# A review and evaluation of safety culture and safety climate measurement tools

Trang Vu

Helen De Cieri

3 December 2015

**Research report:** 060-1215-R03

A joint initiative of

This research report was prepared for WorkSafe Victoria by Dr Trang Vu and Professor Helen De Cieri, Faculty of Business and Economics, Monash University, P O Box 197, Caulfield East, Victoria 3145, Australia.

## Acknowledgements

The research team would like to thank WorkSafe Victoria's project reference group, in particular Ross Pilkington, Jamie Swann and Warwick James for their continuing support of the project and assistance in facilitating contact with project stakeholders. We are very grateful to Dr Tracey Shea and Associate Professor Brian Cooper, Faculty of Business and Economics, Monash University for their expert advice on technical matters.

Thank you to Franz Carrillo and Trisha Pettit, Faculty of Business and Economics, Monash University, who have provided research assistance for this report.

The research team would like to acknowledge the generosity of authors who provided their questionnaires, research results and publications for inclusion in the review.

### Disclaimer



### Contact

Dr Trang Vu
Monash University
PO Box 197, Caulfield East, Victoria 3145, Australia
Phone +613 9903 4229
Email trang.vu@monash.edu

# Table of Contents

# List of Figures



# List of Tables

# List of Definitions

Several terms used in this report may be unfamiliar to some readers so we have provided in the table below a list of definitions of terms used in this report.

**Table 1: List of definitions of terms used in this report**

| Term | Definition |
|---|---|
| Construct validity* | Construct validity refers to what the scale or tool measures and this is based on the scales conceptual definition. This process typically involves conducting correlations between scale scores to determine what a scale measures (convergent validity) and what it does not measure (discriminant validity). This is a theoretically driven process and care must be taken to determine the expected relationships for a newly developed scale. |
| Content validity* | Content validity, in the context of this report, refers to the care taken to define and operationalise the construct for which a scale is being generated. This is a qualitative process that can include a literature review, consultation with subject matter experts and/or pre-testing of the scale. |
| Convergent validity* | A component of construct validity that is conducted to determine what a scale measures. This is evaluated using correlations between scale scores for the construct under investigation and another scale that has been demonstrated to measure a similar (or the same) construct. |
| Criterion validity* | Criterion validity comprises two forms of validity, <u>concurrent</u> and <u>predictive</u>, that describe the relationship between two scale scores. It is usually focused on the prediction of future behaviour or a future outcome (predictive validity) but often evaluates these relationships concurrently. Criterion validity is evaluated using correlations between a scale under development and a specified criterion such as OHS outcomes (e.g., OHS incident, near miss). |
| Discriminant validity* | A component of construct validity that determines what a scale does not measure. This is evaluated using correlations between scale scores for the construct under investigation and another scale that it should not be related to. When examining discriminant validity the magnitude of the correlation is expected to be low. |
| Face validity* | Face validity is the subjective judgement as to whether a scale is seen to measure the construct that it was developed to measure. |
| Known-groups validity* | A component of construct validity that is examined when finding established scales that can be used to evaluate construct validity is problematic. In known-group validity scale scores are compared for groups that are predicted to differ on the construct under investigation (e.g., industry, job role). |
| Reliability* | A reliability coefficient of over 0.7 is indicative of an acceptable level of internal consistency suggesting that the scale is reliable and has low levels of measurement error. |
| Safety climate# | Safety climate is the temporal state measure of safety culture, subject to commonalities among individual perceptions of the organisation. It is therefore situationally based, refers to the perceived state of safety at a particular place at a particular time, is relatively unstable, and subject to change depending on the features of the current environment or prevailing conditions. |
| Safety culture$ | Safety culture reflects the attitudes, values, and priorities of management and employees and their impact on the development, implementation, performance, oversight, and enforcement of safety and health in the workplace. |

References:

* DeVellis (2003); MacKenzie, Podsakoff, and Podsakoff (2011).
# Wiegmann, Zhang, von Thaden, Sharma, & Mitchell (2002, p. 10).
$ National Occupational Research Agenda (NORA) Construction Sector Council (2008, p. 65).

# Accompanying Documents

To cite this report:

Vu, T. & De Cieri, H., 2015a. *A review and evaluation of safety culture and safety climate measurement tools.* Melbourne, Australia: Institute for Safety, Compensation and Recovery Research (ISCRR), Monash University, Research Report number 060-1215-R03.

Accompanying documents available at www.iscrr.com.au:

Vu, T., & De Cieri, H. 2014. *Safety culture and safety climate definitions suitable for a regulator: A systematic literature review*. Melbourne, Australia: Institute for Safety, Compensation and Recovery Research (ISCRR), Monash University, Research Report 0414-060-R2C.

Vu, T. & De Cieri, H., 2015b. *Conceptual foundations of safety culture and safety climate measurement: A snapshot review.* Melbourne, Australia: Institute for Safety, Compensation and Recovery Research (ISCRR), Monash University, Research Report 060-1215-R04.

# 1. Executive Summary

## 1.1. Key messages

In recent times, the world of work has changed significantly, leading to new health and safety challenges for regulators, employers and workers. Hence, there is growing interest among researchers, practitioners and regulators in the field of occupational health and safety (OHS) in the development and use of tools[1] to measure safety culture and safety climate. The prospect of being able to use measurement data for action, particularly to inform policy and program development, is enhanced by the availability of a variety of tools developed to measure safety culture or safety climate.

Safety culture and safety climate are concepts that are often used interchangeably. In this report, however, they are considered as two distinct but related constructs. Most researchers would agree that safety culture is difficult to measure and it is unlikely that any of the tools available in the public domain to date are capable of measuring safety culture in its entirety. Efforts to measure safety culture are likely to benefit from a mixed methods approach drawing both qualitative and quantitative data from multiple sources and at multiple levels.

The majority of publicly available tools are designed to measure safety climate. Some of these tools also include items pertaining to some aspects of safety culture. This systematic review identified 18 publicly available tools that met evaluation criteria for reliability and validity.

## 1.2. Purpose

This review of tools that measure safety culture and safety climate is part of a larger research program supported by WorkSafe Victoria (WSV), via the Institute for Safety Compensation and Recovery Research (ISCRR), to investigate safety culture and safety climate. This review addressed the following questions:

1. What publicly available tools measure safety culture and/or safety climate?
2. Which tools meet evaluation criteria established for this review regarding reliability and validity?

Internationally, there is no consensus on definitions for safety culture and safety climate.  For the purpose of this report, safety culture is defined following the (US) National Occupational Research Agenda (NORA) Construction Sector Council (2008) and safety climate is defined following Wiegmann et al. (2002).

---

[1] In this report the terms `instrument', `measure', `assessment scale' and `tool' are sometimes used interchangeably. In the social science literature, the terms `measure' and `scale' are often used interchangeably (Kaplan, 2004). A scale contains items that collectively measure a construct.

*Safety culture reflects the attitudes, values, and priorities of management and employees and their impact on the development, implementation, performance, oversight, and enforcement of safety and health in the workplace (NORA Construction Sector Council, 2008, p.65).*

*Safety climate is the temporal state measure of safety culture, subject to commonalities among individual perceptions of the organisation. It is therefore situationally based, refers to the perceived state of safety at a particular place at a particular time, is relatively unstable, and subject to change depending on the features of the current environment or prevailing conditions (Wiegmann et al., 2002, p.10).*

## 1.3. Rationale

Tools for measuring safety culture and/or safety climate are helpful for the translation of safety culture and safety climate research knowledge into practice. If measurement data are to provide the basis for policy and practice decisions then reliable and valid tools are a prerequisite.

## 1.4. Methods

A systematic review of publicly available measures of safety culture and/or safety climate was conducted. The scope of the review was determined in consultation with stakeholders of the research. The search strategy covered both the grey and academic literature. Pre-defined selection criteria (see Table 2) were used to select publications for inclusion in the review.

**Table 2: Selection criteria**

| Outcome | Criteria |
|---|---|
| Included | The tool was developed to measure (worker) safety culture or safety climate. |
| | |
| Excluded | The tool was not available publicly (i.e., the tool was described in a conference abstract or thesis under an embargo). |
| | The tool is already included in the review because it has been reported in another eligible publication (i.e. duplicate reporting of the same validation results). |
| | The tool was developed for a specific safety issue in an organisation (e.g., railway safety) or a specific worker population (e.g., drivers) and cannot be generalised. |
| | The tool is a proprietary tool for which licence agreements and licence fees are required. |
| | Items in the tool not rated on a Likert or Likert-type rating scale. |

Tools developed to measure safety culture and safety climate that met the selection criteria were reviewed. The evaluation criteria (see Table 3) were developed with advice from a technical advisory group with expertise in psychometric principles and knowledge of OHS subject matter.

**Table 3: Evaluation criteria for tools reviewed in this report**

| Category | Criteria |
|---|---|
| **Item generation** | Uses multiple methods to generate and revise items to establish content validity:<br>• literature review;<br>• subject matter expert review;<br>• pre-testing of the scale. |
| **Latent structure (dimensionality)** | Reports underlying dimensions and evidence of latent structure such as variance, factor loadings, error terms and fit indices. |
| **Reliability** | Cronbach's alpha ≥0.7 for the scale if unidimensional or each subscale if multidimensional. |
| **Construct validation** | Includes at least one of the following validation processes: |
| - Convergent validity | Evidence of correlation between the scale and another that purports to measure a similar (or the same) construct. |
| - Discriminant validity | Low or no correlation to a scale that purports to measure a different construct. |
| - Known-group validity | t-test or analysis of variance conducted to test for differences in mean scores for two or more groups that are predicted to be different. |
| - Concurrent validity | Correlational analysis against an outcome variable that is expected to be associated with the new scale. |
| - Predictive validity | Correlational analysis against an outcome variable that is measured at a later time point and is expected to be associated with the new scale. |

Information collected using the evaluation criteria was used to qualitatively rate and compare safety culture and safety climate tools. The rating process considered four key evaluation areas: item generation, latent structure, reliability and construct validation. We used a qualitative rating scheme (Table 4) to rate each tool against the evaluation criteria. A tool was deemed to be satisfactory if all evaluation criteria were rated as satisfactory.

**Table 4: Rating scheme used to rate each evaluation criterion**

| Rating | Definition |
|---|---|
| **Satisfactory** | The tools in this category have been developed and reported in accordance with what could generally be considered best practice. No major issues identified that would significantly affect the soundness of a tool. |
| **Partially satisfactory** | The tools in this category satisfied the criteria for content validity and partially met the criteria regarding latent structure, reliability and construct validation. One or more of the following elements were omitted from the description of the tools: reporting an overall scale reliability rather than reliability coefficients for each subscale (in multidimensional tools), reliability below 0.7, latent structure or evidence of construct validity. |
| **Unsatisfactory** | The tools in this category mostly satisfied the criteria for content validity but they did not meet the majority of criteria regarding dimensionality, reliability and construct validation as outlined in Table 3. |
| **Unrated** | Rating not given due to lack of information about of the tool's reliability or validity in the publication. |

## 1.5. Key findings

Question 1: What publicly available tools measure safety culture and/or safety climate?

To address question 1, the search process found a total number of 412 publications that described a (worker) safety culture or safety climate tool to some extent. In these publications, a total of 220 tools were identified as being publicly available. The majority of these tools have been developed to measure safety climate.

Of the 220 tools identified in the search, 206 met the selection criteria and so were eligible to be included in the review.

Question 2: Which tools meet evaluation criteria established for this review regarding reliability and validity?

The current review evaluated 206 tools for which reliability and validity information was available.

Several observations can be made about the tools identified by the review:

- The tools are publicly available for the purposes of this review. It should be noted that the question of whether a tool is available for other purposes, such as for use in a training workshop or for implementation in a workplace, is beyond the scope of this report. Persons or organisations wishing to use these tools are advised to contact the author(s)/copyright holder(s) in the first instance.
- The majority of the tools developed to measure safety culture are actually measures of safety climate, with items measuring some aspect of safety culture. The term ‘safety culture’ has been used loosely in the labelling of tools, suggesting that the label of a tool might be imprecise and might not accurately reflect its contents;
- Authors’ descriptions of tools by terms such as ‘scorecard’ or ‘checklist’ might not truly reveal the characteristics of the tools;
- Some tools address safety climate as a uni-dimensional construct, while others view safety climate as having multiple dimensions; and
- Some of the tools measure safety climate at separate organisational levels by asking the respondent (e.g., an employee) to complete multiple scales for co-worker level, work group level and organisational level of safety climate.

The 206 tools were comprised of 125 original tools and 81 adapted versions of the original tools already included in the evaluation. The 125 original tools have been grouped into three categories: unsatisfactory (45 tools), partially satisfactory (62 tools) and satisfactory (18 tools). The 18 tools in the satisfactory category (Table 5) fully met evaluation criteria established for this review, which include:

- an item pool generated from literature review, expert consultation and pilot testing;
- a clear factor structure determined by factor analysis;
- reliability established by Cronbach’s alpha $\geqslant$0.7; and

- evidence of one or more forms of validity: convergent validity, discriminant validity, known-group validity, predictive validity and/or concurrent validity.

The 18 tools listed in Table 5 have evidence of both construct and criterion-related validity. These tools contain between four and 110 items and vary widely in the number of dimensions included. Management commitment to safety, employee involvement or empowerment in safety and safety communication are the core safety climate dimensions covered by these tools.

The 18 tools in Table 5 can be classified into two groups: (1) those that ask the respondent to report his or her safety climate perceptions of multiple workplace or organisational levels such as co-workers, supervisors and senior managers (three tools are in this group); and (2) those that do not make a distinction between management levels (15 tools are in this group). The first group might be more suitable for hierarchical and larger organisations while the second group might be potentially appropriate for a wider range of organisations or workplaces.

The decision-making process for selecting one of these measures for use in real world settings requires careful consideration of organisations' measurement needs and objectives, in addition to theoretical and practical factors. For anyone interested in using a safety climate tool in an organisation or workplace, we identify some important questions to ask and offer some guidance to address these questions.

**Table 5: List of safety climate and safety culture tools that meet evaluation criteria**

| Authors (year) | Measurement tool (number of items) |
|---|---|
| **Brondino, Pasini, & da Silva (2013)** | Organisational safety climate (17)<br>Supervisor safety climate (12)<br>Co-worker safety climate (12) |
| **Brown, Willis, and Prussia (2000)** | Safety climate (9) |
| **Díaz-Cabrera, Hernández-Fernaud, & Isla-Díaz (2007)** | Safety culture (44) |
| **Evans, Glendon, & Creed (2007)** | Safety climate (42) |
| **Fernández-Muñiz, Montes-Peón, & Vázquez-Ordás (2007)** | Safety culture (46) |
| **Glendon & Litherland (2001)** | Safety climate (32) |
| **Griffin & Neal (2000)** | Safety climate (18) |
| **Håvold (2007)** | Safety climate and safety culture (97) |
| **Hayes, Perander, Smecko, & Trask (1998)** | Safety climate (50) |
| **Heritage (2012)** | Manager safety climate (22)<br>Supervisor safety climate (18)<br>Co-worker safety climate (18) |

| Authors (year) | Measurement tool (number of items) |
|---|---|
| **Kines, Lappalainen, Mikkelsen, Olsen, Pousette, Tharaldsen, et al. (2011)** | Safety climate (45) |
| **Morrow & Barnes (2012)** | Safety culture (110) |
| **Prussia, Brown, & Willis (2003)** | Safety climate (4) |
| **Seo, Torabi, Blair, & Ellis (2004)** | Safety climate (32) |
| **Silva, Lima, & Baptista (2004)** | Safety climate (46) |
| **Tharaldsen, Olsen, & Rundmo (2008)** | Safety climate (32) |
| **Zohar (2000)** | Group-level safety climate (10) |
| **Zohar & Luria (2005)** | Organisation-level safety climate (16)<br>Group-level safety climate (16) |

## 1.6. Use of the research

The review represents a comprehensive, useful resource for WSV specifically and OHS practitioners, regulators and researchers. On the basis of this review, 18 tools developed to measure safety climate have been identified as fully meeting our evaluation criteria.

The question of whether any of the tools is available for purposes other than academic research is beyond the scope of this report. Persons or organisations wishing to use these tools are advised to contact the author(s)/copyright holder(s) in the first instance.

This review could be used in several ways, including:

- This review can be read alongside Vu and De Cieri's (2015b) snapshot review of the conceptual foundation of safety culture and climate measurement. The two reviews identify the conceptual challenges and methodological demands associated with safety culture and safety climate measurement. The reviews extend previous reviews of safety culture and safety climate tools and enhance the knowledge base for safety culture and safety climate measurement.

- This review could be a helpful guide for knowledge sharing and discussions among regulators and/or people working in OHS.

- People looking for a measure of (worker) safety climate to use in their workplace or with a group of workers could inspect the satisfactory tools, consider the practical issues, and select a tool that would be suitable for their work context, goals, priorities and challenges.

- In addition to viewing the tools with regard to their reliability and validity, contextual information relating to their development and validation could provide useful guidance. This includes: the industry in which these tools have been validated; whether these tools have been developed for a specific worker population or a particular organisational level; study sample (characteristics and sample size); study design (cross sectional or longitudinal); adaptations; data collection method (online survey, paper survey or interview); and whether scoring and data analysis guidelines are available.

- People responsible for developing a safety culture and climate research agenda could identify gaps in current knowledge and issues for future research by analysing contextual information relating to the development and validation of the satisfactory tools, and by examining the practical issues and questions raised in the discussion.

Considering these usage potentials, this review and evaluation will facilitate safety culture and climate measurement and contribute to the evidence base for effective OHS.

# 2. Introduction

In recent times, workplaces have changed dramatically in response to economic pressures, demographic challenges, technological innovations and employment shifts (Cappelli & Keller, 2013; Safe Work Australia, 2012; Shaw, 2013). These changes are increasingly being shown to be associated with health and safety concerns that have been influential in increasing interest in the constructs of safety culture and safety climate and their relationship with work-related illnesses and injuries (Nahrgang, Morgeson, & Hofmann, 2010). Further, there has been substantial and growing interest in the measurement of safety culture and safety climate (Silbey, 2009); there is now a wide variety of tools available that have been developed to measure safety culture or safety climate (Bernard, 2014). Nationally and internationally, regulators, researchers and practitioners have a shared interest in identifying and using reliable and valid tools, particularly for policy and program development.

