smith, T., 2010. Assessment of safety culture maturity in a hospital setting. Healthcare Quarterly, 13, 110–115.

4. Wallis, K. and Dovey, S., 2011. Assessing patient safety culture in New Zealand primary care: a pilot study using a modified Manchester Patient Safety Framework in Dunedin general practices. Journal of Primary Health Care, 3, 1, 35-40.

5. Jespersen, L., Griffiths, M., Maclaurin, T., Chapman, B. and Wallace, C. A., 2016. Measurement of food safety culture using survey and maturity profiling tools. Food Control, 66, 174-182.

6. Gordon, R., Kirwan, B. and Perrin, E., 2007. Measuring safety culture in a research and development centre: A comparison of two methods in the Air Traffic Management domain. Safety Science, 45, 669-695.

7. Fleming, M. and Wentzell, N., 2008. Patient safety culture improvement tool: development and guidelines for use. Healthcare Quarterly, 11, 10–15.

8. Kirk, S., Parker, D., Claridge, T., Esmail, A. and Marshall, M., 2007. Patient safety culture in primary care: developing a theoretical framework for practical use. Quality and Safety in Health Care, 16, 313-320.

9. Ashcroft, D.M., Morecroft, C., Parker, D. and Noyce, P. R., 2005. Safety culture assessment in community pharmacy: development, face validity, and feasibility of the Manchester Patient Safety Assessment Framework. Quality and Safety in Health Care, 14, 417–21.

10. Mohamed, S. and Chinda, T., 2010. System dynamics modelling of construction safety culture. Engineering, Construction and Architectural Management, 18, 3, 266-281.

11. Zhang, F., Zhang, Y., Liu, J. and Luo, H., 2013. Reseach on the maturity of real estate enterprises safety culture. Journal of Applied Sciences, 13, 11, 2019-2044.

12. Fleming, M., 2001. Safety culture maturity model. Offshore technology Report 2000/049, Health and Safety Executive, UK.
13. Goncalves Filho, A.P., Andrade, J.C.S. and Marinho, M.M.O., 2010. A safety culture maturity model for petrochemical companies in Brazil, Safety Science, 48, 615-624.

14. Bentley, T. and Tappin, D. 2010. Incorporating organisational safety culture within ergonomics practice. Ergonomics, 53, 10, 1167-1164.

15. Fleming, M. and Meakin, S., 2004. Cultural maturity model: health and safety improvement through involvement. In: The Seventh SPE International on Health, Safety and Environment in Oil and Gas Exploration, Calgary, Alberta, Canada.

16. Wright, M., Leach, P. and Palmer, G., 2012. A tool to diagnose culture in food business operators. Report from Greenstreet Berman Ltd for the UK Food Safety Agency. Reference GSB CL2567 R1 V6 FCA.

17. Phusavat, K., Vongvitayapirom, B. and Hidayanto, A.N., 2015. Enterprise development through the safety culture maturity model. International Journal of Management and Enterprise Development, 14, 2, 89–102.

18. Halligan, M. H., Zecevic, A., Kothari, A. R., Salmoni, A. W. and Orchard, T., 2014. Understanding safety culture in long-term care: a case study. Journal of Patient Safety, 10, 192-201.

19. Bentley, T. and Tappin, D., 2008. Qualitative evaluation of a framework for understanding the development of organisational safety culture. The Journal of Occupational Health and Safety Australia and New Zealand, 24, 3, 231-220.

20. Ohrn, A., Rutberg, H. and Nilsen, P., 2011. Patient safety dialogue: Evaluation of an intervention aimed at achieving an improved patient safety culture. Journal of Patient Safety, 7, 4, 185-192.

21. Xu, W. and Pan, J., 2011. A preliminary study on safety culture maturity model of chinese coal mine. In: Proceedings of 2011 International Conference on Management Science and Intelligent Control (ICMSIC 2011), Beijing, China, pp. 245-247.

22. Sheen, C. and Chung, D.W., 2011. Development of safety culture assessment model using safety culture maturity model and 4P-4C matrix. In: Proceedings of 2011 International Topical Meeting on Probabilistic Safety Assessment and Analysis, Volume 2, pp. 2143-2152.

23. Choi, K.S., Lee, Y.E., Ha, J. T., Chang, H. S. and Kam, S. C., 2010. Application of Safety Maturity Model and 4P-4C Model in Safety Culture Assessment. In: Transactions of the Korean Nuclear Society Autumn Meeting, Jeju, Korea, pp. 1115-1116.

24. Lardner, R., Fleming, M. and Joyner, P., 2001. Towards a mature safety culture. In: Symposium series n. 148 of the Institution of Chemical Engineers (IChemE), pp. 635-642.
25. Lingard, H., Zhang, R., Harley, J., Blismas, N. and Wakefield, R., 2014. Health and Safety. Culture Report 2014 of the RMIT University - Centre for Construction Work Health and Safety Research.

26. Goncalves, A. P., Kanegae, G. and Leite, G., 2012. Safety culture maturity and risk management maturity in industrial organizations. In: XVIII Internacional Conference on Industrial Engineering and Operations Management, pp. ID131.1-ID131.10.

27. Lawrie, M., Parker, D. and Hudson, P., 2006. Investigating employee perceptions of a framework of safety culture maturity. Safety Science, 44, 259-276.

28. Vongvitayapirom, B., Sachakamol, P., Kropsu-Vehkapera, H. and Kess, P., 2013. Lessons learned from applying safety culture maturity model in Thailand. International Journal of Synergy and Research, 2, 1, 5-21.

29. Gotivall, J., 2014. A Framework for Assessing Safety Culture. Master's thesis in Chemical Engineering. Department of Product and Production Development of Chalmers University of Technology, Gothenburg, Sweden.

