Comparative analysis of existing food safety culture evaluation systems

Author names and affiliations

Lone Jespersen,[1], Dr. Mansel Griffiths,[1] and Prof. Carol A Wallace[2].

[1] University of Guelph, 50 Stone Road East, Guelph, ON N1G 2M7, Canada

[2] University of Central Lancashire, International Institute of Nutritional Sciences and Applied Food Safety Studies, Preston, Lancashire PR1 2HE, United Kingdom.

Corresponding author

Lone Jespersen, University of Guelph, 50 Stone Road East, Guelph, ON, N1G 2M7, Canada.

+41792460807. lone@uoguelph.ca

Abstract

The purpose of the research was firstly, to analyze existing culture evaluation systems for commonalities and differences in research quality, applied validation strategies, and content. Secondly, to suggest a simple structure of food safety cultural dimensions to help unify the culture evaluation field. To achieve these goals, a comparison of eight culture evaluation models applied to varing degrees in the food industry was conducted. The systems were found to vary significantly in applied validation strategies but through deductive, textual data analysis, five dimensions were identified that cover elements present in all the models. Transparency is needed when using applied research methodologies to continually increase quality and trustworthiness of culture research in the food safety domain and this field would benefit from both further commonality of approach to validation strategy and structure and adoption of an overarching structural framework.

Keywords

Food safety culture, research quality, trustworthiness, cultural dimensions, culture evaluation

Highlights

- The research discusses standards and guidelines for evaluating research quality and trustworthiness
- The research compares eight models for evaluating culture for validation strategies and content
- It is discovered that common validation techniques are applied but that only two methods make use of predictive validation.
- Based on a qualitative content analysis of each model a suggested framework of five cultural dimensions are proposed to unify the research field.

1.0 Introduction

The problem of food safety culture – what is it and how do you know how good yours is– is probably one of the main issues in modern thinking about food safety (Christopher James Griffith, 2010; Jespersen, Griffiths, Maclaurin, Chapman, & Wallace, 2016; Jespersen & Huffman, 2014; Nyarugwe, Linnemann, Hofstede, Fogliano, & Luning, 2016; Powell, Jacob, & Chapman, 2011; Yiannas, 2009). Although it is now subject to much discussion, the concept of food safety culture is still poorly understood. Incidents that prompted attention to food safety culture include, but are not limited to, the listeriosis outbreak 2008– Canada (Canadian Food Inspection, 2013), Melamine posioning 2007 – China (Gossner et al., 2009; Ingelfinger 2008), EHEC outbreak 2011 – Germany (Bernd Appel, 2011; Weiser et al., 2016), *Clostridium botulinum*, 2013 – New Zealand (Incident, 2014), John Barr 1996 and J.E. Tudor 2005 EHEC outbreaks, U.K. (H. Pennington, 2009; T. H. Pennington, 2014).

Culture, be it organisational, employee health and safety or food safety culture, can be described through Schein's organizational culture definition as *a pattern of shared basic assumptions that was learned by a group as it solved its problems of external adaptation and internal integration, that has worked well enough to be considered valid and, therefore, to be taught to new members as the correct way to perceive, think, and feel in relation to those problems* (Schein, 2004). As such, culture is formulated in groups, the same groups who often make decisions on how to strengthen culture and invest their resources based on attitudes and assumptions and perhaps the results of culture evaluation systems. Food safety culture has been defined as *the aggregation of the prevailing, relatively constant, learned, shared attitudes, values and beliefs contributing to the hygiene behaviours used in a particular food handling environment* (C. J. Griffith, K. M. Livesey, & D. Clayton, 2010a). The definition relates cultural values, beliefs, and learned hygiene behaviour to food. It is often a breakdown of learned and shared attitudes, values, and beliefs that causes deadly failures in food safety management systems and for such

failures to become embedded in food safety in the food manufacturer's psyche. Evaluating culture is one approach to bring transparency to cultural strengths and weaknesses, which may help prevent consumer illness and mortality due to foodborne illness. Some food manufacturers use cultural evaluation systems to avoid consumers from being exposed to foodborne hazards and for brand protection and employment security (Cameron, 2006; Jespersen & Huffman, 2014; Seward, 2012). Thus, if cultural evaluation systems are used in this way, it is necessary to determine the trustworthiness of results against accepted methods to assess their validity and reliabililty.

Current systems for evaluating culture have been referred to as fragmented and built on disparate scientific theories (De Boeck, Jacxsens, Bollaerts, & Vlerick, 2015; F. W. Guldenmund, 2000). Many of these evaluation systems are developed using qualitative research methodologies. There is an ongoing need for qualitative research to be demonstrated as a scientifically-based, learned, and robust methodology and this can only be done by recording, systematizing, and disclosing methods of analysis (Attride-Stirling, 2001). It is a concern that some researchers in the culture domain, embrace qualitative methods but do not provide adequate insight into, nor justifications for selected methods, findings, or conclusions; not unlike the approaches taken in other domains, e.g., educational research and employee health and safety (F. W. Guldenmund, 2000; Howe & Eisenhart, 1990)

The objectives of this research were, firstly, to analyze existing culture evaluation systems for commonalities and differences in research quality, applied validation strategies, and content; secondly, to suggest a simple structure of food safety cultural dimensions to help unify the culture evaluation field.