This systematic review of tools that measure safety culture and safety climate is part of a larger research program supported by the WorkSafe Victoria (WSV), via the Institute for Safety Compensation and Recovery Research (ISCRR), to investigate safety culture and safety climate. Reports arising from the research to date include a stakeholder consultation report (Vu & De Cieri, 2013), a systematic review of safety culture and safety climate definitions from the perspective of a regulator (Vu & De Cieri, 2014) and a report on the conceptual foundations of safety culture and safety climate measurement (Vu & De Cieri, 2015b).

Safety culture and safety climate are concepts that are often used interchangeably. In this report, however, they are considered as two distinct concepts; this approach is consistent with scholars such Flin, Mearns, O'Connor, & Bryden (2000) and Lingard, Cooke, & Blismas (2009). In the absence of consensus definitions and for the purpose of this report, safety culture is defined following the (US) National Occupational Research Agenda (NORA) Construction Sector Council (2008) and safety climate is defined following Wiegmann et al. (2002).

> *Safety culture reflects the attitudes, values, and priorities of management and employees and their impact on the development, implementation, performance, oversight, and enforcement of safety and health in the workplace* (NORA Construction Sector Council, 2008, p.65).
>
> *Safety climate is the temporal state measure of safety culture, subject to commonalities among individual perceptions of the organisation. It is therefore situationally based, refers to the perceived state of safety at a particular place at a particular time, is relatively unstable, and subject to change depending on the features of the current environment or prevailing conditions* (Wiegmann et al., 2002, p.10).

## 2.1. Background and rationale

Scholars and investigators studying the causes of catastrophic events around the world in the 1980s realised that in order to fully explain and understand these events they needed to go beyond identifying traditional causes of disasters, viz. engineering failure and individual human error, to investigate the role of organisational cultural factors in workplace safety

(Health and Safety Commission (HSC), 1993). This realisation marks the beginning of the conceptualisation of safety culture and safety climate. The International Nuclear Safety Advisory Group (INSAG) introduced the term ‘safety culture’ to the nuclear industry in 1986 in its review of the Chernobyl nuclear disaster (INSAG, 1986). Since then, the safety culture construct has been applied in both nuclear and non-nuclear industries (Silbey, 2009).

Closely related to the safety culture construct is the construct of safety climate. Zohar (1980) coined the term ‘safety climate’ to describe a safety-specific organisational climate associated with employees' safety behaviour. A significant body of empirical evidence across several industries and countries accumulated since the 1980s suggests that safety climate is a good indicator of safety outcomes and ultimately workplace injury and illness (Christian, Bradley, Wallace, & Burke, 2009; Cox & Cheyne, 2000; Coyle, Sleeman, & Adams, 1995; Nahrgang et al., 2010; Neal, Griffin, & Hart, 2000; Seo et al., 2004; Zohar, 1980; Zohar, 2000; Zohar & Luria, 2005). Zohar (1980) showed that safety-specific organisational climate strongly correlated with industrial safety program effectiveness. Subsequent work by Zohar and Luria (2005) demonstrated that safety climate operationalised at both organisational and group levels, with organisational level safety climate (perceptions of senior managers) predicted group-level safety climate (perceptions of workgroup supervisors) which in turn predicted employee safety behaviour. A recent meta-analysis of 203 independent samples across four primary industries - construction, health care, manufacturing/processing and transportation - by Nahrgang et al. (2010) found that safety climate was positively related to compliance with safety and preventative measures; and negatively associated with worker anxiety, health, depression and work-related stress; and injury rates and injury severity.

Occupational health and safety (OHS) regulators are thus keen to translate research findings into policy and practice (Health and Safety Executive, 2005a; Independent Transport Safety Regulator (ITSR) & Transport Safety Victoria (TSV), 2012). This interest, however, has been hampered by conceptual and methodological ambiguity. Scholars disagree on how to define safety culture and whether or not it is a distinct concept from safety climate (Edwards, Davey, & Armstrong, 2013; Guldenmund, 2000; Wiegmann et al., 2002).

Conceptual ambiguity regarding these two constructs has led to divergent thinking on what to measure and differing practices for tool development and validation (Flin et al., 2000; Guldenmund, 2007). Consequently, a large number of tools developed to measure safety culture and/or safety climate is available. These tools vary in their levels of complexity and might be simple tick-box checklists or measures that address safety climate from a uni- or multi-dimensional perspective. Additionally, these tools vary in how they tap safety climate, with some developed to measure safety climate at multiple organisational levels (for example Zohar & Luria, 2005). To make well-informed choices regarding the suitability of these tools for use in real world settings, an evaluation of their reliability and validity is needed. The aims of this review are to identify and evaluate tools that measure safety culture and/or safety climate with a view to answering the following questions:

1. What publicly available tools measure safety culture and/or safety climate?
2. Which tools meet evaluation criteria established for this review regarding reliability and validity?

## 2.2. Organisation of this report

This report has several sections. First, the need for measurement of constructs such as safety culture and safety climate and the importance of selecting reliable and valid measures are discussed. Second, the methods used to systematically identify, evaluate and compare publicly available safety culture and safety climate tools are described. Third, findings pertaining to Question 1 are presented, beginning with a presentation of search results followed by an overview of publicly available tools found in the search. Fourth, the application of criteria to select safety culture and safety climate tools identified in Question 1 for inclusion in the evaluation (Question 2) are presented, including the identification of tools with satisfactory psychometric qualities. Finally, the review findings and implications are discussed.

## 2.3. Measuring safety culture and safety climate

Tools for measuring safety culture and/or safety climate are likely to facilitate the translation of safety culture and safety climate research knowledge into practice. Safety culture and/or safety climate measurement serves a number of purposes, including:

- gaining insights into where organisations are placed with respect to OHS matters and where best to focus resources to maintain and/or enhance safety culture and safety climate;
- establishing a baseline against which to evaluate the success or otherwise of safety culture and/or safety climate interventions; and
- monitoring progress towards safety culture and/or safety climate goal attainment (Eeckelaert, Starren, van Scheppingen, Fox, & Brück, 2011).

If safety climate measurement is to be used for these purposes and measurement outcomes provide the basis for policy and practice decisions, then reliable and valid safety culture and safety climate tools are a prerequisite. The attitudes, beliefs and norms implicit in the safety climate and safety culture constructs are subjective in nature and the constructs as a whole are not able to be directly observed. Consequently, a series of items must be developed that tap all elements of a latent construct and these items must correspond directly to the underlying construct they are said to represent (Adcock & Collier, 2001). Readers interested in the conceptualisation and approaches to measurement of safety culture and safety climate could consult a companion report by Vu and De Cieri (2015b).

Procedures for developing reliable and valid tools to measure constructs such as safety culture and safety climate are well-established (DeVellis, 2003; Hinkin, 1995; MacKenzie et al., 2011). The development of such tools is usually conducted over three stages using both qualitative and quantitative methods: item development, scale development and scale evaluation (Schwab, 1980 cited in Hinkin, 1985). The process of validating newly developed tools entail not only the initial assessment of their validity and reliability, but also ongoing confirmation from subsequent studies (DeVellis, 2003; Hinkin, 1995; MacKenzie et al., 2011). Another point to note is that the validity and reliability of any tool may need to be re-assessed when the tool is used in populations that differ from the original populations for which their validity and reliability have been established (MacKenzie et al., 2011). With regard to safety culture and safety climate, an additional consideration is whether tools developed and validated in one industry can be as valid and as reliable in other industries.

Zohar (2010) suggests that industry-specific safety climate tools may be preferable because safety climate perceptions could be context-dependent.

Best practice for item development involves identifying a relevant conceptual model or theory and developing items that tap into the key elements and relationships within the construct. The items are subsequently reviewed for content validity by experts and then ideally, administered to a development sample in a pilot study (DeVellis, 2003; Hinkin, 1995).

Best practice for scale development requires that an examination of the latent structure (dimensionality) using techniques such as factor analysis or Rasch analysis be conducted for a new tool (DeVellis, 2003; Hinkin, 1995). The latent structure of a tool refers to the underlying dimensions represented by the items in that tool. Once the latent structure of a tool is identified, the tool's internal reliability, also known as internal consistency, is evaluated. This is performed by calculating a reliability coefficient, also known as Cronbach's alpha coefficient (Cronbach, 1951). In most social science research, a reliability coefficient of over 0.7 is indicative of an acceptable level of internal consistency and suggests that the scale is reliable and has low levels of measurement error (DeVellis, 2003).

The final stage in the development and validation of a tool is to establish its validity. This process helps to establish whether the tool actually measures what it is intended to measure. The key components of the validation process are: construct validity (convergent, discriminant, known groups) and criterion validity (concurrent, predictive) (see Table 1). A more detailed discussion of validation issues can be found in DeVellis (2003).

In summary, safety culture and safety climate are abstract constructs with multiple dimensions that cannot be directly observed. While there remains some debate about the concept of safety culture, there is more clarity about safety climate and its measurement (Yule, 2003; Zohar, 2010). It is therefore necessary to ensure that, as far as practicable, the measurement of safety culture and safety climate uses a rigorous process that allows for the development of valid and reliable tools. In the following section, the methods used to identify publicly available safety culture and safety climate tools and to evaluate such tools are discussed.

# 3. Methods

The current review employed standard protocols for systematic reviews (Centre for Reviews and Dissemination (CRD), 2009). The scope of the review was determined in 2013 in consultation with stakeholders of the review, including WSV employees and union and employer representatives (Vu & De Cieri, 2013). After additional advice from a technical advisory group (see description below), the following review parameters were used:

- Search period: between January 1980 and June 2014;
- Type of literature: grey and published literature;
- Publications in English; and
- Websites: key Australian and overseas OHS regulators, OHS organisations, and OHS research centres.

These parameters were operationalised in the search strategies and selection criteria described below. It should be noted that non-English language publications were outside the review parameters due to a lack of time, resources and facilities for translation.

## 3.1. Literature search

Six electronic databases were searched for publications on safety culture and safety climate measurement tools published between January 1980 and June 2014:

- Business Source Complete;
- Embase;
- Ovid Medline;
- ProQuest;
- Psychinfo; and
- Safetylit.

Keywords used in the search were `questionnaire’, `survey’, `checklist’, `measures’, `tools’ and `scales’. These key words were combined with the terms `safety culture’ and `safety climate’.

The websites of key Australian OHS regulators, national and international OHS organisations and OHS research centres for were also searched. The key websites searched were as follows.

- Australian OHS regulators: WSV, Workcover NSW, WorkSafe WA, SafeWork SA, WorkSafe Tasmania, WorkSafe ACT, NT WorkSafe, Workplace Health and Safety Queensland and Comcare.
- National OHS organisations: Safe Work Australia (previously known as National Occupational Health and Safety Commission).
- International OHS organisations: WorkSafeNB (New Brunswick, Canada), European Agency for Safety and Health at Work (EU-OSHA), Health and Safety Executive (HSE), United Kingdom (UK), National Institute for Occupational Safety and Health (NIOSH) (US), SAI Global.
- OHS research centres: Monash Injury Research Institute; CARRS-Q, Queensland University of Technology; Institute for Work & Health, Canada; Liberty Mutual Research Institute for Safety, US.
- Other: National Offshore Petroleum Safety and Environmental Management Authority, Australian Council of Trade Unions, International Labour Organisation, World Health Organisation.

A manual search of the references cited in the retrieved papers and reports was also undertaken to identify additional publications not identified by the electronic searches. The table of contents of journal special issues on safety culture and/or safety climate were also searched for relevant articles. This search was performed on the following journals: Safety Science; Work and Stress; and Journal of Occupational Health and Safety - Australia and New Zealand (now Journal of Health, Safety and Environment).

## 3.2. Selection criteria

A technical advisory group was established to help develop criteria for selecting publications for inclusion in the review. The technical advisory group comprises a WSV representative working in the social research area and experts in psychometric tools at Monash University who are also familiar with the OHS literature. The technical advisory group reached consensus on the selection criteria listed below (Table 6), taking into consideration that the safety culture and safety climate literature is still evolving and that construct validation is a continuous process.

The abstracts of publications retrieved through the searches were reviewed. The full publications were subsequently examined if the abstracts were judged to be within the scope of the review.

**Table 6: Selection criteria**

| Outcome | Criteria |
| --- | --- |
| Included | The tool was developed to measure (worker) safety culture or safety climate. |
| | |
| Excluded | The tool was not available publicly (i.e., the tool was described in a conference abstract or thesis under an embargo). |
| | The tool is already included in the review because it has been reported in another eligible publication (i.e. duplicate reporting of the same validation study results). |
| | The tool was developed for a specific safety issue in an organisation (e.g. railway safety) or a specific worker population (e.g., drivers) and cannot be generalised. |
| | The tool is a proprietary tool for which licence agreements and licence fees are required. |
| | Items in the tool not rated on a Likert or Likert-type rating scale. Tools such as this are not standardised and cannot be assessed with regard to their reliability and validity. |

It should be noted that a large number of proprietary safety culture and safety climate tools are available to organisations via purchase or subscription (Cooper, 2002; Gilkey, Puerto, Keefe, Bigelow, Herron, Rosecrance, et al., 2012; National Safety Council of Australia Ltd, n.d.). Some of the safety culture and safety climate tools developed by OHS regulators are in this category, for instance the safety climate tool developed by the Health and Safety Laboratory (Health and Safety Laboratory, 2013). Safety culture applications (apps) for smartphones or devices are also available but these tools and apps are generally proprietary tools and considered to be outside the scope of this review. The focus of the review was on safety culture and safety climate tools which were publicly available at the time of the review and may be obtained free of charge.

## 3.3. Appraisal methods

It should be noted that construct validation is a continuing, complex process which may take several empirical studies to establish (MacKenzie et al., 2011). It follows that the various forms of validity presented in Table 7 might not all be investigated in a single study and hence the requirement for at least one form of validity evidence rather than all forms of validity evidence. This requirement was aimed at minimising any bias toward tools that have been developed and validated many years ago and therefore more likely to have been used more extensively than those developed more recently.

Safety culture and safety climate tools that met the above selection criteria were reviewed. Of the tools examined, those that measured the construct of safety culture or safety climate using multiple items with Likert or Likert-type response options, were evaluated with regard to their reliability and validity. Evaluation criteria were developed with advice from a technical advisory group with expertise in psychometric principles and knowledge of OHS subject matter. The evaluation criteria takes into account best practice in construct measurement and validation (Hinkin, 1995; MacKenzie et al., 2011).

**Table 7: Evaluation criteria for tools reviewed in this report**

| Category | Criteria |
|---|---|
| **Item generation** | Uses multiple methods to generate and revise items to establish content validity:<br>• literature review;<br>• subject matter expert review;<br>• pre-testing of the scale. |
| **Latent structure (dimensionality)** | Reports underlying dimensions and evidence of latent structure such as variance, factor loadings, error terms and fit indices. |
| **Reliability** | Cronbach's alpha ≥0.7 for the scale if unidimensional or each subscale if multidimensional. |
| **Construct validation** | Includes at least one of the following validation processes: |
| - Convergent validity | Moderate correlation to a scale that measures a similar construct. |
| - Discriminant validity | Low or no correlation to a scale that measures a different construct. |
| - Known-group validity | t-test or analysis of variance conducted to test for differences in mean scores for two or more groups that are predicted to be different. |
| - Concurrent validity | Correlational analysis against an outcome variable that is expected to be associated with the new scale. |
| - Predictive validity | Correlational analysis against an outcome variable that is measured at a later time point and is expected to be associated with the new scale. |

Convergent and discriminant validity are collectively known as construct validity (DeVellis, 2003; MacKenzie et al., 2011). Where possible we collected evidence of convergent, discriminant and known-group validity along with the results of factor analysis that were presented in each publication. Concurrent and predictive validity are also referred to as criterion-related validity (DeVellis, 2003; MacKenzie et al., 2011). During the appraisal of tools included in the current review, the `criterion' used to validate these tools was noted as to whether it is related to objective or subjective safety outcomes or both types of safety outcomes. See the list of definitions in Table 1 for more detail on the different components of the validation process.

The appraisal also documented the temporal relationship between safety culture or safety climate scores and the `criterion' (Haynes, Richard, & Kubany, 1995) where possible. Information on this temporal relationship not only helps determine the criterion-related validity of tools but also provides information on the potential utility of such tools. Evidence of a temporal relationship between safety climate or safety culture scores and the `criterion' is usually determined statistically by calculating a Pearson correlation coefficient. This coefficient represents the strength and direction of the association between safety climate/safety culture and the `criterion'. The current review used the guide for interpreting correlation coefficients proposed by Cohen (1988):

- small, where correlation coefficients have an absolute value of ≤ 0.3
- moderate, where correlation coefficients have an absolute value between 0.3 and 0.5
- large, where correlation coefficients have an absolute value of > 0.5

The evaluation criteria enabled the appraisal to be conducted in a systematic manner and facilitated the comparison of measurement tools to identify validated and reliable tools currently available in the public domain. Where a publication presents summary information on a tool rather than a complete tool, the first author was emailed to obtain a complete tool for appraisal purposes using a predefined protocol. When the first author did not reply within our follow-up timeframe (two weeks) or the email bounced back due to an incorrect email address the research team emailed the second author and so on. Authors were reminded twice about our request during a one month period. It should be noted that this process was not applied to publications reporting the development and validation of tools for use in workplace settings in less developed countries due to time constraints.

Information collected using the evaluation criteria was used to qualitatively rate and compare safety culture and safety climate tools included in the current review. The rating scheme shown in Table 8 was applied to each of the following key evaluation areas: item generation, latent structure (dimensionality), reliability and scale validity. A tool was rated as satisfactory if all evaluation criteria were rated as satisfactory. The rating scheme was qualitative instead of being quantitative because a difference in the total scores of two tools could reflect differences in one or more parameters in the key evaluation areas and it would not be meaningful to summarise these differences into a score. Furthermore, it would be difficult to define cut-off scores with which to categorise tools as satisfactory, partially satisfactory or unsatisfactory. As a comparison, we note that among the 21 reviews of safety culture and safety climate tools identified by the current review (see Figure 1 below), only one of these reviews used a quantitative scoring system to rank tools (The Keil Centre, 2003a).