30. McGeorge, D., Sunindijo, R.Y. and Zou, P.X.W., 2011. Criteria for the Development of a Safety Culture Maturity Model for the Construction Industry. In Proceedings of the International Symposium on Information Technologies in Safety Management of Large Scale Infrastructure Projects, Hubei, China.

31. Lunt, J., Bennett, V., Hopkinson, J., Holroyd, J., Wilde, E., Bates, S., Bell, N. and Webster, J., 2011. Development of the people first toolkit for construction small and medium sized enterprises. Research Report RR895 for the UK Health and Safety Executive.

32. Bergersen, C. E. B., 2003. Tool to be used to survey and improve safety culture in the european railway industry. Thesis in Safety, Health and Environment. Faculty of Social Sciences and Technology Management, Department of Industrial Economics and Technology Management at Norwegian University of Science and Technology. Trondheim, Norway.

33. Njeru, P. M., Mailutha, J. T., Gatebe, E. and Mburu, C., 2012. Evaluation of safety culture maturity levels of the Universities in Kenya. In: Proceedings of the 2012 JKUAT Scientific, Technological and Industrialization Conference, Nairobi, Kenya, pp. 349-361.

34. Fleming, M., 2007. Developing safety culture measurement tools and techniques based on site audits rather than questionnaires. Final project report Saint Mary's University in Halifax, Nova Scotia, Canada. http://www.pr-ac.ca/files/files/133_FinalReport_12May07.pdf (August, 04, 2016).

35. Vongvitayapirom, B. and Phusavat, K., 2013. Safety culture maturity in upstream: oil and gas industry in Thailand. In: Proceedings of 2013 International Conference on Technology Innovation and Industrial Management, Phuket, Thailand, pp. S4-264-290.

36. Guldenmund, F. W., 2008. Safety culture in a service company: a case study. Journal of Occupational Health and Safety – Australia and New Zealand, 24, 3, 221-235.

37. Ayers, G. F., Culvenor, J. F., Sillitoe, J. and Else, D., 2013. Meaningful and effective consultation and the construction industry of Victoria, Australia. Construction Management and Economics, 31, 6, 542-567.

38. Law, M. P., Parker, D., Lawrie, M. and Baker, G. R., 2008. A multidimensional approach to studying safety culture in health care. In: Proceedings of 2008 the International Conference of Healthcare Systems and Ergonomics and Patient Safety (HEPS), Strasbourg, France, pp. 1-4.

39. Law, M. P., 2011. Understanding and Changing the Patient Safety Culture in Canadian Hospitals. Thesis submitted for the degree of Doctor of Philosophy Department of Health Policy, Management and Evaluation University of Toronto, Canada.

40. Chen, S., 2016. Towards the definition of an OHS management maturity model and assessment tool. Thesis submitted for the degree of Master in Mechanical Engineering of School of Industrial and Information Engineering, Politecnico di Milano, Italy.

41. Jabonete, F. G. and Concepcion, L. R., 2016. Perceived safety culture of healthcare providers in hospitals in the Philippines. Journal of Sciences, Technology and Arts Research, 2, 1-14.

Table 1: Typology of organisation cultures (adapted from Westrum, 2004)

Typology of organisation	Characteristic	Example
Pathological	Preoccupation with personal power, needs, and glory	Information is hidden, messengers are "shot", responsibilities are shirked, bridging is discouraged, failure is covered up, new ideas are actively crushed
Bureaucratic	Preoccupation with rules, positions, and 'departmental turf'	Information may be ignored, messengers are tolerated, responsibility is compartmentalised, bridging is allowed but neglected, organisation is just and merciful, new ideas create problems
Generative	Concentration on the mission itself, as opposed to a concentration on people or positions	Information is actively sought, messengers are trained, responsibilities are shared, bridging is rewarded, failure causes inquiry, new ideas are welcomed

Table 2: Types of validity and testing (Trochim, 2006)

Type of validity	Definition	Testing
Face	Refers to the degree to which a measurement tool subjectively appears to measure the construct that it is supposed to measure (e.g., safety culture). In other words, face validity is when a measurement appears to do what it claims to do.	To assess the face validity of a safety culture measure, it would be sent the measure to a carefully selected sample of experts on safety culture measure and they all reported back with the judgment that the measure appears to be a good measure of safety culture.
Content	Refers to how accurately a measurement tool taps into the various aspects of the specific construct in question (e.g., safety culture). In other words, do the questions really measure the construct in question.	If a particular measurement tool is designed to measure safety culture, a group of experts on safety culture would evaluate each dimension chosen to measure safety culture and provide an opinion on how well each dimension taps into measuring of safety culture.
Convergent	Examines the degree to which a measurement is similar to (converges on) another measurement that it theoretically should be similar to.	Measurement tool should be able to show a correspondence or convergence between similar safety culture.
Discriminant	Examine the degree to which measurement is not similar to (diverges from) another measurement that it theoretically should be not be similar to.	Measurement tool should be able to discriminate between dissimilar safety culture.