2.1 Research quality

The National Research Council (NRC) and others have described guidelines that shape scientific understanding and that are frequently used to frame the discourse on the quality of research. This has led

to the term ‘scientifically-based research’ being used in some settings to address research quality (Gersten, 2000; Greenhalgh, 1997). Principles for assessing the quality of research include the following:

1. Pose a significant, important question that can be investigated empirically and that contributes to the knowledge base;
2. Test questions that are linked to relevant theory;
3. Apply methods that best address the research questions of interest;
4. Base research on clear chains of inferential reasoning supported and justified by a complete coverage of the relevant literature;
5. Provide the necessary information to reproduce or replicate the study;
6. Ensure the study design, methods, and procedures are sufficiently transparent and ensure an independent, balanced, and objective approach to the research;
7. Provide sufficient description of the sample, the intervention, and any comparison groups;
8. Use appropriate and reliable conceptualization and measurement of variables;
9. Evaluate alternative explanations for any findings;
10. Assess the possible impact of systemic bias;
11. Submit research to a peer-review process;
12. Adhere to quality standards for reporting (i.e., clear, cogent, complete).

While there is no consensus on a specific set of guidelines that will ensure the quality of research, the more research studies are aligned with or respond to these principles, the higher will be the value of the research (Feuer, 2002; Richard J. Shavelson and Lisa Towne, 2002).

1.2 Research trustworthiness

Research should be as trustworthy as possible and every research study must be evaluated in relation to the methods used to generate the results. Describing trustworthiness of qualitative research is different than that of quantitative research. Some believe alternative terms are required to evaluate qualitative research (Graneheim, 2004). Regardless of research type, ambiguous or meaningless findings may result in wasted time and effort, while findings that are simply wrong could result in adoption of

dangerous and harmful practices (Long and Johnson, 2000). Untrustworthy research can be caused by error and bias related to both participants and observers (Robson, 2011) and it is the responsibility of the researcher to have a rigorous and transparent validation strategy to eliminate untrustworthy results. Krippendorf reflects on the trustworthyness of scientiffcally-based research by stating *quoting from the works of other scholars cannot absolve anyone from the responsibilities for investigating and judging what they thereby enter into literature* (Krippendorf, 2004).

1.3 Valid and reliable research

The qualitative researcher must constantly search for techniques to demonstrate rigour of the research process to ensure trustworthiness and usefulness of research findings and to avoid misleading those who use the outcomes of the research (Louis Cohen, 2007; Roberts, 2006). Applying traditional quantitative tests of validity and reliability to qualitative research methods can be difficult, as the subjectivity of respondents' and observers' opinions, attitudes, and perspectives contribute to a degree of bias. Thus, the validity of qualitative research must be seen in relative rather than absolute terms (Gronlund, 1990). This suggests that the quality of trial design could influence the reliability of the final results, which are crucial for their interpretation and subsequent recommendations and implementation (Armijo-Olivo, 2012).

Results and inferences made from any culture evaluation methodology are impacted by the validity and reliability of the research. Researchers must strive to balance the impossibility of reaching 100% validity with avoidance of untrustworthy, invalid research (Louis Cohen, 2007). Validity indicates the system's accuracy; whereas reliability indicates the system's ability to produce consistent and repeatable results (Trochim, 2006). Reliability measures are important parts of the system's overall validity. Reliability is necessary but cannot be considered in isolation. For a system to be reliable, it must be valid. A system can be reliable e.g., through internal consistency but not valid but mostly if it is valid it

is more often than not reliable (Louis Cohen, 2007). However, there is some disagreement with these statements as Robson states that *unless a measure is reliable, it cannot be valid* (Robson & Robson, 2011). There is general concensus that reliability is necessary but not sufficient. Reliability can also be seen to relate to the coherence theory of truth in social research, i.e. a statement is considered a true representation of a socially constructed reality when it is confirmed by several reports (Richie, 2003) . The concepts of validity and reliability have great importance for evaluating the trustworthiness of any results and inference generated by a culture evaluation system (Robson & Robson, 2011). The applicability of different validity and reliability measures is directly linked to the type of research conducted. It is therefore important, when developing a culture evaluation system by which organizations plan to introduce change, to consider the full research process including e.g., research questions, data collection, and data analysis, before designing the final research study, selecting methods, and the specific validation and reliability measures necessary to ensure trustworthiness and usefulness of the results (Louis Cohen, 2007; Meyrick, 2006; Robson & Robson, 2011; T. Long, 2000).

1.3.1 Validation and validity

Validity theory has evolved over time (Shepard, 1993, 2016) and it is important to note that many have engaged in the discussion on defining and selecting the most appropriate validation measures and concluded that not one approach fits all situations. It is not the intent of this paper to give a comprehensive review of all validity and reliability measures but a broad enough view to compare and contrast validity of existing culture evaluation systems. Concepts of measuring validity have been applied to various fields of research and it is clear how quantifiable validation tests are a fit with quantitative scientific research but it is less clear what validity measures are a fit for naturalistic and qualitative research carried out by sociologists, psychologists, and other researchers using qualitative techniques. Here it becomes more difficult to meet the expectations of validity as applied to quantitative data. Borsboom, Mellenbergh, and Van Heerden (2004) discuss this and highlight how, in some cases, outright

mistakes were made when applying quantitative validity measures to qualitative research. They also challenge the dependency on correlation as a proof of validity and argue that, simply put, *a test is valid for measuring an attribute if variation in the attribute causes variation in the test score*. As such, they present an argument for causation and not correlation (Borsboom, 2004). Validity thereby expresses the degree to which the system accurately reflects the value or the change in the measure. Views also exist that qualitative and quantitative data do not calibrate exactly but that this does not undermine either tradition. This view underlines the value of combined approaches, using different forms of evidence for *complementary extension* of insight of the social world (Richie, 2003). Nevertheless, it is important that each facet of evidence is as valid and reliable as it can be based on the research design and methods used.