**Table 8: Rating scheme used to rate each evaluation criterion**

| Rating | Definition |
|---|---|
| Satisfactory | The tools in this category have been developed and reported in accordance with what could generally be considered best practice. No major issues identified that would significantly affect the soundness of the tools. |
| Partially satisfactory | The tools in this category satisfied the criteria for content validity and partially met the criteria regarding latent structure, reliability and construct validation. One or more of the following elements were omitted from the description of the tools: reporting an overall scale reliability rather than reliability coefficients for each subscale (in multidimensional tools); reliability below 0.7; latent structure; or evidence of construct validity. |
| Unsatisfactory | The tools in this category mostly satisfied the criteria for content validity but they did not meet the majority of criteria regarding dimensionality, reliability and construct validation as outlined in Table 7. |
| Unrated | Rating not given due to lack of discussion of the tool's reliability or validity in the publication. |

It should be noted that while the focus of the evaluation of safety culture and climate tools was on their reliability and validity, contextual information relating to their development and validation was also documented. This includes: the industry in which these tools have been validated; whether these tools have been developed for a specific worker population; study sample (characteristics and sample size); study design (cross sectional or longitudinal design); adaptations; data collection method (online survey, paper survey or interview); and whether scoring and data analysis guidelines are available.

# 4. Results

## 4.1. Search results: Safety culture and safety climate tools

Using the search parameters discussed in Section 3.1 the review identified a total of 412 publications that included some description of (worker) safety culture and safety climate tools. The research team contacted nearly 100 authors of safety culture and safety climate tools via email to request them to provide their tools, if the tools have not been presented in full in their publications.  This resulted in 31 tools being made available by authors. The reasons authors gave for declining to provide their tools include tools being proprietary products, tools being lost and tool developer(s) not contactable.

It is noteworthy that a considerable number of publications on safety culture and safety climate tools were found in grey (i.e., not peer-reviewed) literature, suggesting that a substantial proportion of publications on safety culture and safety climate tools would have been missed had the review focussed on academic peer-reviewed literature only. The majority of publications from the grey literature are reports obtained by searching the websites of key OHS organisations and research centres identified in Section 3.1.

In Australia, a number of safety and OHS regulators have recently been involved in or commissioned the development of generic tools developed to measure safety culture or safety climate. These tools include:

- Organisational Safety Culture Appraisal Tool (OSCAT) by the Independent Transport Safety Regulator and Transport Safety Victoria (Independent Transport Safety Regulator & Transport Safety Victoria, 2012),
- Safety culture survey by WorkCover NSW (WorkCover NSW n.d.), and
- Safety culture checklist by Workplace Health and Safety Queensland (Workplace Health and Safety Queensland, 2013).

It is not known how these tools have been used in practice and whether an evaluation of users' experience has been performed. The current review also identified a number of toolkits which were noted for reference but were considered beyond its scope. This type of publication typically provides descriptive information on a limited range of tools developed to measure safety climate and/or safety culture and discusses questions of why, what and how to measure. Some tools and toolkits have been produced by OHS regulators such as the HSE 2005 toolkit for rail organisations (Health and Safety Executive, 2005a) and the HSE 1999 toolkit for offshore industries (Davies, Spencer, & Dooley, 2001).

### *4.1.1. Application of inclusion and exclusion criteria*

With regard to the inclusion criterion, it should be noted that the word 'worker' was used to make clear that the focus of this review is on workers', or employees' safety culture or safety climate, as distinguished from other forms of safety culture or safety climate.  For example, in healthcare, the concept of 'patient safety culture' has received much attention from academics and healthcare professionals (Flin, 2007). The term 'worker' is used to represent all people engaged in a work organisation, regardless of job type or employment status.

The publications found in the search included 21 published reviews of safety culture and safety climate measurement tools, which were not eligible for assessment yet could be used to identify issues pertinent to the evaluation and comparison of tools (Blewett & Flower, 2011; Davies, Spencer & Dooley, 2001; Health and Safety Executive, 2005a; The Keil Centre Ltd, 2003a). The 21 reviews were used to cross-check the search results with previously conducted searches. Some of these reviews have been commissioned by OHS or safety regulators, including the European Agency for Safety and Health at Work (Eeckelaert et al., 2011), the Health and Safety Executive (UK) (Health and Safety Executive, 2001), the Rail Safety and Standards Board (UK) (The Keil Centre Ltd., 2003b), and SafeWork SA (Blewett & Flower, 2011).

Another 29 publications were excluded because they provided general information about a tool or tools related to worker safety culture or climate, yet did not fully describe the tool(s).

The selection criteria were applied in order to determine the number of tools that were publicly available (Question 1) and which met the criteria to be included in the evaluation (Question 2).

Figure 1 summarises the selection process, which resulted in the exclusion of 194 publications.

- 28 tools were excluded because items in the tools are not rated on a Likert or Likert-type rating scale.
- 18 tools were excluded because at the time of the search (June 2014) the tools were not available publicly.
- 30 publications were excluded due to duplicate reporting, i.e. the same validation results were reported in other publications which had already been included in the review.
- 34 tools were excluded on the criterion that the tools were based on specific safety issues in organisations or were designed for a specific work context and could not be generalised.
- 34 tools were excluded because the tools are proprietary tools for which licence agreements and licence fees are required.

After all exclusions were made, a total of 220 tools were available to be included in the review and evaluation.

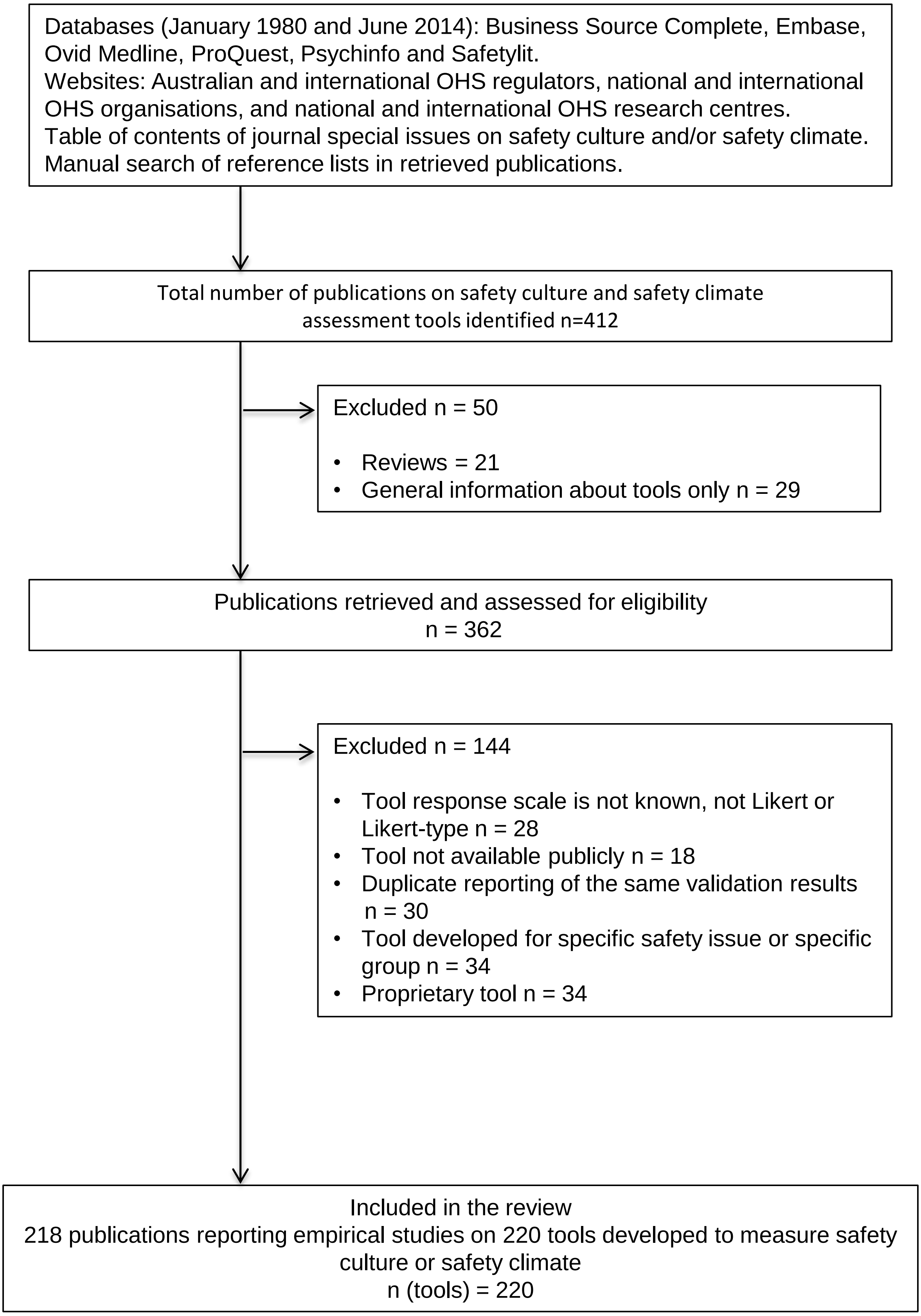


**Figure 1: Flow chart of publication selection process**

### *4.1.2. Comments on tools not rated on Likert or Likert-type response scales*

Twenty-eight tools were excluded because items in the tool are not rated on a Likert or Likert-type rating scale. It is necessary for these tools to be excluded from the evaluation because the tools cannot be assessed with regard to their reliability and validity. These tools may be useful for other purposes but are beyond the scope of this evaluation. These tools may be well-known to people interested in safety culture and safety climate, therefore, some additional information is offered here.

Two safety culture inspection/audits were found and these tools contain both open and closed questions for use during an inspection or audit to elicit information about key aspects of an organisation's safety culture (Health and Safety Executive, 2005a). This information is then evaluated against predefined criteria for good practice or predefined criteria for a particular level of safety culture maturity (Fleming, 2009) (see entry below regarding safety culture maturity model).

Nine checklists were found and these tools contain a set of questions or statements for which a tick box, a two-point response scale (yes/no) or a three-point response scale (e.g., yes/no/maybe) is provided. The questions or statements are aimed at determining if key processes for managing safety risks are in place. For those tools with a response scale, the total score is the sum of the scores for all questions or statements. For example, the Score Your Safety Culture checklist developed by (Reason, 2001) has 20 statements, with each statement being rated by a three-point response scale (yes/no/maybe). It has been suggested that the higher the total score the stronger the safety culture (Reason, 2001).

Two scorecards were found and these tools contain a mixture of measures that provide a framework for monitoring and benchmarking safety performance against safety goals. For example, Mohamed (2003) developed a scorecard for benchmarking construction safety culture based on the balanced scorecard model for managers developed by Kaplan and Norton (1991). These tools also provide a scoring system for monitoring progress towards safety culture maturity. For example, Bergersen (2003) developed the Track to Safety Culture Score Card which has 21 items to be scored by a five maturity-level rating scale: 'Denial', 'Reactive', 'Rule-based', 'Proactive', and 'Ideal'.

Two interview schedules were found and these tools contain a set of in-depth questions to explore basic assumptions about an organisation's safety culture and/or to gain a better understanding of results obtained from other tools such as questionnaires or checklists. For example, Bergh (2011) developed an interview schedule that has 14 questions.

Eight safety culture maturity models were found and these tools contain statements about an organisation's level of maturity for which respondents select one response from a list of safety culture indicators. For example, Lawrie, Parker, and Hudson (2006) developed a five-level safety culture maturity model. Each level of maturity is characterised by a specific set of indicators based on the behavioural and cultural elements of an organisation's safety culture.

## 4.2. Review Results: Tools measuring safety culture and/or safety climate

A total of 220 tools met the selection criteria and were eligible to be included in the review. Several observations can be made about the tools included in the review:

- The tools are publicly available for the purposes of research such as this review. It should be noted that the question of whether a tool is available for other purposes, such as for commercial purposes, for use in a training workshop or for implementation in a workplace, is beyond the scope of this report. Persons or organisations wishing to use these tools are advised to contact the author(s)/copyright holder(s) in the first instance.

- The majority of the tools developed to measure safety culture are actually measures of safety climate, with items measuring some aspect of safety culture. The term 'safety culture' has been used loosely in the labelling of tools, suggesting that the label of a tool might be imprecise and might not accurately reflect its content;

- Some tools address safety climate as a uni-dimensional construct while others view safety climate as having multiple dimensions; and

- Some of the tools measure safety climate at separate organisational levels by asking the respondent (e.g., an employee) to complete multiple scales for co-worker level, work group level and organisational level of safety climate.

### *4.2.1. Theoretical foundations*

Two main approaches have been used in the development of the 220 tools evaluated in the current review, including:

- a model-based approach in which the framework for measurement is a specific conceptual model or theory of safety culture, safety climate and/or related construct(s).
- a pragmatic approach in which elements of different conceptual models and theories of safety culture, and/or safety climate and/or related construct(s) are used to inform the development of tools.

The theories and models used in the development of these tools come from diverse academic disciplines (psychology, management, sociology, anthropology), including Schein's (1985) model of organisational culture, the theory of high reliability organisation (Weick & Sutcliffe, 2001), Cooper's (2000) reciprocal safety culture model and Zohar's multilevel safety climate model (Zohar & Luria, 2005). Each theory or model identifies a slightly different set of potentially relevant attributes for measurement. It is these differences in the conceptualisation of safety culture and safety climate that, in part, account for the proliferation of safety culture and safety climate tools.

The most frequently used theories and models are discussed elsewhere (Vu & De Cieri, 2015b). To date, none of these theories or models has been universally accepted as clearly articulating the construct domain of safety culture or safety climate and no single model or theory may be applicable to all types of organisations (Guldenmund, 2000). Nevertheless, it is generally accepted now that safety culture is a subset of organisational culture which is unconscious and invisible whereas safety climate is a snapshot or manifestation of safety culture that is observable, temporal in nature, subject to change and close to the surface (Guldenmund, 2000; Flin et al., 2000). Researchers seem to have greater clarity about safety climate and its measurement while recognising that it is related to the concept of safety culture (e.g., Cox & Flin, 1998; Flin, 2007; Zohar, 2010).

A common feature across the different theories and models used inform the development of safety culture and safety climate tools is that they support the conceptualisation of safety culture and safety climate as multidimensional constructs. According to Law, Wong, and Mobley (1998, p.741) a multidimensional construct:

> *…consists of a number of interrelated attributes or dimensions […] conceptualised under an overall abstraction and it is theoretically meaningful and parsimonious to use this overall abstract as a representation of the dimensions.*

Recent development of a multi-level safety climate model by Zohar and Luria (2005) defining safety climate as a multidimensional, multi-level construct has spurred the development of separate tools for examining safety climate at different organisational levels such as co-worker, workgroup and organisation (e.g, Brondino et al., 2013; Geddes, 2012; Zohar & Luria, 2005). These tools facilitate the examination of the key drivers of safety climate at each organisational level and the investigation of cross-level influences.

#### *4.2.2. What are the tools intended to measure?*

Of the 220 tools included in the review:

- 150 were developed to measure safety climate;
- 54 were developed to measure safety culture;
- 11 were developed to measure safety attitudes;
- 4 were developed to measure both safety culture and safety climate; and
- 1 was developed to measure safety attitudes and safety climate.

The majority of the tools that were described as being developed to measure safety culture were, upon inspection, focused on safety climate with the incorporation of some attitudinal items. This suggests that the tools measure visible aspects of safety culture rather than evaluate the deep, underlying assumptions characterising this construct.

#### *4.2.3. Target populations*

Among the 220 tools in the review, two types of tools were found: those which are aimed at specific worker populations (e.g., drivers, mobile/remote workers) or work roles (e.g., safety officers, OHS inspectors) and those which are generic and applicable across diverse workplace contexts. The majority of the tools are generic tools designed for administration to any worker population or industry. Generic tools, such as the organisation-level and group-level safety climate scales by Zohar and Luria (2005), have items relating to the core dimensions of safety climate, including management commitment to safety and employee involvement in safety, which would be applicable across industries. In contrast, tools aimed at a specific worker population or industry includes population or industry-specific items in addition to generic items. Huang, Zohar, Robertson, Garabet, Lee, and Murphy (2013a) developed the Trucking Safety Climate scale for measuring safety climate among lone workers, using truck drivers as exemplars, with 35 items to measure organisation-level and 26 items to measure supervisor/group-level safety climate. The scale includes six generic items each for organisation- and group-level safety climate. The remaining items in this scale cover industry-specific issues relating to job functions, communication patterns, work priorities and supervisory interactions. Huang et al. (2013a) show that industry-specific items have significantly stronger predictive validity over generic items. Zohar (2010, p.1521) suggests that the inclusion of industry-specific items would help identify "new, context-dependent targets of climate perceptions in respective industries".

#### *4.2.4. Additional information about the studies in the review*

Additional information about the studies reporting the reviewed tools is shown below:

- The earliest empirical study reporting the development and validation of a safety climate tool was published in 1980 (Zohar, 1980).
- Fourteen of the empirical studies have more than one independent sample.
- Study samples come from diverse industries or sectors, including agriculture, construction, education, electricity services, food services, healthcare, manufacturing, military, mining, nuclear, offshore oil and gas installations, public administration, retail, shipping and transport.
- Study samples from manufacturing constitute the largest category. Very few study samples were drawn from healthcare, small business and retail. The majority of studies have been conducted in a single industry in one country.

### 4.3. Evaluation results

Two hundred and twenty publicly available tools were included in the current review and were assessed on the basis of information presented in respective publications. An extensive process was undertaken to determine which tools met evaluation criteria as described in Section 3.3. The process (summarised in Figure 2) involved extracting and evaluating the following key information from the respective publications: items generation, dimensionality, reliability and construct validation (see Table 7 for full details). Where a publication reports more than one empirical study on a particular tool, information regarding all studies was extracted and recorded. For tabulation and discussion purposes, the earliest publication describing the development and validation of a tool was used as the reference point. Where this earliest publication is written in a language other than English, information provided in a subsequent English-language publication was used as the reference point.

Each of the four key areas evaluated was then qualitatively rated according to the rating scheme described in Table 8. A tool was considered to have fully met the evaluation criteria if it was rated as satisfactory in all evaluation areas. For tools that have been validated in several studies, information from all studies was taken into account in the evaluation. Finally, if a tool had been adapted for use in a workplace setting substantially different from that in which it was developed and the level of adaption was beyond minor changes in item wording and/or removal of non-applicable item(s) it was considered a `new' tool and evaluated as such. Information on the development and validation of 14 tools was not available. This resulted in the removal of 14 tools from the evaluation process, as they could not be rated.