Table 3: Applying the categorisation framework

Title/Author	Source	Country of origin	Domain	Method			Type of publication (Wendler, 2012)				Level descriptor
				D	**E**	**A**	**D**	**A**	**V**	**R**	
1. A framework for understanding the development of organisational safety culture Parker, D., Lawrie, M. and Hudson, P. (2006)	Safety Science	United Kingdom	Oil and gas	Literature review and interview	Focus group	Questionnaire or matrix	X				Level 1: Pathological Level 2: Reactive Level 3: Calculative Level 4: Proactive Level 5: Generative
2. An implementation evaluation of a qualitative culture assessment tool Tappin, D. C., Bentley, T. A. and Ashby, L. E. (2015)	Applied Ergonomics	New Zealand	Non-specific domain	Literature review	Interview and documental analysis	Interview and documental analysis	X				Level 1: Pathological Level 2: Reactive Level 3: Calculative Level 4: Proactive Level 5: Generative
3. Assessment of Safety Culture Law, M. P., Zimmerman, R., Baker, G. R., and Smith, T. (2010)	Healthcare Quarterly	Canada	Healthcare	Questionnaire	Questionnaire	Questionnaire	X				Level 1: Pathological Level 2: Reactive Level 3: Bureaucratic Level 4: Proactive Level 5: Generative
4. Assessing patient safety culture in New Zealand primary care: a pilot study using a modified Manchester Patient Safety Framework in Dunedin general practices Wallis, K. and Dovey, S. (2011)	Journal of Primary Health Care	New Zealand	Healthcare	Literature review and meeting group	Not reported	Matrix and work group	X				Level A Level B Level C Level D Level E
5. Measurement of food safety culture using survey and maturity profiling tools Jespersen, L., Griffiths, M., Maclaurin, T., Chapman, B. and Wallace, C. A. (2016)	Food Control	Canada	Food	Literature review and consultation with expert	Questionnaire	Questionnaire	X				Level 1: Doubt Level 2: React Level 3: Know Level 4: Predict Level 5: Internalize
6A. Measuring safety culture in a research and development centre: A comparison of two methods in the Air Traffic Management domain Gordon, R., Kirwan, B. and Perrin, E. (2007)	Safety Science	France	Air traffic control	Literature review	Questionnaire	Questionnaire	X				Level 1: Emerging Level 2: Managing Level 3: Involving Level 4: Cooperating Level 5: Continually improve
6B. Measuring safety culture in a research and development centre: A comparison of two methods in the Air Traffic Management domain Gordon, R., Kirwan, B. and Perrin, E. (2007)	Safety Science	France	Air traffic control	Literature review	Interview	Interview	X				Level 1: Uncertainty Level 2: Core Level 3: Extended Level 4: Leading Edge Level 5: World Class

Note: D= Development E= Evaluation A=application V= Validity R= Reliability

Table 3: Applying the categorisation framework (cont.)

Title/Author	Source	Country of origin	Domain	Method			Type of publication (Wendler, 2012)					Level descriptor
				D	**E**	**A**	**D**	**A**	**V**	**R**		
7. Patient Safety Culture Improvement Tool: Development and Guidelines for Use Fleming, M. and Wentzell, N. (2008)	Healthcare Quarterly	Canada	Healthcare	Literature review	Interview	Matrix	X					Level 0: Pathological Level 1: Reactive Level 2: Calculative Level 3: Proactive Level 4: Generative
8. Patient safety culture in primary care: developing a theoretical framework for practical use Kirk, S., Parker, D., Claridge, T., Esmail, A. and Marshall, M. (2007)	Quality and Safety in Health Care	United Kingdom	Healthcare	Literature review, consultation with expertise and interview	Interview and focus group	Matrix	X					Level 1: Pathological Level 2: Reactive Level 3: Calculative Level 4: Proactive Level 5: Generative
9. Safety culture assessment in community pharmacy: development, face validity, and feasibility of the Manchester Patient Safety Assessment Framework Ashcroft, D. M., Morecroft, C., Parker, D. and Noyce, P. R. (2005)	Quality and Safety in Health Care	United Kingdom	Healthcare	Focus group	Focus group	Matrix	X					Level 1: Pathological Level 2: Reactive Level 3: Calculative Level 4: Proactive Level 5: Generative
10. System dynamics modelling of construction safety culture Mohamed, S. and Chinda, T. (2010)	Engineering, Construction and Architectural Management	Thailand	Construction	Literature review	No reported	Questionnaire	X					Level 1:Uncommitted Level 2: Drifters Level 3: Improvers Level 4: Award winners Level 5: World Class
11. Research on the Maturity of Real Estate Enterprises Safety Culture Zhang, F., Zhang, Y., Liu, J. and Luo, H. (2013)	Journal of Applied Sciences	China	Construction	Literature review	Not reported	Questionnaire	X					Level 1: Emerging Level 2: Managing Level 3: Involving Level 4: Cooperating Level 5: Continually improve
12. Safety culture maturity model Fleming, M. (2001)	Offshore technology report n. 2000/049 for Health and Safety Executive	United Kingdom	Oil and gas	Literature review and focus group	Interview	No reported	X					Level 1: Emerging Level 2: Managing Level 3: Involving Level 4: Cooperating Level 5: Continually improve
13. A safety culture maturity model for petrochemical companies in Brazil Goncalves Filho, A. P., Andrade, J. C. S. and Marinho, M.M.O. (2010)	Safety Science	Brazil	Oil and gas	Literature review	Interview and questionnaire	Questionnaire	X					Level 1: Pathological Level 2: Reactive Level 3: Bureaucratic Level 4: Proactive Level 5: Generative

Table 3: Applying the categorisation framework (cont.)