2.0Materials and methods

2.1 Sample

Eight culture evaluations systems already applied within the food industry were included in the comparative analysis. Five systems, referred to in this paper by the name of the authors who first described them: Ball, Denison, De Boeck, Jespersen, and Wright, as well as three systems referred to in this paper by their commercial names, CEB, TSI, and NSF. It is important to note that other commercial evaluation systems are available but these three were the most often applied in the food industry and have been included here for reference. Content from these systems was included in the analysis however, details regarding validation strategies either do not exist or are not available and therefore not included in the comparison of applied validation strategies.

2.2 Method

A six step process was used for the comparative analysis (Figure 1). To obtain data from existing culture evaluation systems specific to the quality, trustworthiness, and content of each system a structured content analysis was completed. Publically available material was gathered e.g., peer reviewed papers,

white papers, and books and imported into NVivo 11 [Computer Software] QSR International, Doncaster, Australia, for deductive, textual data analysis. The NVivo software is designed specifically for qualitative coding of textual and other types of qualitative data. A content analysis framework (Table 1) was developed based on initial reading and descriptions of the eight systems. The framework consists of three levels of textual analysis: basic themes, organizing themes, and global themes (Attride-Stirling, 2001).

Table 1: Content Coding Framework

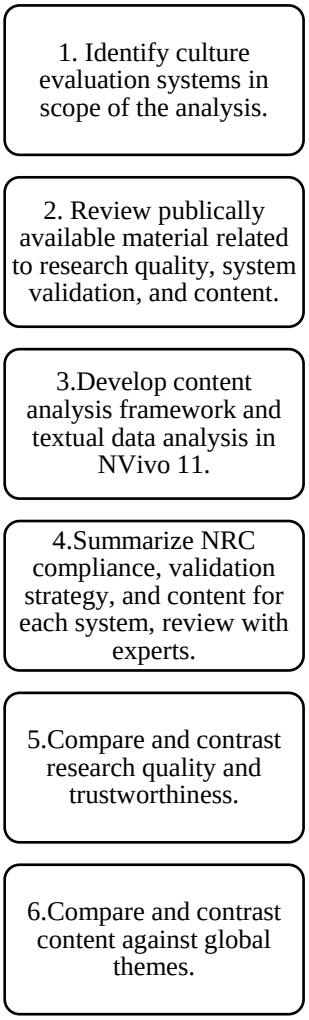
Global	Organizing	Basic
Values and Mission	Statements related to; direction, goals, compliance, improvement, measures, metrics, plan, roadmap, long-term plan, long-term direction. Also, statements as “just the right thing to do…”	• Compliance. • Measures/metrics/KPIs. • Mission, vision, goals. • Ownership/owning, • Plan/roadmap, direction. • Recall/recalls/withdrawals. • Responsibility, accountability, commitment. • Direction
People Systems	Statements related to; role, group, team, accepted behaviour, rejected behaviour, accepted practices, rejected behaviours, training, education, learning, consequences, escalation, celebrations, punishment, communication, group and individual pride.	• Any reference to person’s role/education/job and group or team (e.g., name of team, established teams, established groups). • Behaviour/practice, work routine. • Communications (e.g., written, spoken, and dialog (e.g., interview)) and involvement. • Consequence, escalation. • Pride. • Rewards and celebration. • Training, education, learning, proficiency. • Cross-functional • Unionized. • Rotation and retention.
Consistency	Statements related to; leader communication of system, leader communication of system, due date met, due date missed, tasks, projects, basic tools missing, basic tools available, data collection, data collection tools, data usage,	• Actioned data and performance metrics. Actions, tasks, action due date. • Non-conformance, reoccurring, • Technology • Tools and infrastructure (missing infrastructure, appropriate/right tools, appropriate infrastructure, missing tools). • References to third party standards.

	performance reports, performance, decisions and use of technology.	
Adaptability	Statements related to, improvement, continuous improvement, system improvement, change readiness, change challenges. Also, look for readiness in adoption of new tools, technology, or processes. E.g., will adopt when there is a legal requirement but will not take a broader look to process improvement investments.	• Change readiness, open to change, change ready. • Improvement, must improve, continuous improvement, improvement process, improvement system, continuous improvement, Six Sigma, Lean manufacturing.
Risk awareness	Statements related to; review of risks, hazards identification and assessments, risk assessments, leaders pay attention to control of risks and hazards, operators pay attention and speak up if a risk or hazard gets out of control.	• Leaders risk awareness and perception. • Operator risk awareness and perception. • Risks, hazards.

186
187 The aggregation of basic and organizing themes was used to define the global themes later
188 translated into suggested food safety culture dimensions. The degree of agreement or disagreement in
189 content of each system was evaluated using the coverage of each system for each of the global themes.

Coverage was evaluated by quantifying manifest codes per system per global theme and qualitatively by looking for latent meaning of detailed system content. A summary was written for each system which focused on structure, compliance to the National Research Council (NRC) guidelines, validation strategies, and references to the material included in the comparative analysis. By summarizing the manifest and latent codes per system a comparative figure was developed to visually analyze for agreement or disagreement across the five suggested food safety cultural dimensions.