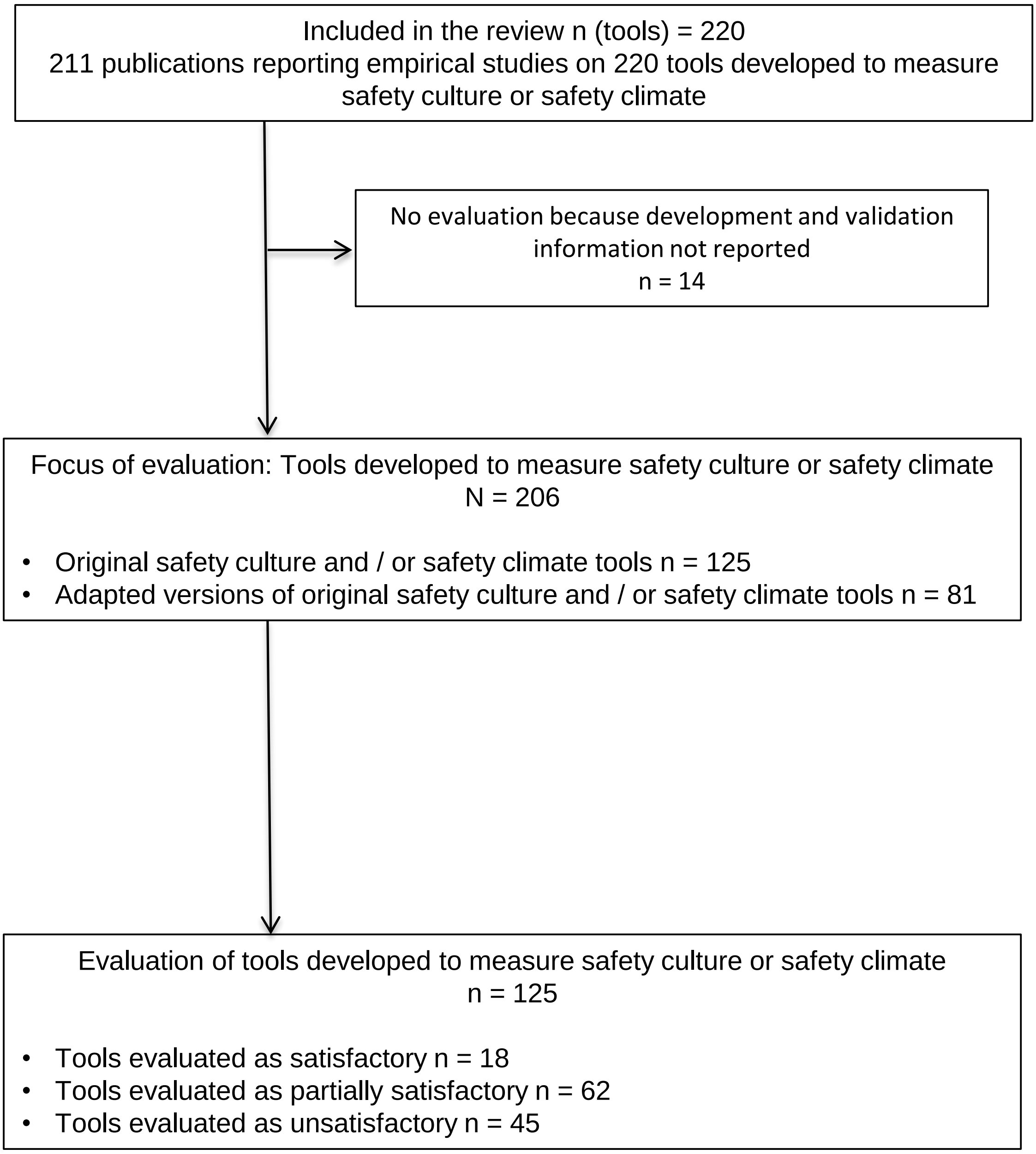


**Figure 2: Flow chart of review and evaluation process**

The focus of the evaluation of tools, therefore, was on the remaining 206 tools comprising of 125 original tools and 81 adapted versions of the original tools already included in the evaluation. Adaptations considered as minor for the purpose of classifying tools for the current evaluation include minor changes in item wording (such as replacing `employee' with `worker'), and/or addition of a small number of context-specific items, and/or deletion of items redundant to context of use. Reliability and validity results for adapted versions of the original tools, were used to assist the evaluation of the original tools.

The evaluation and rating of the 125 original safety culture and safety climate tools was mainly performed by one researcher. Three categories of tools were identified through the evaluation process: unsatisfactory (45 tools), partially satisfactory (62 tools) and satisfactory (18 tools). Each of these categories is presented below. *It should be noted that the current review does not present the actual tools in any categories due to copyright restrictions. The reader who is interested in these tools should contact relevant authors shown in the reference list.*

According to the evaluation, only a small number of extant safety culture/safety climate tools met our evaluation criteria. The evaluation found that evidence of construct validity was not reported for nearly 25 percent of the tools evaluated and criterion-related validity was also under-reported. Self-reported safety outcomes, such as workers' evaluation of workplace environmental conditions, safety behaviour, safety compliance, work-related injury and illness, near misses, and lost work days due to work-related injury and illness, have been commonly used as the criterion in construct validation. Few safety culture and safety climate tools have been validated using objective safety outcomes as the criterion (for example Silva et al., 2004; Zohar, 2000; Zohar & Luria, 2005). Objective safety outcomes included observer-reported safety behaviour; work-related injury, illness and mortality; and lost work days due to work-related injury and illness. Some safety climate tools have been validated using both subjective and objective safety outcomes as the criterion (for example Huang et al., 2013a; Tharaldsen et al., 2008).

### *4.3.1. Tools rated as unsatisfactory*

A total of 45 safety culture and safety climate tools were rated as unsatisfactory by using the evaluation criteria presented in Table 7 and the qualitative rating scheme shown in Table 8. These tools are questionnaires developed to measure safety attitudes (n=3), safety climate (n=27) and safety culture (n=15).

The tools in this category mostly satisfied the criteria for content validity but they did not meet the majority of criteria regarding dimensionality, reliability and construct validation as outlined in Table 7. The dimensionality of the construct under consideration was not reported for the majority of the tools in this category. Similarly, an evaluation of the reliability of these tools was not reported for the majority of them. For a small number of tools for which reliability information was reported, the reporting was inadequate such that an overall reliability coefficient, not reliability coefficients for each dimension of the construct under consideration, was presented. In short, the dimensionality, reliability and validity of the tools in the unsatisfactory category was inadequately investigated and/or inadequately reported.

### *4.3.2. Tools rated as partially satisfactory*

A total of 62 safety culture and safety climate tools were rated as partially satisfactory by using the evaluation criteria presented in Table 7 and the qualitative rating scheme shown in Table 8. These tools measure safety attitudes (n=6), safety climate (n=39) and safety culture (n=17).

The tools in this category satisfied the criteria for content validity and partially met the criteria regarding dimensionality, reliability and construct validation. Inadequacies in the testing and/or reporting of the psychometric properties of these tools include one or more of the followings:

- estimating and/or reporting an overall reliability coefficient only rather than reliability coefficients for each dimension of the construct under consideration;
- Reliability coefficient(s) below 0.7;
- not reporting the dimensionality of the construct under consideration;
- estimating and reporting test-retest reliability instead of reliability; and
- not presenting other evidence of construct validity such as criterion-related validity, known-group validity and discriminant validity.

Reporting quality accounted for approximately 50% of the partially satisfactory rating, indicating a need for reporting guidelines for publications presenting safety culture and safety climate measurement studies. In cases where reporting quality was the reason for the partially satisfactory rating, we identified no particular aspect of reporting that was consistently inadequate across respective studies. Rather, it was incomplete reporting on at least one evaluation area (item generation, dimensionality, reliability and construct validation) that was the reason for the rating. For example, Nielsen (2014) developed a safety climate tool based on published safety climate tools and tested the tool in a longitudinal study which found a variety of evidence supporting safety climate score reliability and predictive ability. The tool was rated as partially satisfactory, however, because no information on its dimensionality (factorial validity) was provided in the respective publication.

With regard to safety climate dimensions included in tools belonging to the partially satisfactory category, Huang, Chen, DeArmondc, Cigularov, & Chen (2007) and Huang, Ho, Smith, & Chen (2006) incorporated items on return-to-work policies as a dimension in their tools. This dimension is not among the safety climate dimensions identified by previous reviews of safety culture and safety climate theories (Guldenmund, 2000; Health and Safety Executive, 2005b). Huang and colleagues did not provide a rationale for this inclusion. The only other tool among the tools found by this review that incorporates such a dimension is the safety culture survey by WorkCover NSW (n.d.) (which was not evaluated due to lack of information on its development and validation).

Other related constructs found in the tools in the partially satisfactory category include:

- continuous improvement attitude (Hsu, Lee, Wu, & Takano, 2008);
- job hazards (Fang, Chen, & Wong, 2006);
- fatalism (Rundmo & Hale, 2003);
- impulsivity (DePasquale & Geller, 1999); and
- turnover intentions (Cree & Kelloway, 1997).

Beus, Payne, Bergman, & Arthur, (2010, p.713) refers to the inclusion of unrelated construct(s) such as these in a safety culture or safety climate measure as “content contamination”. The conceptual domain of safety culture and safety climate is not clearly defined and such inclusions are probably a consequence of conceptual ambiguity. The impact of content contamination on the measurement of safety climate will be discussed in the section below.

#### *4.3.3. Tools rated as satisfactory*

A total of 18 safety culture and safety climate tools were rated as satisfactory by using the evaluation criteria presented in Table 7 and the rating scheme shown in Table 8. The 18 tools (see Table 9) are developed to measure safety climate (n=14), safety culture (n=3), and both safety culture and safety climate (n=1). The tools developed to measure safety culture mainly contain items associated with safety climate and hence could be considered safety climate tools. Therefore, all of the tools in Table 9 can be viewed as safety climate tools.

The 18 tools in the satisfactory category contain between four and 110 items and vary widely in the number of dimensions included. Management commitment to safety, employee involvement or empowerment in safety and safety communication are the core safety climate dimensions covered by these tools.

The 18 tools in Table 9 can be classified into two groups:

(1) Tools that ask the respondent to report his or her safety climate perceptions of multiple workplace or organisational levels such as co-workers, supervisors and senior managers (three tools are in this group). These tools are based on a multilevel conceptualisation of safety climate. This group includes tools to measure safety climate perceptions at three workplace/organisation levels: co-workers (e.g., Brondino et al., 2013), workgroup supervisors (e.g., Brondino et al., 2013; Zohar & Luria, 2005) and senior managers (e.g., Brondino et al., 2013; Zohar & Luria, 2005). These tools enable the collection of safety climate perception data for each workplace/organisation level and may be used to explore cross-level interactions (Brondino et al., 2013).

(2) Tools that do not make a distinction between management levels (15 tools are in this group). These tools focus on the role of management in safety climate and do not make a distinction between management levels (supervisor vs senior management) (e.g., Griffin & Neal, 2000). This type of tools varies greatly in the number of dimensions included, tool lengths and the number of items for a particular dimension. These tools might be potentially appropriate for a wide range of organisations or workplaces.

The majority of the 18 tools have been developed and validated for manufacturing settings in the US, Europe and Australia. Since their initial validation, many of the 18 tools have been adapted for use in further empirical studies, either in the original setting(s) for which they were developed or in new settings. It should be noted that the data generated through the use of the adapted versions were also utilised where necessary to assist the evaluation of these tools.

These tools all have Likert or Likert-type response scales. The most common response scale has five points: strongly disagree, disagree, neither agree nor disagree, agree, strongly agree.

**Table 9: Summary of safety climate and safety culture tools that meet evaluation criteria**

| Author(s)<br>Measure<br>Number of items<br>Study sample | Items generation | Dimensions (number of items) | Latent structure | Reliability | Construct validation | Additional validation |
|---|---|---|---|---|---|---|
| **Brondino et al. (2013)**<br>**Safety climate**<br>**41 items**<br>**Manufacturing, Italy** | Items taken from literature and generated from interviews with stakeholders of the research. Questionnaire tested in a pilot study. | Co-worker safety climate scale<br>1. Co-workers' safety values (3)<br>2. Safety systems (3)<br>3. Safety communication (3)<br>4. Safety mentoring (3) | MCFA<br>Second-order factor structure<br>TLI = 0.94, CFI = 0.95,<br>SRMR = 0.04,<br>RSMEA = 0.096.0 | 0.83 to 0.91 | MCFA results indicate that the factor structure of these scales operates as expected at the organisational, supervisor and co-worker levels, providing evidence to support construct validity. Retrospective validity established with the higher the safety climate scores, the fewer number of self-reported injuries and micro-accidents in previous 6 months. | Manufacturing, Italy (Brondino, Silva, & Pasini, 2012) |
| | | Organisational safety climate scale<br>1. Management safety values (4)<br>2. Safety systems (5)<br>3. Safety communication (4)<br>4. Safety training (4) | MCFA<br>A second-order factor structure comprising of first order factors<br>TLI = 0.94, CFI = 0.95 | 0.76 to 0.81 | | |
| | | Supervisor safety climate scale<br>1. Supervisor's safety values (3)<br>2. Safety systems (3)<br>3. Safety communication (3)<br>4. Safety coaching (3) | MCFA<br>Two-factor structure<br>TLI = 0.92, CFI = 0.94<br>RMSEA >0.08,<br>SRMR 0.05. | 0.90 to 0.92 | | |

| Author(s)<br>Measure<br>Number of items<br>Study sample | Items generation | Dimensions (number of items) | Latent structure | Reliability | Construct validation | Additional validation |
|---|---|---|---|---|---|---|
| **Brown et al. (2000)**<br>**Safety climate**<br>**9 items**<br>**Manufacturing, US** | Plant tours and in-depth interviews in several plants were conducted to help generate survey items. The survey questionnaire was pilot-tested in a focus group. | 1. Supervisory safety climate (5)<br>2. Managerial safety climate (4)<br><br>Note the questionnaire also includes pressure (3 items), cavalier attitude (3 items), safety efficacy (3 items) and safe work behaviour (2 items). | PCA<br>Two-factor structure | 0.94 and 0.86 | Known-group validity - vocal union officials gave lower safety climate ratings of the department in which they worked compared to safety climate ratings of departments which had no vocal union officials | No |
| **Díaz-Cabrera et al. (2007)**<br>**Safety culture**<br>**44 items**<br>**Various industries, incl. aviation, gas and transport, Spain** | Literature review conducted to identify suitable items. HR experts provided feedback on the items. Questionnaire was pilot tested. | 1. Training program content (4)<br>2. Incident and accident reporting systems (12)<br>3. Orientation of safety rules and procedures (4)<br>4. Performance appraisal and safety promotion strategies (8)<br>5. Motivation patterns used (4)<br>6. Information and communication systems (4)<br>7. Leadership styles (8) | PCA<br>Six-factor structure explaining 67.8% of variance<br>$x^2$ (861) = 6824.67, p≤0.001 | >0.80 | Known-group validity with profiles of safety culture factors differed by type of company. | Manufacturing, Canada and US (Chenhall, 2010) |

| Author(s)<br>Measure<br>Number of items<br>Study sample | Items generation | Dimensions (number of items) | Latent structure | Reliability | Construct validation | Additional validation |
|---|---|---|---|---|---|---|
| **Evans et al. (2007)**<br>**Safety climate**<br>**42 items**<br>**Aviation, Australia** | Literature review, experts consultation and pre-testing | 1. Management commitment (7)<br>2. Safety training (7)<br>3. Equipment and maintenance (7)<br>4. Rules and procedures (7)<br>5. Schedules (7)<br>6. Communication (7) | EFA<br>Three-factor structure explaining 68% of total variance.<br>CFA<br>$x^2$ (128) = 322.26, p <0.001<br>RMSEA = 0.06<br>CFI = 0.97, IFI = 0.97 | 0.86 to 0.93 | Factor analysis results provide evidence of construct validity. Concurrent validity established with significant positive correlations between safety climate and perceived safety performance. | No |
| **Fernández-Muñiz et al. (2007)**<br>**Safety culture**<br>**57 items**<br>**Various industries, incl. construction, and services sectors, Spain** | Literature review, published scales, domain experts, pilot test | Safety Management System scale:<br>1. Safety policy (4)<br>2. Incentives for participation (5)<br>3. Training (9)<br>4. Communication (4)<br>5. Planning and control (11)<br>Managers' commitment scale:<br>1. Managers' attitudes (4)<br>2. Managers' behaviours (5)<br>3. Employees' involvement (4) | EFA<br>Factor loading and Cronbach alphas NR<br>CFA<br>Six-factor structure<br>$x^2$ (349) = 736.01, p=0.001. RMSEA = 0.058, NFI = 0.901, CFI = 0.910, GFI = 0.864, AGFI = 0.841 | ≥0.70 | The authors argue that there was evidence of convergent validity because standardised factorial regression coefficients relating each variable observed with the latent one was >0.5 and significant. Discriminant validity was considered to have been met because confidence intervals around the parameters that indicate correlations between factors do not include 1. | No |

| Author(s)<br>Measure<br>Number of items<br>Study sample | Items generation | Dimensions (number of items) | Latent structure | Reliability | Construct validation | Additional validation |
|---|---|---|---|---|---|---|
| **Glendon and Litherland (2001)<br>Safety climate<br>40 items<br>Roads construction and maintenance, Australia** | Items adapted from Glendon, Stanton, and Harrison (1994) and pilot tested on 10 employees. | 1. Communication & support (10)<br>2. Adequacy of procedures (6)<br>3. Work pressure (6)<br>4. Personal protective equipment (4)<br>5. Relationships (3)<br>6. Safety rules (3) | PCA<br>Six-factor structure explaining 69% of variance. | 0.72 to 0.93. | No relationship found b/w safety climate and observed safety behaviour. Known-group validity demonstrated with safety climate found to vary between subgroups. | Government services, UK (Birkbeck, 2010); rail, Australia (Darling, Edkins, Glendon, Lee, Lewis & Thompson, 2004) |
| **Griffin and Neal (2000)<br>Safety climate<br>18 items (Study 2)<br>Manufacturing & mining, Australia** | Literature review, sorting process to match items with theoretical domain, pre-testing | 1. Manager values (4)<br>2. Safety communication (4)<br>3. Safety practices (3)<br>4. Personnel training (4)<br>5. Safety equipment (3)<br><br>Note:<br>Other scales tested in this study: safety knowledge (4), compliance motivation (3), participation motivation (3). | CFA<br>Five-factor structure: four first order factors, all of which are contained in a second-order general factor reflecting the extent safety is valued within the organisation.<br>$\Delta X^2$ (4, N = 326) = 55.93, p <0.001 | 0.74 to 0.90 | CFA results provide evidence of construct validity. Concurrent validity demonstrated with safety knowledge shown to be positively related to self-reported safety compliance and safety participation. | Retail, US (Sinclair, Martin, & Sears, 2010); healthcare, Australia (Neal, Griffin & Hart, 2000); manufacturing, Australia (Wallace , 2005); manufacturing, US (Probst, 2004); petroleum & telecommunications, Canada (Zacharatos, Barling, & Iverson, 2005) |