Title/Author	Source	Country of origin	Domain	Method			Type of publication (Wendler, 2012)				Level descriptor
				D	**E**	**A**	**D**	**A**	**V**	**R**	
14. Incorporating organisational safety culture within ergonomics practice **Bentleya, T. and Tappin, D. (2010)**	Ergonomics	New Zealand	Utilities company	Literature review, interview and documental analysis	Not reported	Interview and documental analysis	X				Level 1: Pathological Level 2: Reactive Level 3: Calculative Level 4: Proactive Level 5: Generative
15. Cultural Maturity Model: Health and Safety Improvement through Involvement **Fleming, M. and Meakin, S. (2004)**	The Seventh SPE International Conference Calgary, Alberta, Canada	Canada	Non-specific domain	Literature and consultation with experts	Workshop	Matrix	X				Level 1: Documenting Level 2: Controlling Level 3: Engaging Level 4: Participating Level 5:Institutionalising
16. A Tool to Diagnose Culture in Food Business Operators **Wright, M., Leach, P. and Palmer, G. (2012)**	Report from Greenstreet Berman Ltd for the Food Standards Agency, Reference GSB CL2567 R1 V6 FCA	United Kingdom	Food	Literature review	Not reported	Documental analysis and observation	X				Level 1: Amoral calculators Level 2: Dependent Level 3: Doubters Level 4: Proactive compliers Level 5: Leaders
17. Enterprise development through the safety culture **Phusavat, K., Vongvitayapirom, B. and Hidayanto, A. N. (2015)**	International Journal of Management and Enterprise Development	Thailand	Non-specific domain	Literature review, consultation with experts, documental analyses and questionnaire	Questionnaire and consultation with expert	Questionnaire		X			Level 1: Pathological Level 2: Reactive Level 3: Calculative Level 4: Proactive Level 5: Generative
18. Understanding Safety Culture in Long-Term Care: A Case Study **Halligan, M.H., Zecevic, A., Kothari, A.R., Salmoni, A. W. and Orchard, T. (2014)**	Journal Patient Safety	Canada	Healthcare	Application of the maturity model developed in Publication 7	See Publication number 7	Interview, Focus group and questionnaire	X				Level 1: Pathological Level 2: Reactive Level 3: Calculative Level 4: Proactive Level 5: Generative
19. Qualitative evaluation of a framework for understanding the development of organisational safety culture **Bentleya, T., and Tappin, D. (2008)**	The Journal of Occupational Health and Safety Australia and New Zealand	New Zealand	Utilities company	Literature review, interview and documental analysis	Not reported	Interview; focus group and documental analysis	X				Level 1: Pathological Level 2: Reactive Level 3: Calculative Level 4: Proactive Level 5: Generative

Table 3: Applying the categorisation framework (cont.)

Title/Author	Source	Country	Domain	Method	Type of	Level descriptor

Title/Author	Source	Country of origin	Domain	Method			Type of publication (Wendler, 2012)				Level descriptor
				D	**E**	**A**	**D**	**A**	**V**	**R**	
20. Patient safety dialogue: Evaluation of an intervention aimed at achieving an improved patient safety culture Ohrn, A., Rutberg, H. and Nilsen, P. (2011)	Journal of Patient Safety	Sweden	Healthcare	Literature review	Not reported	Interview and documental analysis	X				Level 1: Pathological Level 2: Reactive Level 3: Calculative Level 4: Proactive Level 5: Generative
21. A preliminary study on safety culture maturity model of Chinese coal mine Xu, W. and Pan, J. (2011)	Proceedings of 2011 International Conference on Management Science and Intelligent Control	China	Coal mining	Literature review	Not reported	Not reported	X				Level 1: Initial Level 2: Managed Level 3: Preventive Level 4: Matured Level 5: Optimized
22. Development of safety culture assessment model using safety culture maturity model and 4P-4C matrix Sheen, C. and Chung, DW. (2011)	Proceedings of 2011 International Topical Meeting on Probabilistic Safety Assessment and Analysis	South Korea	Nuclear power	Literature review	Not reported	Not reported	X				Level 1: Pathological Level 2: Reactive Level 3: Calculative Level 4: Proactive Level 5: Generative
23. Application of Safety Maturity Model and 4P-4C Model in Safety Culture Assessment Choi, K.S., Lee, Y.E., Ha, J. T., Chang, H. S. and Kam, S. C. (2010)	Transactions of the Korean Nuclear Society Autumn Meeting	South Korea	Nuclear power	Literature Review	Not reported	Not reported	X				Level 1: Pathological Level 2: Reactive Level 3: Calculative Level 4: Proactive Level 5: Generative
24. Towards a mature safety culture Lardner, R., Fleming, M. and Joyner, P. (2001)	Symposium series n. 148 of the Institution of Chemical Engineers	United Kingdom	Oil and gas	Application of the maturity model developed in Publication 12	See Publication number 12	Workshop		X			Level 1: Emerging Level 2: Managing Level 3: Involving Level 4: Cooperating Level 5: Continually improve
25. Health and Safety Culture Lingard, H., Zhang, R., Harley, J., Blismas, N. and Wakefield, R. (2014)	Report 2014 of the RMIT University - Centre for Construction Work Health and Safety Research	Australia	Construction	Literature review	Not reported	Not reported		X			Level 1: Pathological Level 2: Reactive Level 3: Calculative Level 4: Proactive Level 5: Generative

Table 3: Applying the categorisation framework (cont.)

Title/Author	Source	Country of origin	Domain	Method	Type of publication (Wendler,	Level descriptor