Figure 1: Six step process applied in the comparative analysis



2.3 Research trustworthiness

A four step validation strategy was applied to assess trustworthiness of the research results;

(1) External population and ecological validity to ensure generalizability. This was achieved through literature review, expert consultation, and coverage of global codes found in each system, analysis and inclusion of language found in content analysis and currently applied in the food safety domain. The experts were chosen for their expertise in the area of culture and independent of the research and/or the models. The experts were trained in the NRC guidelines and reviewed samples of documentation aginst these.

(2) Internal face and construct validity to ensure the content reflects what it is intended to and that the evaluation construct is robust. Peer review was conducted, adding representative quotes linked to the global themes from each evaluation system (Graneheim, 2004). Each system owner was given the opportunity to approve the system summary and assessments.

(3) Internal validity to ensure replicability. This was accomplished through the systemic search and inclusion of most often cited evaluation systems both academic and commercial. Although this, might be influenced by how well a given method is marketed through commercial channels.

(4) Reliability through evaluation of technical accuracy. Group discussions with practitioners and academicians were conducted to ensure that the construct and results are representative of the concept it is intended to measure.

3.0 Results

3.1 Summary of in-scope culture evaluation systems

The following provides a short summary of the eight culture evaluation systems; focus domain, structure, general adherence to the twelve principles in the NRC guidelines, validation strategies, and references.

3.1.1 Ball model

This system was developed by Brita Ball in 2009 as part of her doctorate work at the University of Guelph (Ball, Wilcock, & Aung, 2009; Wilcock, Ball, & Fajumo, 2011). The system is focused on the food safety domain, specifically food safety climate, and was tested with five food manufacturers in Canada. There is not reference to a definition in the papers and this is likely to be due to the early date of the research when no formal definition had been published. Thirteen in-depth interviews were conducted with five small to medium sized processing plants; together with two focus group sessions with interest groups. The model consists of six themes and 20 sub-themes, each measured through a self-assessment survey, in-depth interviews, field observations, and a second self-assessment survey. Analysis of data generated was conducted in NVivo 7 [Computer Software] QSR International, Doncaster, Australia, by applying validated content analysis principles using both deductive and inductive analysis. Multiple methods of data collection allowed researchers to apply some triangulation. A model was developed following Fishbein and Ajzen's reasoned action model (Ajzen, 2011). The model showed strong significance of work unit commitment to food safety as a key driver of the food safety behaviours of food handlers. The system research meets 11 of the 12 NRC guidelines (Figure 2) as it is not clear from publically available material how alternative explanations of the findings were explored.

3.1.2 De Boeck model

This system was developed by Elien De Boeck in 2015 as part of her doctorate work at the University of Gent (De Boeck, Jacxsens, Bollaerts, Uyttendaele, & Vlerick, 2016; De Boeck et al., 2015). The researchers use the terms *culture* and *climate* interchangeably in their research. Food safety culture, they defined as *the interplay of the food safety climate as perceived by the employees and the managers of a company (so called 'human route') and the context in which a company is operating, the current implemented FSMS, consisting out of control and assurance activities (so called 'techno-managerial route') resulting in a certain (microbiological) output* whilst food safety climate was considered *as employees' (shared) perception of leadership, communication, commitment, resources and risk awareness concerning food safety and hygiene within their current work organization* (De Boeck et al., 2015). However, the authors themselves state that the concepts remain vague and with no unanimous definitions (De Boeck et al., 2015); therefore, for the purpose of this analysis no differentiation is made between the terms in evaluating the De Boeck model. The system is focused on the food safety domain and piloted at eight affiliates of a large, centrally coordinated meat distribution company in Belgium. The model consists of five indicators, with 27 sub-indicators, assessed through a self-assessment survey. A detailed study was completed in eight butcheries and butcher shops in Belgium, and though a small sample, some statistical differences were detected in the food safety climate of the participating organizations. The authors define culture and climate and it is not apparent how precisely these definitions are based on existing research (C. J. Griffith et al., 2010a; Frank W. Guldenmund, 2007; Schein, 2004) in the domains of culture and climate. The research meets nine of the 12 NRC guidelines (Figure 2) as it is not clear from publically available material how comprehensive the literature review was that lead to the model, how alternative explanations of the findings are explored, and if there is a potential impact of systematic bias.

3.1.3 Denison model

This system was developed by Dan Denison in 1989 and applied extensively with global organizations since as the Denison Model (D. R. Denison, 1997; D. R. Denison & Mishra, 1995; D. R. H. N. L. Denison, and Colleen Lief, 2012). The system is focused on organizational culture with a branch in the people safety domain and is therefore broader than food safety culture. It consists of four traits assessed through a self-assessment survey. Details of the research have been widely published in books and peer reviewed papers. This method represents the strongest proof of validity based on both quantitative and qualitative research and documented evidence against all 12 NRC guidelines are available in publically available material (Figure 2).