| Author(s)<br>Measure<br>Number of items<br>Study sample | Items generation | Dimensions (number of items) | Latent structure | Reliability | Construct validation | Additional validation |
|---|---|---|---|---|---|---|
| **Håvold (2007)**<br>**Safety climate and safety culture**<br>**97 items**<br>**Shipping, Norway** | Items derived from a number of published scales, including Cox and Cheyne (2000); Glendon and Litherland (2001); Lee and Harrison (2000); Mearns, Whitaker, Flin, Gordon, and O'Connor (2003); Rundmo and Hale (2003); Williamson, Feyer, Cairns, and Biancotti (1997).<br>Questionnaire pilot tested. | 1. Management and employee commitment to safety (15)<br>2. Compliance to rules/safety norms/<br>occupational risk behaviour (11)<br>3. Workload/work pressure/stress (5)<br>4. Fatalism (6)<br>5. Knowledge/competence (5)<br>6. Safety values (6)<br>7. Conflict between safety and work/priorities (5)<br>8. Reporting culture (6)<br>9. Job satisfaction (3)<br>10. Officers awareness of risk (2)<br>11. Safety communication (2)<br>12. Learning culture (2)<br>13. Actions based on accidents (2)<br>14. Perception of safety instructions (2)<br>15. Work itself (2)<br>16. Safety behaviour (3) | PCA<br>Individual level: 15 factors accounting for 57% of variance.<br>Group level: 4 factors explaining 84% of variance. | Individual-level analysis: 10 out of 15 factors, Cronbach's alpha > 0.70. Another five factors 0.58 to 0.70.<br>Group-level analysis Cronbach's alpha >0.90<br>. | Known-group validity shown with significant differences in safety climate scores between nationalities. | No |

| Author(s)<br>Measure<br>Number of items<br>Study sample | Items generation | Dimensions (number of items) | Latent structure | Reliability | Construct validation | Additional validation |
|---|---|---|---|---|---|---|
| **Hayes et al. (1998)**<br>**Safety climate**<br>**50 items**<br>**Healthcare, US** | Items generated from a literature review.<br>Questionnaire pilot tested. | 1. Global perception of job safety (5)<br>2. Co-worker safety (5)<br>3. Supervisor safety (5)<br>4. Management safety practices (5)<br>5. Satisfaction with the safety program (5) | PCA<br>Five-factor structure explaining 63% of variance. | ≥0.91 | Convergent validity demonstrated through significant correlations between safety climate and safety awareness training, safety orientation and compliance with safety behaviours. Evidence of predictive validity - predictors of accidents were co-worker safety, management safety, satisfaction with safety program and job safety. | No |
| **Heritage (2012)**<br>**Safety climate**<br>**58 items.**<br>**Oil and gas, Australia** | Items were derived from a literature review and company's competency framework; and were pilot tested. | Manager scale (22 items in total)<br>1. Standards<br>2. Communication<br>3. Risk Management<br>4. Involvement | EFA<br>Four-factor structure | ≥ 0.70 | Supervisor and manager scales shown to have predictive validity through significant negative associations with self-reported near misses. All scales shown to have convergent validity through significant positive correlations with scores obtained from Zohar's (2000) safety climate scale | No |
| | | Supervisor scale (18 items in total)<br>1. Standards<br>2. Communication<br>3. Risk Management<br>4. Involvement | EFA<br>Four-factor structure | ≥ 0.70 | | |

| Author(s)<br>Measure<br>Number of items<br>Study sample | Items generation | Dimensions (number of items) | Latent structure | Reliability | Construct validation | Additional validation |
|---|---|---|---|---|---|---|
| | | Co-worker scale (18 items in total)<br>1. Standards<br>2. Communication<br>3. Risk Management<br>4. Involvement | EFA<br>Four-factor structure | ≥0.70 | | |
| **Kines et al. (2011)<br>Safety climate<br>45-65 items<br>(Current version has 50 items)<br>Various industries, incl. construction and food, Nordic countries** | Literature review, group consensus. Items pilot tested. | 1. Management safety priority, commitment and competence (9)<br>2. Management safety empowerment (4)<br>3. Management safety justice (4)<br>4. Workers' safety commitment (6)<br>5. Workers' safety priority and risk non-acceptance (7)<br>6. Safety communication, learning and trust in co-worker safety competence (8)<br>7. Workers' trust in the efficacy of safety systems (7). | CFA<br>Three-factor structure<br>$X^2$ (296) = 970.2, $p<0.001$<br>CFI = 0.90, RMSEA = 0.06<br><br>Group-related items<br>$X^2$ (344) = 1081.4, $p < 0.001$, CFI = 0.88, RMSEA = 0.06 | ≥0.71 | Convergent validity demonstrated by correlation patterns between scales. Known-group validity shown with significant differences in safety climate scores between organisational units. Concurrent validity demonstrated with safety climate dimensions correlating well with workers' safety motivation, with fewer self-reported safety violations and self-rated safety behaviour. | Manufacturing, Sweden (Bergh, 2011) |

| Author(s)<br>Measure<br>Number of items<br>Study sample | Items generation | Dimensions (number of items) | Latent structure | Reliability | Construct validation | Additional validation |
|---|---|---|---|---|---|---|
| **Morrow and Barnes (2012)**<br>**Safety culture**<br>**110 items**<br>**Nuclear industry, US** | Inductive and deductive based on INPO's principles for a strong nuclear safety culture, experts opinion and literature. | 1. Management commitment (36)<br>2. Willingness to raise concerns (9)<br>3. Decision-making (10)<br>4. Supervisor responsibility for safety (11)<br>5. Questioning attitude (9)<br>6. Safety communication (13)<br>7. Personal responsibility for safety (6)<br>8. Prioritising safety (6)<br>9. Training quality (6) | PCA<br>Nine-factor structure accounting for 58% of variance. 50 items were deemed redundant and a PCA of 60 remaining items identified a 7-factor structure (the `safety communication' and 'prioritising safety' factors were found to be very similar, statistically, to the `decision-making factor'. | 0.77 to 0.98 for 60-item survey | PCA results provided evidence for construct validity. Predictive validity demonstrated by (weak) correlations between safety culture and some safety performance indicators. | No |
| **Prussia et al. (2003)**<br>**Safety climate**<br>**4 items**<br>**Aviation, US** | Items based on Zohar's scale (1980) and information obtained from interviews conducted in several organisations. The survey questionnaire was pilot-tested. | Safety climate (4)<br><br>Note: scale used in a manager survey. See Brown et al., 2000 for employee survey scale. Other scales included in questionnaire are: safety versus production (3), cavalier attitude (3) and safety efficacy (3) | Single factor for safety climate. | 0.88 for safety climate. 0.53 for safety vs production and 0.87 for cavalier attitudes | Evidence of known-group validity: differences in safety climate perceptions were detected across departments and between managers and employees. | No. |

| Author(s)<br>Measure<br>Number of items<br>Study sample | Items generation | Dimensions (number of items) | Latent structure | Reliability | Construct validation | Additional validation |
|---|---|---|---|---|---|---|
| **Seo et al. (2004)**<br>**Safety climate**<br>**32 items**<br>**Agriculture, US** | Literature review, experts' review, pilot study | 1. Management commitment (7)<br>2. Supervisor support (5)<br>3. Co-worker support (6)<br>4. Employee participation (7)<br>5. Competence level (7) | CFA<br>Five-factor structure accounting for 52% of variance.<br><br>Calibration sample<br>$X^2$ (598) = 1.63<br>CFI = 0.98, IFI= 0.99, NNFI = 0.98, RMSEA = 0.05.<br><br>Validation sample had similar fit. | 0.78 to 0.84 | CFA results provide evidence of construct validity.<br>Evidence of concurrent validity shown by significant positive correlation of study's safety climate scores with those obtained from Brown et al. (2000), with self-reported time used to follow safety procedures and with safety behaviour. | No |
| **Silva et al. (2004)**<br>**Safety climate**<br>**46 items**<br>**Various industries, incl. utility, public administration, manufacturing and healthcare, Portugal** | Items generated through literature review and adaption of items from Ostrom, Wilhelmsen, and Kaplana (1993). Item pool evaluated by organisational psychology senior students to determine compatibility with chosen theoretical framework. | 1. Support, goals, innovation and rules (11)<br>2. Safety as an organisational value (5)<br>3. Organisation safety practices (22)<br>3. Personal involvement with safety (8) | CFA<br>Safety climate (2nd order)<br>$X^2$/df = 4.39<br>RSMR = 0.03, GFI = 0.91, NNFI = 0.95, CFI = 0.96, RMSEA = 0.07 | 0.72 to 0.83 | Authors suggest that construct validity was demonstrated on theoretical grounds. Predictive validity shown – organisations with strong safety climate had the lowest frequency and severity of accidents. | No |

| Author(s)<br>Measure<br>Number of items<br>Study sample | Items generation | Dimensions (number of items) | Latent structure | Reliability | Construct validation | Additional validation |
|---|---|---|---|---|---|---|
| **Tharaldsen et al. (2008)**<br>**Safety climate**<br>**32 items**<br>**Offshore oil industry, Norway** | Literature review to identify relevant items. Health and safety representatives, unions and industry provided feedback on chosen items. Questionnaire pilot tested. | 1. Safety prioritisation (8)<br>2. Individual motivation (5)<br>3. Safety management and involvement (11)<br>4. Safety versus production (4)<br>5. System comprehension (4) | PCA<br>Five-factor structure.<br>RMSEA = 0.039 (2001) and 0.041 (2003),<br>CFI = 0.97 (2001) and 0.98 (2003), GFI = 0.95 (2001) and 0.95 (2003), AGFI = 0.94 (2001) and 0.94 (2003) and<br>CN = 661.35 (2001) and 702.55 (2003). | ≥0.70 | Predictive validity shown with safety climate negatively associated with accident rates two years after the survey. Known-group validity shown by variations of safety climate between platforms and between 2 time points. | International shipping (Borgersen, Hystad, Larsson, & Eid, 2013); shipping, Norway (Nielsen, Eid, Hystad, Sætrevik & Saus, 2013); offshore industry, Norway (Nielsen, Eid, Mearns, & Larsson, 2013; Hope, Øverland, Brun, & Matthiesen, 2010; and Høivik, Tharaldsen, Baste, & Moen, 2009) |
| **Zohar (2000)**<br>**Safety climate (group-level)**<br>**10 items**<br>**Manufacturing, Israel** | Items were derived from interviews with workers and piloted with 123 workers. | 1. Supervisors' attitude to towards safety standards when under time pressure (2)<br>2. Supervisors' response when witnessing safe/unsafe behaviours (4)<br>3. Supervisors' communication and feedback on safety issues (4) | PCA<br>Two-factor structure | 0.87 to 0.90 | PCA results provide evidence for construct validity. Predictive validity shown with safety climate perceptions significantly predicted micro-incidents. | Construction, Australia (Lingard, Cooke, & Blismas, 2012); Manufacturing, UK (Clarke & Ward, 2006); oil refinery and manufacturing, Israel (Zohar & Luria, 2003); manufacturing, Israel (Zohar, 2002b; Zohar, 2002a); transport, US (Wallace, Popp, & Mondore, 2006); army, Israel (Zohar & Luria, 2004) |

| Author(s)<br>Measure<br>Number of items<br>Study sample | Items generation | Dimensions (number of items) | Latent structure | Reliability | Construct validation | Additional validation |
|---|---|---|---|---|---|---|
| **Zohar and Luria (2005)**<br>**Safety climate**<br>**32 items**<br>**Manufacturing, Israel** | Organisation-level safety climate scale (OSCS): Item pool developed from activities outlined in the British Standards Institute's safety management code. Items retained on the basis of representation of content themes and item loadings.<br><br>Group-level safety climate scale (GSCS): Items derived from a previously published GSCS (Zohar, 2000) (see above). Items retained on the basis of representation of content themes and item loadings. | OSCS<br>1. Active practices (7)<br>2. Proactive practices (4)<br>3. Declarative practices (5)<br><br>GSCS<br>1. Active practices (7)<br>2. Proactive practices (4)<br>3. Declarative practices (5) | EFA<br>OSCS: three-factor structure. Model fit statistics not reported.<br>GSCS: three-factor structure. Model fit statistics not reported. | 0.92 for OSCS;<br>0.95 for GSCS | OSCS: Predictive validity shown with moderate correlation between organisational safety climate and safety engineering audit scores given by an independent safety inspector.<br><br>GSCS: Predictive validity shown with moderate correlation between group-level safety climate and safety behaviour observations.<br><br>Note construct validity of adapted versions reported in other publications such as Huang et al. (2013a). | Power industry, US (Huang et al., 2013a); nuclear industry, Spain (Navarro, Lerín, Tomás, & Silla, 2013 and Martínez-Córcoles, Gracia, Tomás, & Peiró, 2011); manufacturing, UK (Nazaruk, 2011); manufacturing, US (Johnson, 2007); army, Israel (Luria, 2010) |

Abbreviations: AGFI: Adjusted goodness-of-fit index. CFA: Confirmatory factor analysis. CFI: Comparative fit index. EFA: Exploratory factor analysis. GFI: Goodness-of-fit index. GSCS: Group-level safety climate scale. HR: Human resource. INPO: Institute of Nuclear Power Operations. MCFA: Multilevel confirmatory factor analysis. NFI: Normed fit index. OSCS: Organisation-level safety climate scale. PCA: Principle component analysis. RSMEA: Root mean square error of approximation. SRMR: Standardized root mean squared residual. TLI: Tucker-Lewis index. $X^2$: Chi-square test.

*4.3.3.1 Item generation*
Items contained in the 18 tools in the satisfactory category were generated by:

- adapting items from published scales and adding newly developed items. For example, the 16-item group safety climate scale by Zohar and Luria (2005) validated in manufacturing was derived from the 10-item group safety climate scale by Zohar (2000), also validated in manufacturing.
  or
- using information from a combination of sources, such as literature review and feedback from stakeholders (workers, health and safety representatives, unions or industry) (e.g., Tharaldsen et al., 2008).
  or
- using safety management standards (e.g., the organisational level safety climate scale by Zohar & Luria, 2005) or safety culture principles as the framework for developing items (e.g., Morrow & Barnes, 2012).

All of the tools in the satisfactory category were also pilot tested before use in empirical studies. The sample size used in empirical studies was at least 200 respondents. It has been recommended that a sample size of at least 200 is required for an empirical study on psychometric tool development and validation (Hinkin, 1995).

The final versions of the tools in the satisfactory category vary widely in the total number of items. For example, the safety climate tool by Griffin and Neal (2000) has 18 items covering nine dimensions whereas the safety culture tool by Håvold (2007) has 97 items covering 16 dimensions. Some of the tools have fewer than 10 items in total (for example Prussia et al., 2003) while at the other extreme some tools have more than 100 items (for example Morrow & Barnes, 2012; Brown et al., 2000). The majority of tools have at least 32 items.

Another issue relating to the construct domain covered by the tools in the satisfactory category is that some tools include constructs that might not be universally accepted as dimensions of safety culture or safety climate. Debatable constructs include the following:

- job factors: job safety (Hayes et al., 1998), job satisfaction (Håvold, 2007).
- personality dispositions: fatalism (Håvold, 2007).

The potential impact of including debatable constructs on the measurement of safety climate (content contamination) will be discussed in the section below. Content contamination appears to be less of an issue for tools in the satisfactory category than those in the partially satisfactory category – which does seem to suggest that our evaluation criterion for item generation worked as intended. It should be noted that the dimensions included in the tool developed by Håvold (2007, p. 178) are contained in a safety orientation model and referred to as “factors or antecedents to safety culture or safety climate”.  It should be further noted that the tool developed by Hayes et al. (1998) is described in the respective publication as the Work Safety Scale.

*4.3.3.2 Dimensionality*
Various factor structures have been extracted from empirical data collected using the 18 tools in the satisfactory category. Among the structures reported, those having two (for example the safety climate scale by Brown et al., 2000) to six factors (for example the safety culture scale by Díaz-Cabrera et al., 2007) are the most common. Håvold (2007) (see Table

9) reported differences in the factor structure depending on the level of analysis: a 15-factor structure when safety climate data from individual workers were assumed to be completely independent and analysed at individual level[2] and a four-factor structure when the same data were aggregated and analysed at the employee's country of origin level[3]. This approach takes into account the multilevel nature of safety climate and represents one of a few instances in which the relationship between national culture and safety culture is investigated.

From the organisational behaviour literature, James (1982) explains that while the unit of observation for climate is the individual, it is individual responses appropriately aggregated to a higher level of analysis (teams, work groups, departments, etc.) that demonstrate shared perceptions of organisational climate. This explanation is particularly relevant here because safety climate has been defined by some as shared perceptions among workers about organisational commitment to safety. Readers interested in level-of-analysis issues should consult references such as James (1982) and Flin, Burns, Mearns, Yule, and Robertson (2006).

The level-of-analysis consideration was also addressed by Brondino et al. (2013) in their development and validation of the Integrated Organisational Safety Climate Questionnaire which comprises three scales: organisational safety climate scale, supervisor's safety climate scale and co-workers' safety climate scale. The authors used multilevel confirmatory factor analysis to investigate the factor structure of the scales and argued that the scales' factor structures align with the safety climate model by Griffin and Neal (2000), viz. safety climate is a higher-order factor comprising of specific first order factors (dimensions) at the individual and group levels (Brondino et al., 2013). The authors strongly advocated the use of these scales for exploring workers' perceptions of safety climate at each level and investigating the relationships and interactions between levels.