Title/Author	Source	Country of origin	Domain	Method			Type of publication (Wendler, 2012)				Level descriptor
				D	E	A	D	A	V	R	
26. Safety Culture Maturity and Risk Management Maturity in Industrial Organizations Goncalves, A. P., Kanegae, G. and Leite, G. (2012)	XVIII International Conference on Industrial Engineering and Operations Management	Brazil	Oil and gas, footwear Manufacturing and infrastruture	Application of the maturity model developed in Publication 13	See Publication number 13	Questionnaire			X		Level 1: Pathological Level 2: Reactive Level 3: Bureaucratic Level 4: Proactive Level 5: Generative
27. Investigating employee perceptions of a framework of safety culture maturity Lawrie, M., Parker, D. and Hudson, P. (2006)	Safety Science	United Kingdom	Oil and gas	Application of the maturity model developed in publication 1	Questionnaire	See Publication number 1				X	Level 1: Pathological Level 2: Reactive Level 3: Calculative Level 4: Proactive Level 5: Generative
28. Lessons Learned from Applying Safety Culture Maturity Model in Thailand Vongvitayapirom, B., Sachakamol, P., Kropsu-Vehkapera, H. and Kess, P. (2013)	International Journal of Synergy and Research	Thailand	Oil and gas	Literature review	Questionnaire	Questionnaire			X		Level 1: Pathological Level 2: Reactive Level 3: Calculative Level 4: Proactive Level 5: Generative
29. A Framework for Assessing Safety Culture Gotivall, J. (2014)	Master's thesis in Chemical Engineering University of Technology Gothenburg	Sweden	Non-specific domain	Literature review	Accident analysis	Not reported			X		Level 1: Pathological Level 2: Reactive Level 3: Calculative Level 4: Proactive Level 5: Generative
29. A Framework for Assessing Safety Culture Gotivall, J. (2014)	Master's thesis in Chemical Engineering University of Technology Gothenburg	Sweden	Non-specific domain	Literature review	Accident analysis	Not reported			X		Level 1: Pathological Level 2: Reactive Level 3: Calculative Level 4: Proactive Level 5: Generative
30. Criteria for the Development of a Safety Culture Maturity Model for the Construction Industry McGeorge, D., Sunindijo, R.Y., Zou, P.X.W. (2011)	Proceedings of the International Symposium on Information Technologies in Safety Management of Large Scale Infrastructure Projects	Australia	Construction	Literature review	Not reported	Questionnaire			X		Level 1: Emerging Level 2: Managing Level 3: Involving Level 4: Cooperating Level 5: Continually improve

Table 3: Applying the categorisation framework (cont.)

Title/Author	Source	Country of origin	Domain	D	E	A	D	A	V	R	Level descriptor
31. Development of the people first toolkit for construction small and medium sized enterprises **Lunt, J., Bennett, V., Hopkinson, J., Holroyd, J., Wilde, E., Bates, S., Bell, N. and Webster, J. (2011)**	Research Report RR895 for the Health and Safety Executive	United Kingdom	Construction	Literature review, meeting discussion, consultation with expert, focus group and interview	Interview	Matrix		X			Level 1: Pathological Level 2: Reactive Level 3: Calculative Level 4: Proactive Level 5: Generative
32. Tool to be used to survey and improve safety culture in the European railway industry **Bergersen, C. E. B. (2003)**	Thesis in Safety, Health and Environment University of Science and Technology	Norway	Rail	Literature review, interview and focus group	Workshop	Questionnaire		X			Level 1: Pathological Level 2: Reactive Level 3: Calculative Level 4: Proactive Level 5: Generative
33. Evaluation of safety culture maturity levels of the universities in Kenya **Njeru, P. M., Mailutha, J. T., Gatebe, E. and Mburu, C. (2012)**	Proceedings of the 2012 JKUAT Scientific, Technological and Industrialization Conference	Kenya	Education (Universities)	Literature review	Not reported	Questionnaire		X			Level 1: Emerging Level 2: Managing Level 3: Involving Level 4: Cooperating Level 5: Continually improve
34. Developing safety culture measurement tools and techniques based on site audits rather than questionnaires **Fleming, M. (2007)**	Final project report Saint Mary's University Halifax Nova Scotia	Canada	Oil and gas	Literature review and consultation with experts	Questionnaire, scored card, interview and interrater reliability	Interview		X			Level 1: Pathological Level 2: Reactive Level 3: Calculative Level 4: Proactive Level 5: Generative
35. Safety culture maturity in upstream: oil and gas industry in Thailand **Vongvitayapirom, B. and Phusavat, K. (2013)**	Proceedings of 2013 International Conference on Technology Innovation and Industrial Management	Thailand	Oil and gas	Literature review	Questionnaire	Questionnaire		X			Level 1: Pathological Level 2: Reactive Level 3: Calculative Level 4: Proactive Level 5: Generative
36.Safety culture in a service company: a case study **Guldenmund, F. W. (2008)**	Journal of Occupational Health and Safety – Australia and New Zealand	The Netherlands	Non-destructive research	Application of the maturity developed in publication 1	Questionnaire	See Publication number 1			X		Level 1: Pathological Level 2: Reactive Level 3: Calculative Level 4: Proactive Level 5: Generative

Table 3: Applying the categorisation framework (cont.)

Title/Author	Source	Country of origin	Domain	Method	Type of publication (Wendler, 2012)	Level descriptor

							D	A	V	R	Level
37. Meaningful and effective consultation and the construction industry of Victoria, Australia Ayers, G. F., Culvenor, J. F., Sillitoe, J. and Else, D. (2013)	Construction Management and Economics	Australia	Construction	Literature review	Not reported	Documental analyses		X			Level 1: Pathological Level 2: Reactive Level 3: Calculative Level 4: Proactive Level 5: Generative
38. A multidimensional approach to studying safety culture in health care Law, M. P., Parker, D., Lawrie, M. and Baker, G. R. (2008)	Proceedings of the International Conference of Healthcare Systems and Ergonomics and Patient Safety	Canada and United Kingdom	Healthcare	Application of the maturity model developed in Publication 8	Questionnaire	See Publication number 8				X	Level 1: Pathological Level 2: Reactive Level 3: Calculative Level 4: Proactive Level 5: Generative
39. Understanding and Changing the Patient Safety Culture in Canadian Hospitals Law, M. P. (2011)	Thesis submitted for the degree of Doctor of Philosophy University of Toronto	Canada	Healthcare	Application of the maturity developed in publication 3	Questionnaire	See Publication number 3			X		Level 1: Pathological Level 2: Reactive Level 3: Calculative Level 4: Proactive Level 5: Generative
40. Towards the definition of an OHS management maturity model and assessment tool Chen, S. (2016)	Thesis for the degree of Master in Politecnico di Milano	China	Non-specific domain	Application of the maturity model developed in Publication 13	See Publication number 13	Questionnaire			X		Level 1: Pathological Level 2: Reactive Level 3: Bureaucratic Level 4: Proactive Level 5: Generative
41. Perceived safety culture of healthcare providers in hospitals in the Philippines Jabonete, F. G. and Concepcion, L. R. (2016)	Journal of Sciences, Technology and Arts Research	The Philippines	Healthcare	Not reported	Focus group and consultation with expert	Questionnaire		X			Level 1: Pathological Level 2: Reactive Level 3: Bureaucratic Level 4: Proactive Level 5: Generative