3.1.4 Jespersen model

This system was developed by Lone Jespersen in 2010 as part of her Masters work and subsequently her Doctoral work in 2014 at the University of Guelph (Jespersen et al., 2016; Jespersen & Huffman, 2014). Jespersen et al (2016) state that food safety culture in food manufacturing is rooted in the definition, dimensions, and characteristics of organizational culture, as defined by Schein (2004). The system is focused on the food safety domain and consists of five capability areas. The system was tested with a global food manufacturing company in North America. The evaluation was conducted using triangulation between self-assessment survey; behavioral observations and interviews; and performance assessments and made use of combined deductive and inductive content analysis and quantitative self-assessment data. The research makes use of both quantitative and qualitative research methods and the results were evaluated using a food safety maturity model. The authors openly declare a validation gap since the system was tested in one organization and to demonstrate the validation principle of generalizability the model needs to be tested on other organizations. The system research meets eight of the 12 NRC guidelines (Figure 2) as it is not clear from publically available material how comprehensive

the literature review was that lead to the development of the model, how alternative explanations of the findings are explored, and if there is a potential impact of systematic bias.

3.1.5 TSI model

This system was developed by TSI in 2015 and applied to food service in Dubai and small food manufacturers in the U.K (J. Taylor, Garat, J. P., Simreen, S., & Sarieddine, G., 2015). The authors describe food safety culture as *prevailing attitudes, values and practices related to food safety that are taught, directly and indirectly, to new employees*. The system was built on research in the area of HACCP application conducted in U.K. small and medium size companies and food service restaurants (Gilling, 2001; E. A. Taylor & Taylor, 2004) and the authors also state that their model is based on research from a broad range of academic disciplines and industry sectors (Taylor et al, 2015). The system is an audit tool and focuses on the food safety domain and consists of four categories assessed through a self-assessment survey. Collectively the four categories cover 16 factors and when applied commercially the findings from the self-assessment survey were reported and discussed with clients. It is not clear from the publically available material how the detailed 16 factors were derived and details regarding research methods and validation strategies for this model were not published.

3.1.6 Wright model

This system was developed by Wright, Leach and Palmer on commission for the U.K. Food Standards Agency (FSA) and intended for use by the agency's public health inspectors (Wright & Leach, 2013). The authors use the Griffith et al (2010) definition of food safety culture and the system is focused on the food safety domain and consists of eight elements. The elements are assessed using a self-assessment scale and behavioral observations. The system research meets nine of the 12 NRC guidelines as it is not clear from publically available material how alternate explanations were explored, how potential systemic bias was assessed, and the material was not submitted for peer-review.

3.1.7 CEB model

The system was developed by CEB (CEB, 2016) and makes use of a five level maturity model evaluating quality culture across five categories; organizational scope, employee ownership, peer involvement, message credibility, and leadership emphasis (Srinivasan & Kurey, 2014) The assessment is made thorough employee self-assessment and makes use of a social cognitive model with four characteristics; hear, see, transfer, and feel, in guiding actions based on the assessment. Details regarding validation strategies for this model were not published.

3.1.8 NSF model

The system was developed by NSF (NSF, 2016) in collaboration with Cognisco Ltd. Cranfield, Bedford, U.K.The basis for the system is an NSF assessment of approximately 10,000 food handlers and the theories of social cognitive theory and behavioral science (Fone, 2010) The Culture Maturity system has five phases that go beyond the evaluation of culture into the areas of tactics for changing behaviours and evaluating a company and the efficacy of their food safety and quality management systems. The system evaluates behaviour across six core markers, (1) Regulatory Governance, (2) Management Systems, (3) Policies & Standards, (4) Assessments, (5) Talent Development, and (6) Culture & Behaviours. The evaluation scores are a combination of employee self-assessment and on-site activities and scores are mapped on a scale of four progressive generations, ranging from reactive to core-values. Details regarding validation strategies for this model were not published (D.Fone, Personal communication, November 11, 2016).

3.2 Differences in Validation strategies

Each of the scientifically-based culture evaluation models make use of unique validation strategies and, in exploring the differences, it was found that many models make use of internal face and

331 construct validation but only two show predictive validation (Ball and Denison). Reliability testing is
332 shared in two culture evaluation systems and not clear in the remaining (Ball and Denison) (Table 2).

333 **Table 2: Differences between the validation strategies applied in the eight culture**
334 **evaluation models.**

Culture evaluation model	Validity methodology
Ball	Internal face and construct validation through expert solicitation. External population validation through focus groups and peer review. Respondent validation through responds transcript validation. Internal consistency through Cronbach's Alpha. Predictive validation through behavioural model and triangulation.
De Boeck	External population validation through peer review. Internal, face and construct validation through twenty experts, cross-sector, from Belgium.
Denison	External population and historical validation through analysis of existing performance data. Internal validation through peer review and expert solicitation. Internal construct and predictive validation through correlation analysis using the Survey of Organizations and The Organizational Survey Profile data. Reliability through stability of time.
Jespersen	External population and ecological validation through review of existing food safety performance data and adoption language

Culture evaluation model	Validity methodology
	from existing food safety standards. Internal construct and face validity through an 18 member expert panel cross-sectional, from US, UK, and Canada, consisting of academicians and practitioners. Predictive validation through behavioural model and triangulation.
Wright	External population and ecological validation through focus groups. Internal construct and face validation through expert solicitation and transparency in audit trail through publically available reports.

335

336 **3.3 System Content comparison**

337 The finidings from the content analysis provided data for comparison of the content between the
338 eight culture evaluation systems. It should be noted that not all systems apply the food safety culture
339 definition introduced earlier. The content was grouped to provide a graphical representation of dimension
340 coverage by each culure evaluation system (Figure 3).