Until recently it has been a common practice in safety climate research for supervisor and management commitment to safety to be assessed using the same scale (for example Cooper & Phillips, 2004; Cox & Cheyne, 2000). Meliá and Sesé (2007, p. 232) argue that, "the particular contribution of each of these main sources of influence [supervisors, managers, co-workers, safety officers] should be measured separately". Zohar (2008, p. 379) further argues that supervisors have discretion in their interpretation and implementation of safety policy, and that individuals acting as members of an organisation and a work unit of that organisation, distinguish between "global and local emphasis on safety".

Management commitment to safety demonstrated through their attitudes and actions, employee involvement or empowerment in safety, and open and frequent communication between managers and workers are the core safety climate dimensions covered by the tools in the satisfactory category, except for the tool by Prussia et al. (2003) in which safety climate was conceptualised as senior management's commitment to safety. This narrower conceptualisation of safety climate has also been used by other authors (for example Neal & Griffin, 2006), suggesting the existence of two identifiable approaches in safety climate measurement employing narrower or broader conceptualisation of this concept. It should be

---

[2] Factor analysis using raw scores.
[3] Factor analysis using mean scores.

noted that the safety climate tool by Prussia et al. (2003) and that by Neal and Griffin (2006) both have been tested with an accompanying scale that measures safety motivation. This latter scale is considered by other authors as an attribute of safety climate (for example Tharaldsen et al., 2008; Díaz-Cabrera et al., 2007) or as an outcome criterion in construct validation (for example Kines et al., 2011).

*4.3.3.3 Reliability*

All of the empirical studies which were conducted to validate the tools in the satisfactory category reported Cronbach alphas of 0.7 or greater, with 0.7 being the cut-off value used in the current review to denote acceptable reliability, as recommended by experts (Hinkin, 1995; MacKenzie et al., 2011).

*4.3.3.4 Validity*

The majority of the tools in the satisfactory category have more than one source of evidence of validity. Evidence of the convergent and discriminant validity of these tools was established via factor analysis results and/or theoretical grounds. The majority of the tools have evidence for both convergent and discriminant validity. Brown et al. (2000), Díaz-Cabrera et al. (2007), Glendon and Litherland (2001), Håvold (2007) and Prussia et al. (2003) only provided evidence of the discriminant validity of their tools, suggesting that further testing is desirable.

Evidence of the concurrent and predictive validity of the tools in the satisfactory category was evaluated using one or more of the following outcomes:

- self-reported injuries (e.g., Brondino et al., 2013).
- self-reported micro `accidents' (e.g., Brondino et al., 2013; Zohar, 2000).
- self-reported near misses (e.g., Heritage, 2012).
- self-reported safety violations (e.g., Kines et al., 2011).
- safety motivation (e.g., Kines et al., 2011).
- safety behaviour (e.g., Glendon & Litherland, 2001; Seo et al., 2004).
- risk perception (e.g., Tharaldsen et al., 2008).
- safety engineering audit score (Zohar & Luria, 2005).
- `accidents' as recorded in official records (Hayes et al., 1998; Silva et al., 2004; Zohar, 2000).

The outcomes tested are diverse, indicating that the tools' authors have adopted a broad interpretation of what outcomes should be tested as part of construct validation. *It was noted that what might be considered outcomes in one tool are actually considered safety climate dimensions in another tool.* For instance, Kines et al. (2011) (NOSACQ-50) considered safety motivation as an outcome criterion whereas Tharaldsen et al. (2008) regarded safety motivation a dimension of safety climate. Furthermore, some of the outcomes listed above are intermediate outcomes – they are worker perceptions and behaviours that have been assumed to lead to fewer OHS injuries and illness and not actual OHS injuries and illness incidents. Few empirical studies in which the tools in the satisfactory category were tested have employed objective OHS injuries and illness incidence as outcome measures. The three studies using objective outcome measures (Hayes et al., 1998; Silva et al., 2004; Zohar, 2000) reported significant negative associations between safety climate and OHS injuries.

# 5. Discussion

## 5.1. Key findings

This review addressed two questions:

- Question 1: What publicly available tools measure safety culture and/or safety climate?
- Question 2: Which tools meet evaluation criteria established for this review regarding reliability and validity?

The search process identified a total of 412 publications potentially relevant to the review questions. The majority of these publications refer to tools that were developed to measure safety climate. Most of the tools that were described as safety culture tools were actually measures of safety climate, with items measuring some aspect of safety culture. The term ‘safety culture’ has been used loosely in the labelling of tools, suggesting that the label of a tool might be imprecise and might not accurately reflect its content. Similarly, authors’ description of their tools by terms such as ‘scorecard’ or ‘checklist’ might not truly reveal the characteristics of their tools, so further investigation may be advisable.

Among the 412 publications, some were targeted towards specific worker populations (e.g., lone workers, mobile/remote workers, injured workers) or work roles (e.g., managers, safety officers or safety coordinators). Other tools are generic in their targeting and applicable across diverse workplace contexts.

Safety culture and safety climate tools presented in the 412 publications are diverse and include both qualitative and quantitative instruments. The majority of the tools are quantitative instruments for measuring safety climate. A small number of tools are qualitative instruments. While qualitative tools were beyond the scope of this review, they could be used in conjunction with (quantitative) safety climate measures (e.g., Antonsen, 2009; Guldenmund, 2000, 2010) as part of a mixed methods evaluation of safety culture. Such an approach could provide a more in-depth understanding of safety culture than a quantitative survey alone. It has been suggested that organisations may convert information obtained from administering a safety culture or safety climate checklist or questionnaire to a maturity level using a safety culture maturity model, which may provide one way to develop and set safety goals and targets for improvement (Heese, 2012).

It is beyond the scope of this review to evaluate all types of publicly available tools measuring safety culture and safety climate that have been identified. Selection criteria were used to identify tools that were suitable for inclusion in the evaluation. Following application of the selection criteria, 220 tools were eligible for inclusion in a review of their reliability and validity. After further investigation, 14 tools were removed from the review because they could not be rated. After all exclusions, 206 tools were evaluated.

The 206 tools are comprised of 125 original tools and 81 adapted versions of the original tools already included in the evaluation. The tools were grouped into three categories through an extensive evaluation process: unsatisfactory (45 tools), partially satisfactory (62 tools) and satisfactory (18 tools). The tools in the unsatisfactory category mostly satisfied the criteria for content validity but they did not meet the majority of criteria regarding dimensionality, reliability and construct validation. The tools in the partially satisfactory category satisfied the criteria for content validity and partially met the criteria regarding dimensionality, reliability and construct validation. The tools in the satisfactory category met all evaluation criteria and therefore may be suitable for use by organisations or workplaces interested in safety climate measurement. *It should be noted that while all the tools in the satisfactory category fully met the evaluation criteria used in the current review, they vary in the completeness of validity evidence available*.

## 5.2. Scale development and validation issues identified by this evaluation

The results of the current review strengthen the rationale for evaluating publicly available safety culture and safety climate tools. A large number of tools have been developed to measure safety culture and safety climate and are publicly available. The evaluation found that almost 40% of tools did not meet the evaluation criteria while only 14% of the 125 original tools fully met evaluation criteria. This finding indicates that a considerable gap exists in empirical evidence for the reliability and validity of publicly available (quantitative) safety culture and safety climate tools. This evidence may be obtained in future research by applying established principles for measuring multidimensional constructs to the development and validation of safety culture and safety climate tools (Hinkin 1995; MacKenzie et al. 2011). Such principles include the following:

- An item pool adequately covering the conceptual domain of each construct.
- A stable factor structure determined by factor analysis.
- Internal consistency reliability determined by Cronbach's alpha.
- Construct validation to determine convergent validity, discriminant validity, predictive validity and concurrent validity.

The challenges in operationalising each of these principles in safety culture and safety climate measurement are discussed below.

### *5.2.1. Item development*

Item generation begins with a clear definition of the construct of interest (Hinkin, 1995). The definition helps focus measurement objectives and identify a set of potentially relevant attributes or dimensions for measurement. In the case of safety culture, this process is challenging because a consensus does not exist in the literature on what socio-psychological factors (such as attitudes, beliefs, values, norms) and dimensions should be included in a definition of safety culture and many definitions of safety culture are similar to those of safety climate. In fact, the two terms are often used interchangeably (Cox & Flin, 1998; Vu & De Cieri, 2014; Wu, Lin, & Shiau, 2010). The lack of clarity regarding the conceptualisation of safety culture and how it differs from safety climate has affected the labelling of the tools found in the current review, in particular tools developed to measure safety culture. Those tools actually measure safety climate (artefacts and espoused values regarding safety) and only some aspects of safety culture (inference from espoused values and artefacts).

A closely related issue is the lack of a common set of safety culture and safety climate dimensions, resulting in significant differences among tools in the number and nature of dimensions that constitute the conceptual domain. This issue is evident across all tools evaluated in this report and is further complicated by content contamination (Beus et al., 2010), such as the inclusion of personality dispositions, risk perceptions, social emotions and job satisfaction, as dimensions of safety culture and safety climate. Content contamination was found in tools in all three categories: satisfactory, partially satisfactory and unsatisfactory. Flin et al. (2000) argue that while personality dispositions may have a direct effect on safety behaviours, they do not predict safety outcomes. Beus et al. (2010, p.721) agree, noting that “personal safety attitudes are not descriptive of an organization’s safety policies, procedures, or practices, to which safety climate refers even when considered at the psychological level. As for perceptions of risks, Flin and colleagues’ view is that “some workers continue to take risks” despite having “fairly accurate perceptions of the risks they face”. The work of Beus et al. (2010) indicates that individual job safety/risk is the most prevalent “dimension-level contaminant” in safety climate measures. Regarding job satisfaction as a dimension in safety culture measures, Harvey, Erdos, Bolam, Cox, Kennedy, and Gregory (2002, p.22) claimed that job satisfaction is related to safety culture because it is “to some extent a function of management style and commitment”. It could be argued, however, that while job satisfaction may be related to management commitment to safety, it might be independent of it and influenced by a multitude of factors, including the nature of the job, individual factors and other organisational factors (Tietjen & Myers, 1998; Wilkin, 2013; Ziegler, Hagen, & Diehl, 2012).

Beus et al. (2010 p.716 and p.722) argue that the inclusion of debatable constructs “can create “noise” in safety climate-injury relationships and subsequently attenuate effect sizes” and that “greater contamination led to stronger safety climate-injury effects”. Given this problem, Flin (2007) suggests that a set of universal safety climate dimensions across industries or sectors is desirable, supplemented by context-specific set of factors relevant for each industry or sector. Since this set is currently not available, it is important to ensure that the dimensions included in a safety culture or safety climate tool and the tool as a whole correspond with how safety culture or safety climate is defined and conceptualised. Readers interested in the conceptual foundation of safety culture and safety climate measurement could consult the companion report by Vu and De Cieri (2015b).

#### *5.2.2. Latent structure*

Factor analysis is an analytical technique used to validate a scale by demonstrating that its constituent items share some common underlying (latent) factors (dimensions) (Hinkin, 1995; MacKenzie et al., 2011). Among the 125 original safety culture and safety climate tools that were evaluated, information regarding the factor structure of 45 tools (36%) was not reported. Of the 80 remaining original tools, only 18 tools (those in the satisfactory category) have adequate to excellent information relating to the factor analysis techniques used to determine underlying factor structures.

Various factor structures have been derived from empirical data collected using the 18 tools. In other words, different studies have found different factor structures for the same tool. This has been discussed in previous reports (Cox & Flin, 1998; Shannon & Norman, 2009) and is a well-recognised issue requiring further research. It has been hypothesised that the inability to replicate the factor structure of safety climate tools is, in part, due to incorrect handling of multi-level safety climate data (i.e., level-of-analysis issue) (Guldenmund, 2000; Shannon & Norman, 2009). Indeed, Håvold (2007) reported differences in the factor structure of his safety climate/safety culture tool: a 15-factor structure when safety climate data from individual workers were assumed to be completely independent and analysed at individual level and a four-factor structure when the same data were aggregated and analysed appropriately at group level. Cox and Flin (1998) have suggested that differences in factor structures may be partially explained by a lack of transferability of items across industries while Cooper and Phillips (2004) argued that these differences may stem from methodological dissimilarities across studies in terms of item generation, study population and factor labelling. Moreover, the interpretation of factor analysis results might have been affected by subjectivity in the evaluation of factor loadings, cross-factor loadings, and model fit statistics (Hinkin, 1995).

### *5.2.3. Reliability*

The internal consistency reliability of the scales that make up the 125 original tools was evaluated using Cronbach's alpha cut-off value of 0.7 as recommended by experts (Hinkin, 1995; MacKenzie et al., 2011). The tools in the satisfactory category have Cronbach's alpha values ≥0.7, thereby satisfying the criterion for reliability as applied in the current review. Some constituent scales, however, have Cronbach's alpha values exceeding 0.9, indicating that some items in these scales might be redundant. DeVellis (2003) suggests that scales with many items and very high Cronbach's alpha values might have some redundancy and could therefore be shortened.

### *5.2.4. Construct validation*

Adequate coverage of the conceptual domain, stability of the factor structure and reliability are necessary but not sufficient to establish that safety culture and safety climate tools actually measure what they are intended to measure (Hinkin, 1995; MacKenzie et al., 2011). A construct validation process is also required to provide evidence of construct validity and to compare the tools' performance with independent sources of information. This process was not reported for the majority of the 125 original tools, suggesting that it is important to obtain evidence of validity of a safety culture or climate tool before use. For tools that have construct validation details reported, in many cases the only details available are those associated with known-group validity, such as differences in safety climate scores between nationalities within a workplace (Håvold, 2007); differences in safety climate perceptions between workgroups (Cox & Cheyne, 2000). While these details are helpful in terms of mapping safety climate perceptions across workplaces, Cooper and Phillips (2004, p.498) argue that "sub-group differences within the same organization are a given" and "… merely inform about the degree to which the measure has reached its initial design goals".

In the current evaluation, the criterion-related validity of the majority of the tools in the satisfactory category has been established using subjective safety performance criteria. The use of subjective performance criteria may simply reflect difficulties in obtaining objective workplace injury and illness data for validation purposes. Nevertheless, given that subjective safety performance criteria might not compare well with objective safety performance criteria (Lenderink, Zoer, van der Molen, Spreeuwers, Frings-Dresen, & van Dijk, 2012), evidence for the criterion-related validity of most of the tools evaluated in the current review might be weaker than reported. This finding is consistent with previous reports (Cooper & Phillips, 2004; Mkrtchyan & Turcanu, 2012) and suggests that correlations between safety climate scores and subjective safety performance criteria might not necessarily establish criterion-related validity.

Another potential problem associated with subjective performance criteria is social-desirability biases, defined as biases resulting from a tendency by respondents to present themselves in a favourable light (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). Ganster, Hennessey, and Luthans (1983) suggest that social-desirability biases can produce spurious correlations, and/or obscure or moderate relationships between the predictor (safety climate) and the `criterion (safety performance). Obtaining data on the predictor and the criterion variables from different sources has been recommended as one of the ways by which to remedy these biases (Podsakoff et al., 2003). In fact, this recommendation has been adopted in the development and validation of a small number of safety climate tools in the satisfactory category with objective safety performance being used as the `criterion'. These tools have been shown to produce safety climate scores predictive of future workplace injury rates (e.g., Silva et al., 2004; Tharaldsen et al., 2008; Zohar, 2000).

In summary, the construct validation processes performed to date on safety culture and safety climate tools included in this evaluation have either largely focused on the technical attribute of these tools or evaluated their validity using subjective safety performance criteria. A full understanding of validity issues is necessary if safety culture or safety climate measurement is to be used to provide the basis for policy and practice decisions. In this regard, the requirement for tools to have at least one type of validity evidence in terms of the construct validation process was probably not ideal and appears to imply that all forms of validity are `created equal'. Nevertheless, this requirement was chosen taking into consideration the complexity and time-consuming nature of the construct validation process.

To further advance safety culture and safety climate measurement, future research should consider the question: All else being comparable, how can valid and reliable tools of varying lengths be compared and selected for use? It is currently not known how brief safety culture/safety climate tools would compare with longer tools in terms of performance and validity.

## 5.3. Implications for practice

This report is intended as a resource for regulators, people working in OHS, and other parties interested in the measurement of safety culture and safety climate. For anyone interested in using a safety climate tool in an organisation or workplace, there are numerous practical issues to consider. In this section, we identify some important questions to ask and offer some guidance to address these questions.

To select a tool for use in a workplace or organisational setting, some important questions to ask include:

- Can we obtain permission to use this tool?
- What is our aim in using a safety climate tool?
  - If the aim is to investigate relationships with other employee attitudes, then a safety climate tool could be included in a larger survey of the workforce.
  - If the aim is to measure improvements over time, then repeating the survey on a regular basis, e.g., annual basis would be advisable.
- Has the tool been validated for this type of workforce and/or workplace?
  - If not, consider incorporating validation process into your use of the tool.
- Do we have the resources available to use and analyse the results of a safety climate tool?

- Will the tool be suitable for our workforce and/or workplace?
    - Will a very long tool be too much of a burden? Will a very brief tool collect enough information?
    - Are the items appropriate for and relevant to our workforce and/or workplace? If not, can we make adaptations without reducing the reliability and validity of the tool?
- How can we encourage employees to participate in a safety climate survey so that we have a representative sample?

With regard to the first question an important consideration is copyright matters. The current review and evaluation focused on safety culture and safety climate tools that are publicly available for research purposes published between September 2013 and June 2014. It should be noted that there may be restrictions on the use of these tools for non-research purposes. It would be prudent for organisations wishing to use these tools to contact the author(s)/copyright holder(s) of these tools in the first instance.

With regard to the remaining questions, there are excellent resources available that offer practical guidance on the design and conduct of workforce surveys. Detailed information and advice about survey design and methods are beyond the scope of this review; however, the information can be obtained from references such as Dillman, Smyth, and Christian (2008). A brief overview of key practical considerations in survey design and implementation is offered below.

In addition to viewing the tools with regard to their reliability and validity, contextual information relating to their development and validation could provide useful guidance. This includes: the industry in which these tools have been validated; whether these tools have been developed for a specific worker population; study sample (characteristics and sample size); study design (cross sectional or longitudinal design); adaptations; data collection method (online survey, paper survey or interview); and whether scoring and data analysis guidelines are available.