Table 4: Methods used to develop, evaluate and apply the maturity model

Method	Develop	Evaluate	Apply
Literature	1, 2, 4, 5, 6, 7, 8, 10, 11, 12, 13, 14, 15, 16, 17, 19, 20, 21, 22, 23, 25,		

	28, 29, 30, 31, 32, 33, 34, 35, 37		
Interview	1, 8, 14, 19, 31, 32	2, 6, 7, 8, 12, 13, 31, 34	2, 6, 14, 18, 19, 20, 34
Matrix			1, 4, 7, 8, 9, 15, 31
Consultation with expert	5, 8, 15, 17, 31, 34	41	
Focus group	9, 12, 31, 32	1, 8, 9, 36, 41	4, 12, 18, 19
Questionnaire	3, 17, 34	3, 5, 6, 13, 17, 27, 28, 34, 35, 36, 38, 39	3, 5, 6, 10, 11, 13, 17, 18, 26, 28, 30, 32, 33, 35, 40, 41
Documental analysis	17, 19	2	2, 14, 16, 19, 20, 37
Meeting discussion	4, 31		
Observation			16
Scored Card		34	
Inter-rater reliability		34	
Workshop		15, 32	24
Triangulation			2, 4, 14, 16, 18, 19, 20
Accident analysis		29	
Comparison		6, 36, 34, 39	
No reported	41	4, 10, 11, 14, 16, 19, 20, 21, 22, 23, 25, 30, 33, 37	12, 21, 22, 23, 25, 29

Table 5: Type of publication (Wendler, 2012)

Type of publication	Publication	Frequency
Maturity model development	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13,	33

	14, 15, 16, 17, 19, 20, 21, 22, 23, 25, 28, 29, 30, 31, 32, 33, 34, 35, 37, 41	
Maturity model application	18, 24, 26, 40	4
Maturity model validation	36, 39	2
Maturity model reliability	27, 38	2

Table 6: Level descriptor

Publication	Level descriptor	Frequency
1, 2, 3, 7, 8, 9, 13, 14, 17, 18,	Level 1: Pathological	28

19, 20, 22, 23, 25, 27, 26, 28, 29, 32, 34, 36, 35, 37, 38, 39, 40, 41	Level 2: Reactive Level 3: Calculative or bureaucratic Level 4: Proactive Level 5: Generative	
6A, 11, 12, 24, 30, 33	Level 1: Emerging Level 2: Managing Level 3: Involving Level 4: Cooperating Level 5: Continually improve	6
4	Level A Level B Level C Level D Level E	1
5	Level 1: Doubt Level 2: React Level 3: Know Level 4: Predict Level 5: Internalize	1
6B	Level 1: Uncertainty Level 2: Core Level 3: Extended Level 4: Leading Edge Level 5: World Class	1
10	Level 1: Uncommitted Level 2: Drifters Level 3: Improvers Level 4: Award winners Level 5: World Class	1
15	Level 1: Documenting Level 2: Controlling Level 3: Engaging Level 4: Participating Level 5: Institutionalising	1

Table 6: Level descriptor (cont.)

Publication	Level descriptor	Frequency
16	Level 1: Amoral calculators	1

	Level 2: Dependent Level 3: Doubters Level 4: Proactive compliers Level 5: Leaders		
21	Level 1: Initial Level 2: Managed Level 3: Preventive Level 4: Matured Level 5: Optimized	1	
31	Level 1: Starting blocks Level 2: Getting going Level 3: Walking Level 4: Running Level 5: Sprinting	1	

Table 7: Validity and reliability

	Publications with evidence that was carried out	Frequency	Publications without evidence that was carried out	Frequency
Reliability	2, 3, 13, 17, 27, 28, 31, 34, 35, 38, 41	11	1, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 18, 19, 20, 21,	30

			22, 23, 24, 25, 26, 29, 30, 32, 33, 36, 37, 39, 40	
Content Validity	1, 3, 5, 7, 8, 9, 12, 15, 17, 29, 31, 32, 34, 41	14	2, 4, 6, 10, 11, 13, 14, 16, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 30, 33, 35, 36, 37, 38, 39, 40	27
Face Validity	1, 3, 5, 7, 8, 9, 12, 15,17, 29, 31, 32, 34, 41	14	2, 4, 6, 10, 11, 13, 14, 16, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 30, 33, 35, 36, 37, 38, 39, 40	27
Convergent Validity	6, 34, 36, 39	4	1, 2, 3, 4, 5, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 35, 37, 38, 40, 41	37
Discriminant Validity	2, 34	2	1, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 33, 35, 36, 37, 38, 39, 40, 41	39
Reliability + Validity (face, content, convergent and discriminant)	34	1	4, 10, 11, 14, 16, 18, 19, 20, 21, 22, 23, 24, 25, 26, 30, 33, 37, 40	18