Constructs (Ball)	Indicators (De Boeck)	Traits (Denison)	Drivers (CEB)	Capability area (Jespersen)	Markers (NSF)	Categories (TSI)	Elements (Wright)
Management commitment	Leadership	Mission	Leadership emphasis	Perceived Value	Culture and Awareness	Purpose	Perception of safety
Supervisor commitment	Commitment	Involvement	Message credibility	People systems	Management	People	Business priority
Training	Communication	Consistency	Peer involvement	Process thinking	Training	Process	Leadership
Infrastructure support	Resources	Adaptability	Employee ownership	Technology enabler	Regulatory Compliance	Proactivity	Ownership of safety
Worker commitment	Risk awareness			Tools and infrastructure	Policies and standard		Competence
Worker behaviours					Auditing		Employee communication
					Traceability		Employee involvement
					IT Systems		Risk perception

Figure 2: Content comparison of the eight culture evaluation systems. Five affinity groupings emerged from the analysis; red = values and mission, green = people systems, blue = consistency, yellow = adaptability, and purple = risk and hazards

As such, five dimensions of food safety culture were identified, which all culture evaluation systems cover to varying degrees. The dimension are Values and mission, People systems, Consistency, Adaptability, and Risk awareness.

3.3.1 Values and mission

This dimension covers cultural content related to 1) management and employee commitment to food safety (Ball, De Boeck, Taylor), 2) how leadership sets objectives, motivates, and addresses food safety (De Boeck), 3) direction for the organization (Denison), 4) the organization's perceived value and priorities related to food safety (Jespersen, Wright), and 5) food safety ownership (Wright). Wright covers this dimension in four of the eight elements and Ball also covers this in three of six constructs. As such, Ball and Wright have the most detail of any of the cultural evaluation systems in this dimension.

3.3.2 People systems

This dimension covers cultural content related to 1) knowledge, qualifications, and team effectiveness (Ball), 2) training, integration of new employees, and expectations of competency level (Ball, De Boeck, Jespersen, Taylor), 3) leaders and employees' communication of food safety (De Boeck, Wright), 4) actual and expected involvement, autonomy, degree of membership input (Denison, Taylor, Wright), 5) expectations of tasks or behaviours (Jespersen), 6) knowledge of risk (Wright). Wright covers this dimension in three of the eight elements of the model and provides the most detail around this dimension. Wright is also the only one that includes "risk" in the people system dimension. Ball covers this dimension in two of six constructs and is the only one that speaks of "infrastructure" as part of the people system and how this drives food handler food safety behaviours.

3.3.3 Consistency

This dimension covers cultural content related to; 1) degree of following rules (Ball, Taylor), 2) good procedures and instructions are in place (De Boeck), 3) systems are enforced vs. allowance for by-passing (Denison), 4) technology enabled behaviours (Jespersen), 5) access to the right tools and investment in infrastructure (Jespersen). Jespersen covers this in two of five capability areas and Wright does not cover this dimension directly in any elements.

3.3.4 Adaptability

Dimension covers cultural content related to; 1) how the organization embraces or resists change (Denison, Taylor), 2) how problem solving is approached (Jespersen). Three cultural evaluation systems cover this dimension but no direct relation was found in Ball, De Boeck, and Wright.

3.3.5 Risk awareness

Dimension covers cultural content related to; 1) risks are known, under control, and employees are alert to actual and potential food safety risks (De Boeck, Wright). De Boeck and Wright are the only ones that identify this as a separate indicator. Others have risk awareness incorporated in other dimensions but have not assigned as much importance to this dimension as De Boeck and Wright.

3.4 Suggested framework to unify the research field

Based on the analysis of the eight evaluation models and the above discussion a five dimension framework is suggestd to provide some unification of the food safety culture research field. These dimension were found to most extensively cover the content of the existing models and that found in supporting literature (Figure 3).

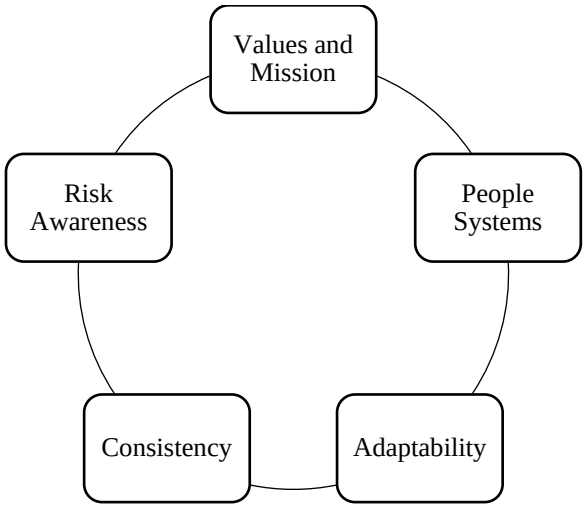


Figure 3: Food Safety Culture - Dimensional Framework

4.0 Discussion and conclusion

Five scientifically-based and three commercial culture evaluation models were within the scope of the comparative analysis; Ball, CEB, Denison, De Boeck, Jespersen, NSF, TSI, and Wright. Analysis included whether the models had been applied in the food industry and therefore directly relevant for the evaluation of food safety culture. Each system was evaluated for compliance to the National Research Council (NRC) guidelines (Richard J. Shavelson and Lisa Towne, 2002), through a comparison of validation strategies, and through results from analysis of available textual data using content analysis. Differences were found in the degree to which the systems were developed according to NRC guidelines, from meeting all to meeting five of the research quality ptinciples (reference). The largest gap was the lack of evidence in the assessment of systemic bias and its documentation. Ball and Denison were found to do this well through transparent assessment of literature and documented path between literature findings and research outcomes. The second area where weaknesses were discovered was related to the assessment of data for alternative interpretation of results. Again, Ball and Denison were found to cover

401 this most comprehensively through discussion of focus group discoveries and missing evidence around
402 impact on culture assessment through organizational levels.