An important consideration is whether or not an organisation or a workplace has in-house or can buy-in expertise to utilise these measures. Requisite expertise includes designing, planning and conducting safety culture or safety climate measurement, analysing the collected data and interpreting the results obtained. Depending on the tool(s) used, as well as factors such as the size of the participating workforce, number of work-groups, and organisational structure, potentially a large amount of psychometric data would be collected. Because of the hierarchical nature of the data—workers in workgroups, workgroups in work units, work units within work sites—appropriate data aggregation procedure and multilevel analyses would be required (Zohar, 2010). This would necessitate the involvement of personnel with expertise in psychometric techniques, such as testing for aggregation, testing for reliability and factor analysis. As discussed earlier, safety climate is estimated by aggregating individual worker evaluations of that climate and thus careful attention should be paid to conceptual and methodological issues in order to avoid improper cross-level inference and misspecification and aggregation biases (Guldenmund, 2000; Schneider, Ehrhart, & Macey, 2013; Zohar, 2003). In this respect the Nordic Occupational Safety Climate Questionnaire (NOSCQ) would appear to be a good choice because a data coding and analysis manual has been developed and is available on the internet. Nevertheless, an in-depth measurement of the safety climate or safety culture of an organisation or workplace would ideally require triangulation of psychometric data with qualitative data gained from direct observations, document analysis, focus group(s) and unstructured interviews (Guldenmund, 2000). These qualitative methods would provide context-specific information that enriches findings from psychometric data.

Dillman et al. (2008) describe a total design method for surveys. Dillman and his colleagues have identified steps that can be taken to improve the completion of a survey. For example, there are numerous techniques to encourage engagement by employees in the survey, to make the survey more acceptable to the target employees, and to improve response rates. Several of these areas are discussed below.

The acceptability of the measurement itself among these respondents also requires consideration. A pending measurement might raise fear, resistance and apprehension in these respondents and possibly the workforce at large, particularly when the measurement is restricted to certain workgroups or work areas (Health and Safety Executive, n.d.). In a `good' safety culture, workers should be able to speak up about safety issues, and report incidents and near misses and discuss lessons learned. The paradox is the presence of fear and apprehension would suggest a less than `ideal' safety culture and a rationale for the measurement. Preparations before a safety culture or safety climate measurement taking place are therefore crucial, not only to ensure that the correct methods and logistics are in place but also to enhance acceptability of the chosen tool(s) and the acceptance of the measurement among the workforce. This issue is discussed in guidelines and toolkits for safety culture and/or safety climate measurement (for instance The Keil Centre Ltd., 2003b; Health and Safety Executive, 2005a). A discussion about these guidelines is outside the scope of this review.

The acceptability of the chosen tool(s) among workers might be enhanced if workers are consulted about who should take part in the measurement, the wording of statements, items or questions in the chosen tool(s), and how the tool(s) could be administered to workers (hard copy, hand-held device or electronic survey) and how their confidentiality would be maintained (Health and Safety Executive, n.d.). Given that workplace hazards and possibly work practices vary by industry, by organisations within a particular industry and even by work areas within an organisation, it might be necessary to modify the wording of items in a tool to ensure relevance and appropriateness among the workforce of interest. This modification process would benefit from worker involvement (Health and Safety Executive, n.d.). Depending on the level of modification required, the modified tool(s) might have to be re-tested for reliability and validity (DeVellis, 2003).

A closely related issue is the demand of time placed on workers participating in the measurement (participant/respondent burden). To ensure that a sufficient number of workers would be willing to participate in a safety climate measurement, and hence an acceptable response rate, consideration should be given to the length of the questionnaire chosen, how this would affect response rate and what strategies can be used to address this limitation (Health and Safety Executive, n.d.; Hinkin, 1995). Tools with a large number of items/questions/statements might potentially deter participation because they are burdensome to complete (Hinkin, 1995). Strategies are therefore required to enhance response rates in measurements that use burdensome tools. Previous studies suggest that response rates would be higher when workers are provided with safety climate questionnaires to complete during safety training, safety meetings or work time (The Keil Centre, 2002; Zohar & Luria, 2005).

Response rates might also be enhanced if the chosen tools are available in languages other than English to facilitate participation by workers from culturally and linguistically diverse backgrounds (CALD). CALD workers now represent 27% of the Australian workforce (Australian Bureau of Statistics (ABS), 2010). It follows that appropriate language provision should be a key consideration in workplace measurement. Among the tools assessed by the current review to be valid and reliable for use, the NOSCQ-50 is the most convenient tool in this regard because it is available in more languages than any other tools (Kines et al., 2011).

Finally, it should be noted that not much is known about the sensitivity or responsiveness of the psychometric measures in the satisfactory category because the majority of these measures have been tested in cross-sectional and not longitudinal studies. The concept of responsiveness originates in the healthcare research domain; it indicates the ability of a measure to detect meaningful changes in the scores of individuals over time that reflect the effects of an intervention (Mokkink, Terwee, Knol, Stratford, Alonso, Patrick, et al., 2006). A measure's responsiveness to change is an important consideration for organisations planning to implement safety intervention(s) and subsequently evaluate changes over time in safety climate, safety culture and work-related injury and illness. It has been observed that a measure might be valid and reliable but shows weak responsiveness to change (Koopmans, Coffeng, Bernaards, Boot, Hildebrandt, de Vet, et al., 2014). Another issue to note is that empirical evidence regarding the effectiveness of safety interventions to prevent work-related injury and illness is scant. Moreover, where the evidence is available, it is mostly derived from studies with no control groups (Blewett & Flower, 2011; Singer & Vogus, 2013).

In summary, a variety of valid and reliable safety climate measures is available to organisations interested in safety climate/safety culture measurement. The decision-making process for selecting one of these measures for use in real world settings requires careful consideration of organisations' measurement needs and objectives, and theoretical and practical factors. The above discussion presents some key considerations for organisations wishing to measure safety culture and/or climate. Further information about survey methods can be obtained from references such as Dillman et al. (2008).

## 5.4. Strengths and limitations of this review

This review employed standard protocols for systematic reviews (Centre for Reviews and Dissemination (CRD), 2009). The search strategy was comprehensive and covered both the grey and published literature. Pre-defined selection criteria were used to select publications for inclusion in the evaluation. The evaluation of safety culture and safety climate tools that met the selection criteria was conducted using pre-defined evaluation criteria developed with advice from a technical advisory group with expertise in psychometric principles and knowledge of OHS.

Potential limitations of this review are associated with the inclusion of a limited number of websites in the search strategy, the use of certain key words in the literature search and the selection of measures for inclusion in the review mainly by one researcher. The review used certain key words to search for publications on safety climate/safety culture measures because it is a standardised approach to locate relevant publications. It is acknowledged, however, that early publications on safety climate and safety culture measures may not have been coded with the key words used in the search. A good case in point is the measure by Glendon et al. (1994) which does not have any of the key words used in the search and was not described as a safety climate measure. However, the manual search conducted on the reference lists of retrieved publications rectified this omission. Finally, even though the selection and evaluation of measures was mainly performed by one researcher, this process was systematic through the use of predefined criteria.

# 6. Use of This Report

The review and evaluation represent a comprehensive, useful resource for WSV specifically and for OHS practitioners, regulators and researchers. On the basis of this review, 18 tools developed to measure safety climate have been identified as fully meeting our evaluation criteria.

The question of whether a tool is available for purposes other than academic research is beyond the scope of this report. Persons or organisations wishing to use these tools are advised to contact the author(s)/copyright holder(s) in the first instance.

This review could be used in several ways, including:

- This review can be read alongside Vu and De Cieri's (2015b) snapshot review of the conceptual foundation of safety culture and climate measurement. The two reviews identify the conceptual challenges and methodological demands associated with safety culture and safety climate measurement. The reviews extend previous reviews of safety culture and safety climate tools and enhance the knowledge base for safety culture and safety climate measurement.

- This review could be a helpful guide for knowledge sharing and discussions among regulators and/or people working in OHS.

- People looking for a measure of safety climate to use in their workplace could inspect the satisfactory tools, consider the practical issues and questions identified in the discussion, and select a tool that would be suitable for their work context, goals, priorities and challenges. In addition to viewing the tools with regard to their reliability and validity, contextual information relating to their development and validation could provide useful guidance. This includes: the industry in which these tools have been validated; whether these tools have been developed for a specific worker population; study sample (characteristics and sample size); study design (cross sectional or longitudinal); adaptations; data collection method (online survey, paper survey or interview); and whether scoring and data analysis guidelines are available.

- People responsible for developing a safety culture and climate research agenda could identify gaps in current knowledge and issues for future research by examining the practical issues and questions raised in the discussion section.

Considering these usage potentials, this review and evaluation will facilitate safety culture and climate measurement and contribute to the evidence base for effective OHS.

# 7. References

Adcock, R. & Collier, D. (2001). Measurement validity: A shared standard for qualitative and quantitative research. *American Political Science review*, 95(3): 529-546.

Antonsen, S. (2009). Safety culture and the issue of power. *Safety Science, 47*(2), 183-191. doi: http://dx.doi.org/10.1016/j.ssci.2008.02.004

Australian Bureau of Statistics (ABS). (2010). Year Book Australia, 2009–10 Canberra: ABS.

Bergersen, C.E.B. (2003). *Tools to be used in survey and improve safety culture in the European railway industry.* (Diploma), Norwegian University of Science and Technology, Trondheim.

Bergh, M. (2011). *Safety climate: An evaluation of the safety climate at Akzonobel Site Stenungsund.* (Master's thesis), Chalmers University of Technology.

Bernard, B. (2014). Safety culture as a way of responsive regulation: Proposal for a nuclear safety culture oversight model. *International Nuclear Safety Journal, 3*(2), 1-11.

Beus, J., Payne, S., Bergman, M., & Arthur, W. (2010). Safety climate and injuries: An examination of theoretical and empirical relationships. *Journal of Applied Psychology, 95*, 713.

Birkbeck, D.J.L. (2010). *Investigating the relationship between LMX, safety climate and the components of safety performance in a high accident environment.* (Doctoral dissertation), University of Manchester, Manchester.

Blewett, V., & Flower, J. (2011). Best practice in OHS culture *Final draft report*: University of South Australia.

Borgersen, H.C., Hystad, S.W., Larsson, G., & Eid, J. (2013). Authentic Leadership and Safety Climate Among Seafarers. *Journal of Leadership & Organizational Studies*. doi: 10.1177/1548051813499612

Brondino, M., Pasini, M., & da Silva, S.C.A. (2013). Development and validation of an Integrated Organizational Safety Climate Questionnaire with multilevel confirmatory factor analysis. *Quality & Quantity, 47*(4), 2191-2223. doi: 10.1007/s11135-011-9651-6

Brondino, M., Silva, S.A., & Pasini, M. (2012). Multilevel approach to organizational and group safety climate and safety performance: Co-workers as the missing link. *Safety Science, 50*(9), 1847-1856.

Brown, K.A., Willis, P.G., & Prussia, G.E. (2000). Predicting safe employee behavior in the steel industry: Development and test of a sociotechnical model. *Journal of Operations Management, 18*, 445-465.

Cappelli, P., & Keller, J. (2013). Classifying work in the new economy. *Academy of Management Review, 38*(4), 575-596. doi: 10.5465/amr.2011.0302

Centre for Reviews and Dissemination (CRD). (2009). Systematic reviews: CRD's guidance for undertaking reviews in health care. York: University of York.

Chenhall, E.C. (2010). *Assessing safety culture, values, practices, and outcomes.* (Doctoral dissertation), Colorado State University, Fort Collins, Colorado.

Christian, M.S., Bradley, J.C., Wallace, J.C., & Burke, M.J. (2009). Workplace safety: A meta-analysis of the roles of person and situation factors. *Journal of Applied Psychology, 94*(5), 1103-1127.

Clarke, S., & Ward, K. (2006). The role of leader influence tactics and safety climate in engaging employees' safety participation. *Risk Analysis, 26*(5), 1175-1185. doi: 10.1111/j.1539-6924.2006.00824.x

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences. 2nd Edition*. New Jersey: Lawrence Erlbaum.

Cooper, D. (2000). Towards a model of safety culture. *Safety Science, 36*(2), 111-136.

Cooper, D. (2002). *Surfacing your safety culture*. Paper presented at the Major Hazard Commission at the Federal Ministry of Environment: Human Factors Conference, 4-6 March, Loccum, Germany.

Cooper, D., & Phillips, R.A. (2004). Exploratory analysis of the safety climate and safety behavior relationship. *Journal of Safety Research, 35*(5), 497-512.

Cox, S., & Flin, R. (1998). Safety culture: Philosopher's stone or man of straw? *Work & Stress, 12*(3), 189-201. doi: 10.1080/02678379808256861

Cox, S.J., & Cheyne, A.J.T. (2000). Assessing safety culture in offshore environments. *Safety Science, 34*(1–3), 111-129.

Coyle, I.R., Sleeman, S.D., & Adams, N. (1995). Safety climate. *Journal of Safety Research, 26*(4), 247-254. doi: http://dx.doi.org/10.1016/0022-4375(95)00020-Q

Cree, T., & Kelloway, E.K. (1997). Responses to occupational hazards: exit and participation. *J Occup Health Psychol, 2*(4), 304-311.

Cronbach, L. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297-334. doi: 10.1007/BF02310555

Darling, C., Edkins, G., Glendon, I., Lee, R., Lewis, K., & Thompson, N. (2004). Safety management systems review of RailCorp and the NSW Independent Transport Safety and Reliability Regulatory. Sydney: Special commission of inquiry into the Waterfall rail accident.

Davies, F., Spencer, R., & Dooley, K. (2001). Summary guide to safety climate tools. *Offshore Technology Report 1999/063, Health and Safety Executive*.

DePasquale, J.P., & Geller, E.S. (1999). Critical success factors for behavior-based safety: A study of twenty industry-wide applications. *Journal of Safety Research, 30*(4), 237-249. doi: http://dx.doi.org/10.1016/S0022-4375(99)00019-5

DeVellis, R.F. (2003). *Scale development: Theory and applications*. Newbury Park, California: SAGE Publications.

Díaz-Cabrera, D., Hernández-Fernaud, E., & 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. doi: http://dx.doi.org/10.1016/j.aap.2007.03.005

Dillman, D.A., Smyth, J.D., & Christian, L.M. (2008). *Internet, mail, and mixed-mode surveys: The tailored design method*: John Wiley & Sons.

Edwards, J.R.D., Davey, J., & Armstrong, K. (2013). Returning to the roots of culture: A review and re-conceptualisation of safety culture. *Safety Science, 55*(0), 70-80. doi: http://dx.doi.org/10.1016/j.ssci.2013.01.004

Eeckelaert, L., Starren, A., van Scheppingen, A., Fox, D., & Brück, C. (2011). Occupational safety and health culture assessment - A review of main approaches and selected tools. Luxembourg: European Agency for Safety and Health at Work (EU-OSHA).

Evans, B., Glendon, A.I., & Creed, P.A. (2007). Development and initial validation of an Aviation Safety Climate Scale. *Journal of Safety Research, 38*(6), 675-682. doi: http://dx.doi.org/10.1016/j.jsr.2007.09.005

Fang, D., Chen, Y., & Wong, L. (2006). Safety climate in construction industry: A case study in Hong Kong. *Journal of Construction Engineering and Management, 132*(6), 573–584.

Fernández-Muñiz, B., Montes-Peón, J.M., & Vázquez-Ordás, C.J. (2007). Safety culture: Analysis of the causal relationships between its key dimensions. *Journal of Safety Research, 38*(6), 627-641. doi: http://dx.doi.org/10.1016/j.jsr.2007.09.001

Fleming, M. (2009). Developing safety culture measurement tools and techniques based on site audits rather than questionnaires *Final project report*: Saint Mary's University, Halifax, Nova Scotia.

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., & Bryden, R. (2000). Measuring safety climate: Identifying the common features. *Safety Science, 34*(1–3), 177-192.

Ganster, D.C., Hennessey, H.W., & Luthans, F. (1983). Social desirability response effects: Three alternative models. *Academy of Management Journal, 26*(2), 321-331. doi: 10.2307/255979

Geddes, F.R. (2012). *Managers, mates and the role of social exchange : A multilevel model of safety climate and proactive safety behaviour.* (Doctoral dissertation), Curtin University, Perth.

Gilkey, D.P., Puerto, C.L.d., Keefe, T., Bigelow, P., Herron, R., Rosecrance, J., & Chen, P. (2012). Comparative analysis of safety culture perceptions among homesafe managers and workers in residential construction. *Journal of Construction Engineering and Management, 138*(9), 1044-1052.

Glendon, A.I., & Litherland, D.K. (2001). Safety climate factors, group differences and safety behaviour in road construction. *Safety Science, 39*(3), 157-188. doi: http://dx.doi.org/10.1016/S0925-7535(01)00006-6

Glendon, A.I., Stanton, N.A., & Harrison, D. (1994). *Factor analysing a performance shaping concepts questionnaire.* Paper presented at the Ergonomics Society's 1994 Annual Conference, University of Warwick, London.

Griffin, M.A., & Neal, A. (2000). Perceptions of safety at work: A framework for linking safety climate to safety performance, knowledge, and motivation. *Journal of Occupational Health Psychology, 5*(3), 347-358.

Guldenmund, F.W. (2000). The nature of safety culture: A review of theory and research. *Safety Science, 34*(1-3), 215-257.

Guldenmund, F.W. (2007). The use of questionnaires in safety culture research – an evaluation. *Safety Science, 45*(6), 723-743. doi: http://dx.doi.org/10.1016/j.ssci.2007.04.006

Harvey, J., Erdos, G., Bolam, H., Cox, M.A.A., Kennedy, J.N.P., & Gregory, D.T. (2002). An analysis of safety culture attitudes in a highly regulated environment. *Work & Stress, 16*(1), 18-36.

Håvold, J.I. (2007). National cultures and safety orientation: A study of seafarers working for Norwegian shipping companies. *Work & Stress: An International Journal of Work, Health & Organisations, 21*(2), 173-195.

Hayes, B.E., Perander, J., Smecko, T., & Trask, J. (1998). Measuring perceptions of workplace safety: Development and validation of the Work Safety Scale. *Journal of Safety Research, 29*(3), 145-161.