Figure 1: Quality Management Maturity Grid (Crosby, 1979) and Capability Maturity Model (Paulk et al., 1993)
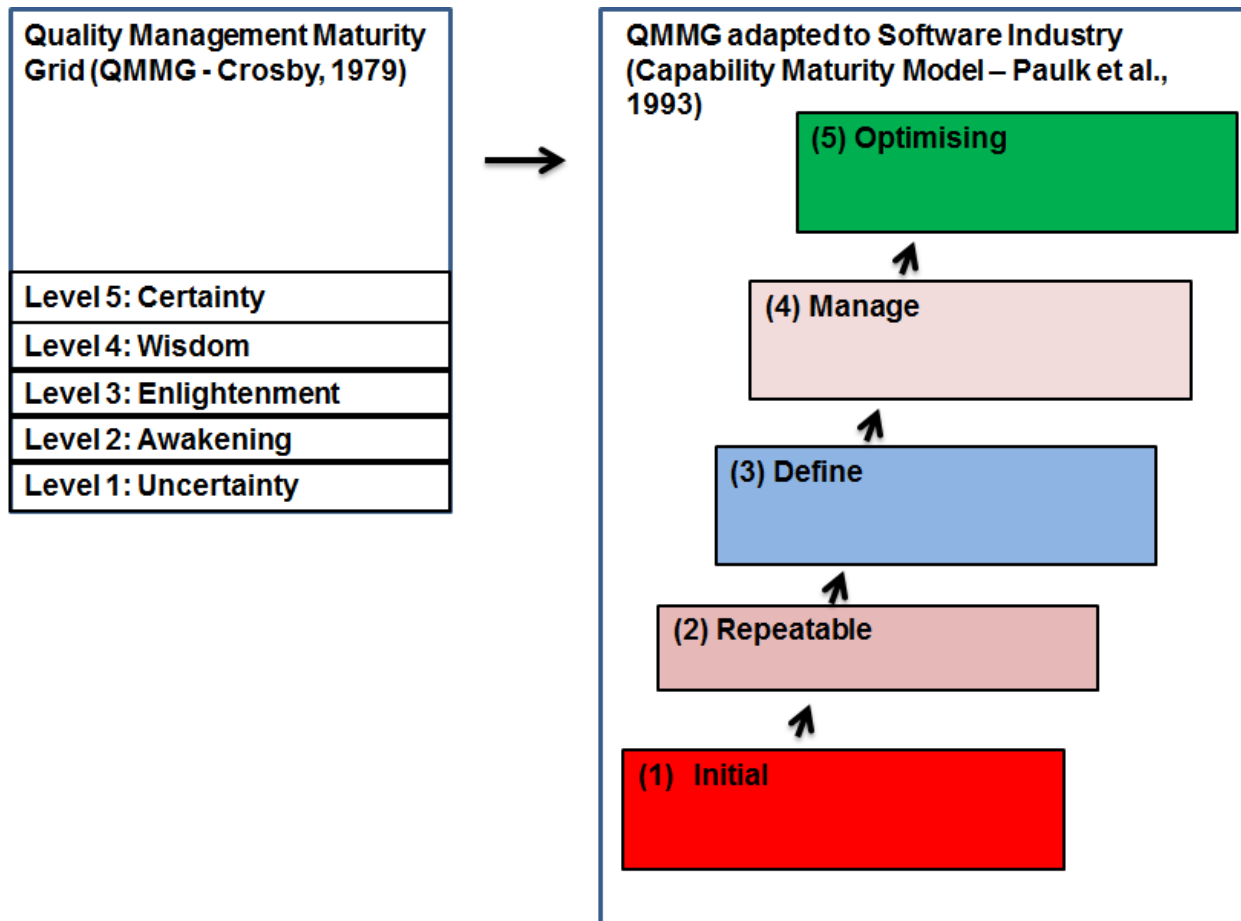
Figure 2: Typology of organisations (Westrum, 1993)

Typology of organisations proposed by Westrum (1993)