403 Validation and reliability measures are important for the validity of any type of research (Louis
404 Cohen, 2007; Robson & Robson, 2011). Validation and reliability measures were reviewed and sorted
405 according to type of research conducted. Each culture evaluation system was analyzed for the validation
406 strategy applied. Most make use of external validation through population, ecology, and peer review.
407 Only the Ball and Denison models document predictive validity. Ball constructed a predictive model
408 based on the Reason Action model and Denison showed a predictive relation between strength of culture
409 assessment and existing financial and organizational performance data. Few models document reliability
410 measures and this is considered a considerable gap. Again, Ball and Denison do incorporate reliability
411 measures into their models and document the method chosen to do this in detail.

412 Content was compared through content analysis of the textual data. It was found that almost all
413 culture evaluation systems contain some content related to an organization's values and mission. The
414 Wright model dedicates four of five elements to this dimension. Five of the six culture evaluation systems
415 cover content related to people systems. Four of the six cultural evaluation systems cover consistency and
416 Jespersen was found to cover this in three of five dimensions. Adaptability was covered by the models of
417 Ball, Denison, whereas the Taylor and Ball models dedicated two of six constructs to this dimension. Risk
418 awareness was only covered in detail by the De Boeck and Wright models. The content findings suggest
419 that the proposed five dimensions cover all of what each system independently cover and it suggests some
420 areas that could be strengthened in some systems e.g., adaptability, consistency and risk awareness. By
421 reviewing the detailed basic, organizing, and global themes, it is clear that by looking at all content from
422 all systems a very strong picture of not simply organizational culture but food safety culture emerges. It is
423 concluded that the five proposed dimensions could be used to unify research in the food safety culture

domain and also provide each system owner with input into the continuous improvement of each system independently.

The analysis of NRC compliance rate and validation strategy provides information about the quality and trustworthiness of the culture evaluation systems; both of which are critical characteristics of research leading to culture evaluation systems by which food manufacturers make decisions regarding resources for culture transformation. It was surprising to find few of the culture evaluation systems had documented reliability measures and predictable validation strategies. Also, few made use of structured triangulation, a method commonly applied to qualitative and mixed method research (Denzin, 2012) and a method to validate research findings. It was also unexpected that systems named *climate* and those named *culture* had such great similarities in content. If these terms were used consistently and according to historical textual data then it might help guide the specialization of content and truly deliver on both climate and culture evaluation systems. It was an unexpected finding that the global themes and suggested food safety culture dimensions resemble organizational culture dimensions with the important exception of the dimension Risk awareness, which appears to be more specific to food safety. One limitation of the research is the lack of detailed access to the methods behind the Taylor and CEB culture evaluation systems. More documented details for each of these two systems could have contributed further to the comparative analysis.

The definition of food safety culture and that of organizational culture suggest that culture is learned and shared among people (C. J. Griffith et al., 2010a; Schein, 2004). It is based on accepted assumptions, values, and beliefs, is dynamic and impacted by an array of factors and situations. By evaluating culture, food manufacturers can get a snap-shot of strengths and weaknesses and make decisions about actions and resources. Such decisions can make the difference between a group's assumptions and beliefs regarding food safety practices; whether or not to implement them; and

subsequently if consumers are put in harms way or not. Hence the research behind a culture evaluation system must optimize quality, trustworthiness, and cover the broadest possible content to inform the food manufacturer correctly. These results must be given the same importance of quality and trustworthiness as, for example, microbiological testing, sampling for presence of allergens, and detecting metal contamination. The lack of an appropriate food safety culture is an emerging risk (C. J. Griffith, K. M. Livesey, & D. A. Clayton, 2010b) and both academicians and practitioners must hold each other to a high standard to minimize this risk. It is suggested that more research is conducted in the field of unifying food safety dimensions through a common glossary, empirical research and predictive studies and to develop models to assess the maturity of food safety culture within organizations based on these dimensions.

5.0Acknowledgements

The authors would like to acknowledge the owners of the systems in scope of this analysis; Brita Ball, Elien De Boeck, Dan Denison, Dan Fone, Ryan Hadley, Joanne Taylor, and Michael Wright,

6.0 References

Ajzen, I. (2011). The theory of planned behaviour: Reactions and reflections. *Psychology & Health, 26*(9), 1113-1127. doi:10.1080/08870446.2011.613995

Armijo-Olivo, S., Stiles, C. R., Hagen, N. A., Biondo, P. D. and Cummings, G. G. (2012). Assessment of study quality for systematic reviews: a comparison of the Cochrane Collaboration Risk of Bias Tool and the Effective Public Health Practice Project Quality Assessment Tool: methodological research. *Journal of Evaluation in Clinical Practice, 18: 12–18.* .