Haynes, S.N., Richard, D.C.S., & Kubany, E.S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment, 7*(3), 238-247. doi: 10.1037/1040-3590.7.3.238

Health and Safety Commission (HSC). (1993). ACSNI Humans Factors Study Group. Third report: Organising for safety. London: Health and Safety Commission.

Health and Safety Executive. (2005a). Development and validation of the HMRI safety culture inspection toolkit. Bristol: Her Majesty's Railway Inspectorate.

Health and Safety Executive. (2005b). A review of safety culture and safety climate literature for the development of the safety culture inspection toolkit. Bristol: Her Majesty's Railway Inspectorate.

Health and Safety Executive. (n.d.). Safety climate measurement. User guide and toolkit: Health and Safety Executive.

Health and Safety Laboratory. (2013). Key Features of the SCT, Retrieved from http://www.hsl.gov.uk/products/safety-climate-tool/key-features-of-the-sct on 11/06/2013. Retrieved from

Heese, M. (2012). Got the results, now what do you do? Safety culture transformation from theory into practice. *Aviation Psychology and Applied Human Factors, 2*(1), 25-33. doi: 10.1027/2192-0923/a000020

Heritage, K.M. (2012). *Lagged relationships between a multilevel model of safety climate and employee safety outcomes.* (Doctoral dissertation), Curtin University, Perth.

Hinkin, T.R. (1995). A review of scale development practices in the study of organizations. *Journal of Management, 21*(5), 967-988.

Høivik, D., Tharaldsen, J.E., Baste, V., & Moen, B.E. (2009). What is most important for safety climate: The company belonging or the local working environment? – A study from the Norwegian offshore industry. *Safety Science, 47*(10), 1324-1331. doi: http://dx.doi.org/10.1016/j.ssci.2009.04.001

Hope, S., Øverland, S., Brun, W., & Matthiesen, S.B. (2010). Associations between sleep, risk and safety climate: A study of offshore personnel on the Norwegian continental shelf. *Safety Science, 48*(4), 469-477. doi: http://dx.doi.org/10.1016/j.ssci.2009.12.006

Hsu, S.H., Lee, C.-C., Wu, M.-C., & Takano, K. (2008). A cross-cultural study of organizational factors on safety: Japanese vs. Taiwanese oil refinery plants. *Accident Analysis and Prevention, 40*(1), 24-34.

Huang, Y.-H., Chen, J.-C., DeArmondc, S., Cigularov, K., & Chen, P.Y. (2007). Roles of safety climate and shift work on perceived injury risk: A multi-level analysis. *Accident Analysis and Prevention, 39*(6), 1088-1096.

Huang, Y.-H., Zohar, D., Robertson, M.M., Garabet, A., Lee, J., & Murphy, L.A. (2013a). Development and validation of safety climate scales for lone workers using truck drivers as exemplar. *Transportation Research Part F, 17*(February), 5-19.

Huang, Y.H., Ho, M., Smith, G.S., & Chen, P.Y. (2006). Safety climate and self-reported injury: Assessing the mediating role of employee safety control. *Accident Analysis and Prevention, 38*(3), 425-433. doi: 10.1016/j.aap.2005.07.002

Independent Transport Safety Regulator (ITSR) & Transport Safety Victoria (TSV). (2012). Organisational safety culture appraisal tool (OSCAT). Victoria.

International Nuclear Safety Advisory Group. (1986). Summary report on the post-accident review meeting on the Chernobyl accident *Safety Series No.75-INSAG-l*. Vienna: International Atomic Enerygy Agency.

James, L. (1982). Aggregation bias in estimates of perceptual agreement. *Journal of Applied Psychology, 67*(2), 219-229.

Johnson, S.E. (2007). The predictive validity of safety climate. *Journal of Safety Research, 38*(5), 511-521. doi: http://dx.doi.org/10.1016/j.jsr.2007.07.001

Kaplan, D. (Ed.). (2004). *The SAGE handbook of quantitative methodology for the social sciences*. Thousand Oaks, California: Sage Publications, Inc.

Kaplan, R.S., & Norton, D.P. (1991). The balanced scorecard - Measures that drive performance. *Harvard Business Review*(January-February), 71-79.

Kines, P., Lappalainen, J., Mikkelsen, K.L., Olsen, E., Pousette, A., Tharaldsen, J., . . . Törner, M. (2011). Nordic Safety Climate Questionnaire (NOSACQ-50): A new tool for diagnosing occupational safety climate. *International Journal of Industrial Ergonomics, 41*(6), 634-646.

Koopmans, L., Coffeng, J.K., Bernaards, C.M., Boot, C.R.L., Hildebrandt, V.H., de Vet, H.C.W., & van der Beek, A.J. (2014). Responsiveness of the individual work performance questionnaire. *BMC Public Health, 14*, 513-513. doi: 10.1186/1471-2458-14-513

Law, K.S., Wong, C.-S., & Mobley, W.M. (1998). Toward a taxonomy of multidimensional constructs. *Academy of Management Review, 23*(4), 741-755. doi: 10.5465/amr.1998.1255636

Lawrie, M., Parker, D., & Hudson, P. (2006). Investigating employee perceptions of a framework of safety culture maturity. *Safety Science, 44*(3), 259-276.

Lee, T., & Harrison, K. (2000). Assessing safety culture in nuclear power stations. *Safety Science, 34*(1–3), 61-97.

Lenderink, A., Zoer, I., van der Molen, H., Spreeuwers, D., Frings-Dresen, M.W., & van Dijk, F.H. (2012). Review on the validity of self-report to assess work-related diseases. *Int Arch Occup Environ Health, 85*(3), 229-251. doi: 10.1007/s00420-011-0662-3

Lingard, H., Cooke, T., & Blismas, N. (2009). *Group safety climate in the construction industry: An analysis of strength and level.* Paper presented at the CIB W099 International Conference 'Working together: Planning, Designing and Building a Healthy and Safe Industry, Melbourne.

Lingard, H., Cooke, T., & Blismas, N. (2012). Do perceptions of supervisors' safety responses mediate the relationship between perceptions of the organizational safety climate and incident rates in the construction supply chain? *Journal of Construction Engineering and Management, 138*(2), 234-241.

Luria, G. (2010). The social aspects of safety management: Trust and safety climate. *Accident Analysis and Prevention, 42*(4), 1288-1295.

MacKenzie, S.B., Podsakoff, P.M., & Podsakoff, N.P. (2011). Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques. *Management Information Systems Quarterly, 35*(2), 293-334.

Martínez-Córcoles, M., Gracia, F., Tomás, I., & Peiró, J.M. (2011). Leadership and employees' perceived safety behaviours in a nuclear power plant: A structural equation model. *Safety Science, 49*(8–9), 1118-1129. doi: http://dx.doi.org/10.1016/j.ssci.2011.03.002

Mearns, K., Whitaker, S.M., Flin, R., Gordon, R., & O'Connor, P. (2003). Factoring the human into safety: Translating research into practice *Research report* (Vol. 1 of 3). London: Health and Safety Executive.

Meliá, J.L., & Sesé, A. (2007). Supervisor's safety response: A multisample confirmatory factor analysis. *Psicothema, 19*(2), 231-238.

Mkrtchyan, L., & Turcanu, C. (2012). Safety culture assessment tools in nuclear and non-nuclear domains *Open report*. Belgium: Nuclear Science and Technology Studies (NST).

Mokkink, L.B., Terwee, C.B., Knol, D.L., Stratford, P.W., Alonso, J., Patrick, D.L., . . . de Vet, H.C. (2006). Protocol of the COSMIN study: COnsensus-based Standards for the selection of health Measurement INstruments. *BMC Med Res Methodol, 6*, 2. doi: 10.1186/1471-2288-6-2

Morrow, S.L., & Barnes, V. (2012). Independent evaluation of INPO's nuclear safety culture survey and construct validation study *Technical paper*: Division of Risk Analysis, Office of Nuclear Regulatory Research.

Nahrgang, J.D., Morgeson, F.P., & Hofmann, D.A. (2010). Safety at work: A meta-analytic investigation of the link between job demands, job resources, burnout, engagement, and safety outcomes. *Journal of Applied Psychology, 96*(1), 71-94. doi: 10.1037/a0021484

National Occupational Research Agenda (NORA) Construction Sector Council. (2008). National Construction Agenda: National Institute for Occupational Safety and Health

National Safety Council of Australia Ltd. (n.d.). Safety Climate Survey. *Retrieved from http://nsca.org.au/products-services/safety-climate-survey/ on 26 may 2014*. Retrieved from

Navarro, M.F.L., Lerín, F.J.G., Tomás, I., & Silla, J.M.P. (2013). Validation of the group nuclear safety climate questionnaire. *Journal of Safety Research, 46*(September), 21-30.

Nazaruk, M. (2011). *Developing safety culture interventions in the manufacturing sector.* (Doctoral dissertation), University of Bath, Bath, United Kingdom.

Neal, A., & Griffin, M. (2006). A study of the lagged relationships among safety climate, safety motivation, safety behavior, and accidents at the individual and group levels. *Journal of Applied Psychology, 91*(4), 946-953.

Neal, A., Griffin, M.A., & Hart, P.M. (2000). The impact of organizational climate on safety climate and individual behavior. *Safety Science, 34*(1–3), 99-109.

Nielsen, K. (2014). Improving safety culture through the health and safety organization: A case study. *Journal of Safety Research, 48*, 7-17.

Nielsen, M.B., Eid, J., Hystad, S.W., Sætrevik, B., & Saus, E.-R. (2013). A brief safety climate inventory for petro-maritime organizations. *Safety Science, 58*(0), 81-88. doi: http://dx.doi.org/10.1016/j.ssci.2013.04.002

Nielsen, M.B., Eid, J., Mearns, K., & Larsson, G. (2013). Authentic leadership and its relationship with risk perception and safety climate. *Leadership & Organization Development Journal, 34*(308-325).

Ostrom, L., Wilhelmsen, C., & Kaplana, B. (1993). Assessing safety culture. *Nuclear Safety, 34*(2), 163-172.

Podsakoff, P.M., MacKenzie, S.B., Lee, J.Y., & Podsakoff, N.P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *J Appl Psychol, 88*(5), 879-903. doi: 10.1037/0021-9010.88.5.879

Probst, T.M. (2004). Safety and insecurity: Exploring the moderating effect of organizational safety climate. *J Occup Health Psychol, 9*(1), 3-10. doi: 10.1037/1076-8998.9.1.3

Prussia, G.E., Brown, K.A., & Willis, P.G. (2003). Mental models of safety: Do managers and employees see eye to eye? *J Safety Res, 34*(2), 143-156.

Reason, J. (2001). Score your safety culture. *Flight Safety Australia Magazine*.

Rundmo, T., & Hale, A.R. (2003). Managers' attitudes towards safety and accident prevention. *Safety Science, 41*(7), 557-574. doi: http://dx.doi.org/10.1016/S0925-7535(01)00091-1

Safe Work Australia. (2012). Australian Work Health and Safety Strategy 2012–2022. Canberra, Australia.: Safe Work Australia.

Schein, E.H. (1985). *Organizational culture and leadership*. San Francisco: Jossey-Bass Publishers.

Schneider, B., Ehrhart, M.G., & Macey, W.H. (2013). Organizational Climate and Culture. *Annual Review of Psychology, 64*(1), 361-388. doi: doi:10.1146/annurev-psych-113011-143809

Seo, D.C., Torabi, M.R., Blair, E.H., & Ellis, N.T. (2004). A cross-validation of safety climate scale using confirmatory factor analytic approach. *Journal of Safety Research, 35*(4), 427-445. doi: 10.1016/j.jsr.2004.04.006

Shannon, H.S., & Norman, G.R. (2009). Deriving the factor structure of safety climate scales. *Safety Science, 47*(3), 327-329. doi: http://dx.doi.org/10.1016/j.ssci.2008.06.001

Shaw, L. (2013). Are we ready to address the new expectations of work and workers in the transforming world of work? *Work, 44*(1), 3-9. doi: 10.3233/wor-121582

Silbey, S.S. (2009). Taming Prometheus: Talk about safety and culture. *Annual Review of Sociology, 35*(1), 341-369. doi: doi:10.1146/annurev.soc.34.040507.134707

Silva, S.l., Lima, M.L.s., & Baptista, C. (2004). OSCI: An organisational and safety climate inventory. *Safety Science, 42*(3), 205-220. doi: http://dx.doi.org/10.1016/S0925-7535(03)00043-2

Sinclair, R.R., Martin, J.E., & Sears, L.E. (2010). Labor unions and safety climate: Perceived union safety values and retail employee safety outcomes. *Accident Analysis and Prevention, 42*(5), 1477-1487. doi: 10.1016/j.aap.2009.11.003

Singer, S.J., & Vogus, T.J. (2013). Safety climate research: Taking stock and looking forward. *BMJ Quality and Safety, 22*(1), 1-4. doi: 10.1136/bmjqs-2012-001572

Tharaldsen, J.E., Olsen, E., & Rundmo, T. (2008). A longitudinal study of safety climate on the Norwegian continental shelf. *Safety Science, 46*(3), 427-439. doi: http://dx.doi.org/10.1016/j.ssci.2007.05.006

The Keil Centre. (2002). Evaluating the effectiveness of the Health and Safety Executive's Health and Safety Climate Survey Tool *Research Report*. London: Health and Safety Executive.

The Keil Centre Ltd. (2003a). Managing safety culture in the UK rail industry: Report on the review of safety culture tools and methods *Research Report, Rail Safety and Standards Board*.

The Keil Centre Ltd. (2003b). Measurement of safety culture in the rail industry *Research Report*. London: Rail Safety and Standards Board.

Tietjen, M.A., & Myers, R.M. (1998). Motivation and job satisfaction. *Management Decision, 36*(4), 226-231. doi: doi:10.1108/00251749810211027

Vu, T., & De Cieri, H. (2013). Defining safety culture from a regulator's perspective: A scoping study. Melbourne, Australia: Institute for Safety, Compensation and Recovery Research (ISCRR), Monash University.

Vu, T., & De Cieri, H. (2014). Safety culture and safety climate definitions suitable for a regulator: A systematic literature review. Melbourne, Australia: Institute for Safety, Compensation and Recovery Research (ISCRR), Monash University, Research Report #0414-060-R2C.

Vu, T., & De Cieri, H. (2015). Conceptual foundations of safety culture and safety climate measurement: A snapshot review. Melbourne, Australia: Institute for Safety, Compensation and Recovery Research (ISCRR), Monash University, Research Report 060-1015-R04.

Wallace, A.M. (2005). *Climate and behaviour: The role of affective experiences in the safety context.* (Master's thesis), University of Queensland.

Wallace, J.C., Popp, E., & Mondore, S. (2006). Safety climate as a mediator between foundation climates and occupational accidents: A group-level investigation. *Journal of Applied Psychology, 91*(3), 681-688.

Weick, K.E., & Sutcliffe, K.M. (2001). *Managing the unexpected: Assuring high performance in an age of complexity* (First ed.). San Francisco: JosseyBass.

Wiegmann, D.A., Zhang, H., von Thaden, T., Sharma, G., & Mitchell, A. (2002). A synthesis of safety culture and safety climate research. Savoy: Aviation Research Lab, Institute of Aviation.

Wilkin, C.L. (2013). I can't get no job satisfaction: Meta-analysis comparing permanent and contingent workers. *Journal of Organizational Behavior, 34*(1), 47-64. doi: 10.1002/job.1790

Williamson, A.M., Feyer, A.-M., Cairns, D., & Biancotti, D. (1997). The development of a measure of safety climate: The role of safety perceptions and attitudes. *Safety Science, 25*(1–3), 15-27.

WorkCover NSW. (n.d.). Safety culture survey. Gosford: WorkCover NSW.

Workplace Health and Safety Queensland. (2013). Understanding safety culture: The State of Queensland, Department of Justice and Attorney-General.

Wu, T.-C., Lin, C.-H., & Shiau, S.-Y. (2010). Predicting safety culture: The roles of employer, operations manager and safety professional. *Journal of Safety Research, 41*(5), 423-431. doi: http://dx.doi.org/10.1016/j.jsr.2010.06.006

Yule, S. (2003). *Senior management influence on safety performance in the UK and US energy sectors.* (Doctoral dissertation), University of Aberdeen, Aberdeen.

Zacharatos, A., Barling, J., & Iverson, R.D. (2005). High-performance work systems and occupational safety. *J Appl Psychol, 90*(1), 77-93. doi: 10.1037/0021-9010.90.1.77

Ziegler, R., Hagen, B., & Diehl, M. (2012). Relationship between job satisfaction and job performance: Job ambivalence as a moderator. *Journal of Applied Social Psychology, 42*(8), 2019-2040. doi: 10.1111/j.1559-1816.2012.00929.x

Zohar, D. (1980). Safety climate in industrial organizations: Theoretical and applied implications. *Journal of Applied Psychology, 65*(1), 96-102.

Zohar, D. (2000). A group level model of safety climate: testing the effect of group climate on microaccidents in manufacturing jobs. *Journal of Applied Psychology, 85*(4), 587-596.

Zohar, D. (2002a). The effects of leadership dimensions, safety climate, and assigned priorities on minor injuries in work groups. *Journal of Organizational Behavior, 23*(1), 75-92. doi: 10.1002/job.130

Zohar, D. (2002b). Modifying supervisory practices to improve subunit safety: A leadership-based intervention model. *Journal of Applied Psychology, 87*(1), 156-163.

Zohar, D. (2008). Safety climate and beyond: A multi-level multi-climate framework. *Safety Science, 46*(3), 376-387. doi: http://dx.doi.org/10.1016/j.ssci.2007.03.006

Zohar, D. (2010). Thirty years of safety climate research: Reflections and future directions. *Accident Analysis and Prevention, 42*(5), 1517-1522. doi: 10.1016/j.aap.2009.12.019

Zohar, D., & Luria, G. (2003). The use of supervisory practices as leverage to improve safety behavior: A cross-level intervention model: Technion Institute of Technology, Haifa, Israel & Institute for Work and Health, Toronto, Canada.

Zohar, D., & Luria, G. (2004). Climate as a social-cognitive construction of supervisory safety practices: Scripts as proxy of behavior patterns. *Journal of Applied Psychology, 89*(2), 322-333. doi: 10.1037/0021-9010.89.2.322

Zohar, D., & Luria, G. (2005). A multi-level model of safety climate: Cross-level relationships between organization and group-level climates. *Journal of Applied Psychology, 90*(4), 616-628.