Generative

Bureaucratic

Pathological

Adapted by Hudson (2001) into 5 stages of maturity

(5) Generative

(4) Proactive

(3) Calculative
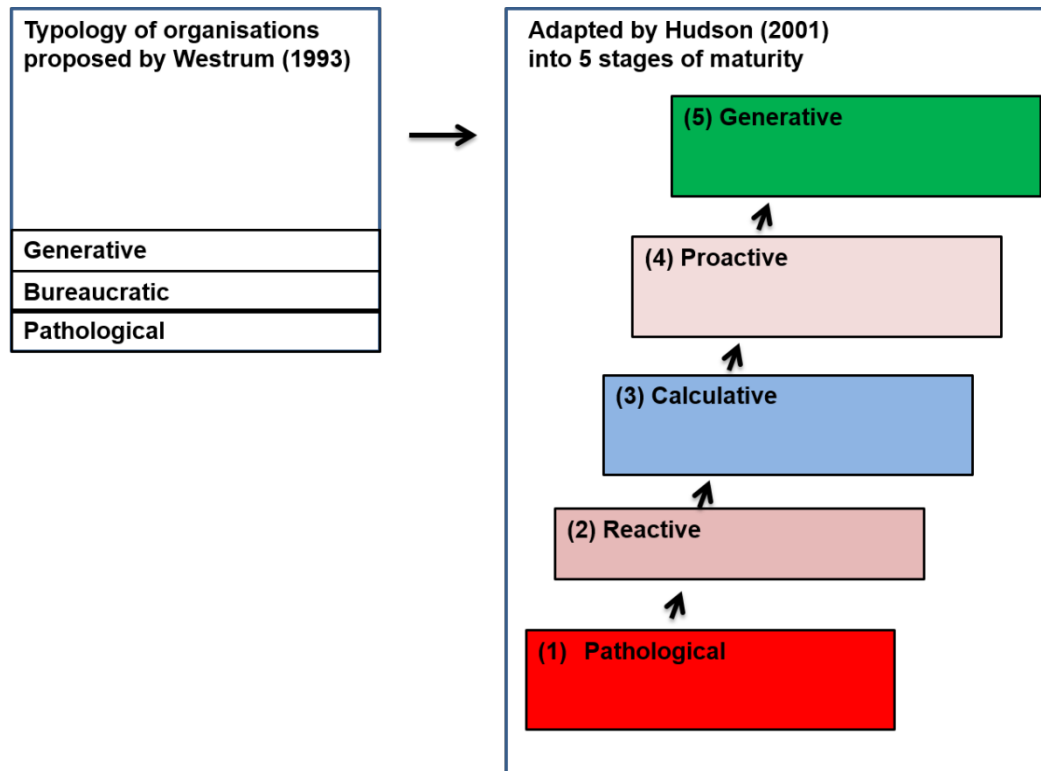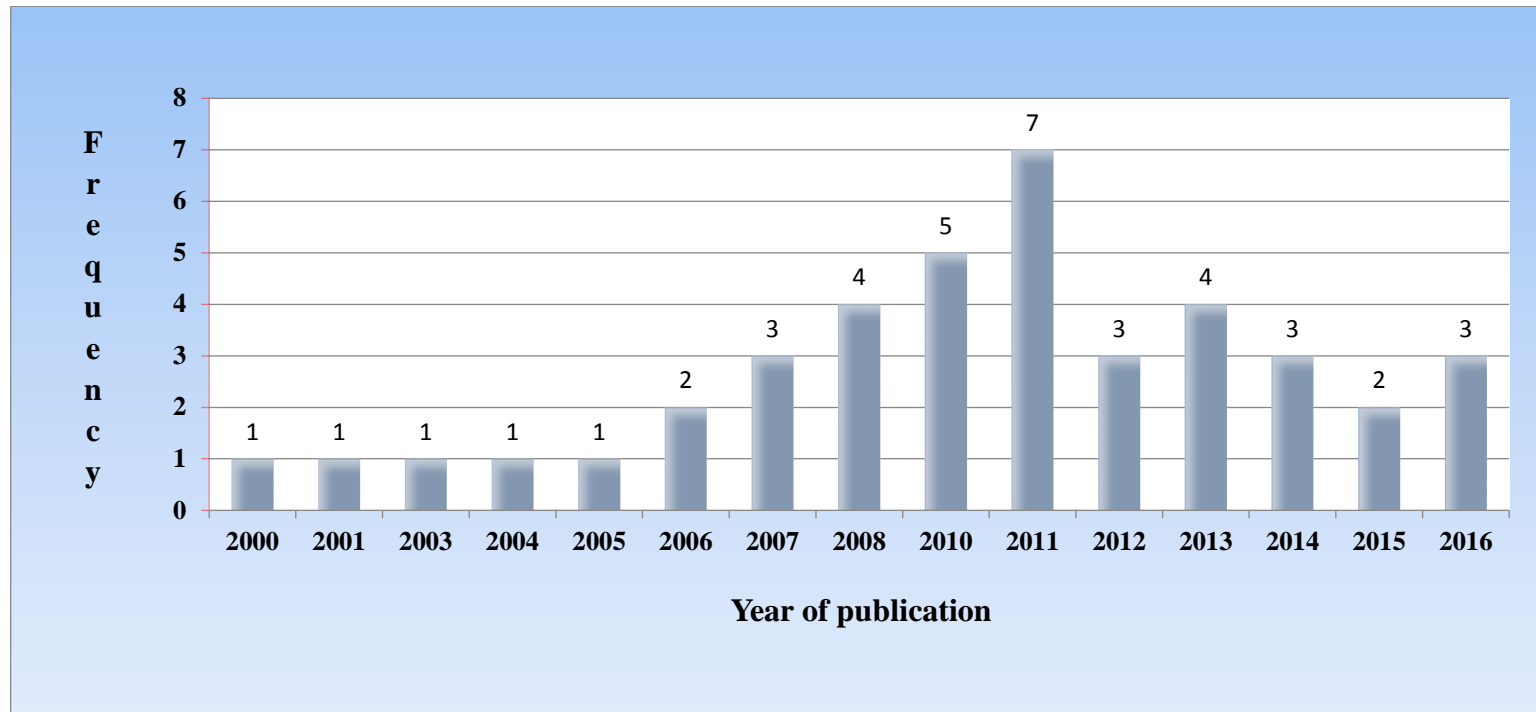
(2) Reactive

(1) Pathological

Figure 3: Year of publication

Figure 4: Maturity model development over time

Developments (vertical axis)

Time (horizontal axis)

Safety Maturity Model (SCMM) – Keil Centre (Fleming, 2001)

Manchester Patient Safety Assessment Framework (MaPSaF) (Ashcroft et al., 2005)

Food Safety Culture Tool (Food Standards Agency - Wright et al., 2012)

Maturity Models and Safety Culture

Oil & Gas (e.g., Fleming, 2000)

Construction (e.g., Mohammed and Chinda, 2010)

Rail (e.g., Bergerson, 2003)

Food (e.g., Wright et al., 2012)

Cultural Maturity Model (Fleming and Meakin, 2004

Nuclear (e.g., Choi et al., 2010)

Occupational Safety & Health (e.g., Chen, 2016)

Quality Management Maturity Grid (QMMG) (Crosby, 1979)

Healthcare (e.g., Ashcroft et al., 2005)

Health and Safety Culture Tool (Lingard et al., 2014)

Capability Maturity Model for Software – (Paulk et al., 1993)

Hearts and Minds Programme (e.g., Hudson and Willekes, 2000)

Air Traffic Control (e.g., Gordon et al. 2007)

Westrum's Typology of Organisational Cultures (Westrum, 1993, 2004)

1979-1990s   2000   2010   2017+