Attride-Stirling, J. (2001). Thematic networks: An analytic tool for qualitative research. *Qualitative Research, 1(3), 385-405.*

Ball, B., Wilcock, A., & Aung, M. (2009). Factors influencing workers to follow food safety management systems in meat plants in Ontario, Canada. *International journal of environmental health research, 19*(3), 201-218. doi:10.1080/09603120802527646

Bernd Appel, G.-F. B. l., Matthias Greiner, Monika Lahrssen- Wiederholt und Andreas Hensel. (2011). *EHEC-Ausbruch 2011*. Retrieved from

Borsboom, D. M., Gideon J.; van Heerden, Jaap. (2004). The Concept of Validity. *Psychological Review, Vol 111(4), Oct 2004, 1061-1071.*

Cameron, K., & Quinn, R. (2006). *Diagnosing and changing organizational culture : Based on the competing values framework* . : San Francisco: Jossey-Bass.

Canadian Food Inspection, A. (2013). [Action on Weatherill Report Recommendations to Strengthen the Food Safety System: Final Report to Canadians]. Web Page.

CEB. (2016). www.cebglobal.com/innovation-strategy/quality/culture-of-quality.html.
De Boeck, E., Jacxsens, L., Bollaerts, M., Uyttendaele, M., & Vlerick, P. (2016). Interplay between food safety climate, food safety management system and microbiological hygiene in farm butcheries and affiliated butcher shops. *Food Control, 65*, 78-91. doi:http://dx.doi.org/10.1016/j.foodcont.2016.01.014
De Boeck, E., Jacxsens, L., Bollaerts, M., & Vlerick, P. (2015). Food safety climate in food processing organizations: Development and validation of a self-assessment tool. *Trends in Food Science & Technology, 46*(2, Part A), 242-251. doi:http://dx.doi.org/10.1016/j.tifs.2015.09.006
Denison, D. R. (1997). *Corporate Culture and Organizational Effectiveness*: Denison Consulting; 2nd edition (May 1997).
Denison, D. R., & Mishra, A. K. (1995). Toward a Theory of Organizational Culture and Effectiveness. *Organization Science, 6*(2), 204-223.
Denison, D. R. H. N. L., and Colleen Lief. (2012). *Leading Culture Change in Global Organizations: Aligning Culture and Strategy*. John Wiley & Sons.
Denzin, N. K. (2012). Triangulation 2.0. *Journal of Mixed Methods Research, 6*(2), 80-88. doi:10.1177/1558689812437186
Feuer, M. J., Towne, L., & Shavelson, R. J. . (2002). Scientific culture and educational research. . *Educational Researcher, 31(8), 4-14.*
Gersten, R., Baker, S., & Lloyd, J. (2000). Designing high-quality research in special education. *The Journal of Special Education, 34(1), 2-18.*
Gilling, S. J. (2001). Successful Hazard Analysis Critical Control Point Implementation in the United Kingdom: Understanding the Barriers through the Use of a Behavioral Adherence Model. *Journal of Food Protection, 64*, 710-715.
Gossner, C. M. E., Schlundt, J., Ben Embarek, P., Hird, S., Lo-Fo-Wong, D., Beltran, J. J. O., . . . Tritscher, A. (2009). The Melamine Incident: Implications for International Food and Feed Safety. *Environ Health Perspect, 117*(12), 1803-1808. doi:10.1289/ehp.0900949
Graneheim, U. L., B. (2004). Qualitative content analysis in nursing research: Concepts, procedures and measures to achieve trustworthiness. *Nurse Education Today, 24*, 105-112.
Greenhalgh, T. (1997). Assessing the methodological quality of published papers. . *BMJ : British Medical Journal, 315(7103), 305–308.*
Griffith, C. J. (2010). Do businesses get the food poisoning they deserve? *British Food Journal, 112*(4), 416-425. doi:10.1108/00070701011034420
Griffith, C. J., Livesey, K. M., & Clayton, D. (2010a). The assessment of food safety culture. *British Food Journal, 112*(4), 439-456. doi:10.1108/00070701011034448
Griffith, C. J., Livesey, K. M., & Clayton, D. A. (2010b). Food safety culture: the evolution of an emerging risk factor? *British Food Journal, 112*(4), 426-438. doi:10.1108/00070701011034439
Gronlund, N. L., Robert L. (1990). *Measurement and Evaluation in Teaching*.
Guldenmund, F. W. (2000). The nature of safety culture: a review of theory and research. *Safety Science, 34*(1–3), 215-257. doi:http://dx.doi.org/10.1016/S0925-7535(00)00014-X
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
Howe, K., & Eisenhart, M. (1990). Standards for Qualitative (and Quantitative) Research: A Prolegomenon. *Educational Researcher, 19*(4), 2-9. doi:10.3102/0013189x019004002
Incident, G. I. i. t. W. P. C. C. (2014). The WPC80 incident: Causes and responses government InquIry Into the Whey ProteIn Concentrate ContamInatIon IncIdent.
Ingelfinger , J. R. (2008). Melamine and the Global Implications of Food Contamination. *New England Journal of Medicine, 359*(26), 2745-2748. doi:doi:10.1056/NEJMp0808410

Jespersen, L., Griffiths, M., Maclaurin, T., Chapman, B., & Wallace, C. A. (2016). Measurement of food safety culture using survey and maturity profiling tools. *Food Control, 66*, 174-182. doi:http://dx.doi.org/10.1016/j.foodcont.2016.01.030

Jespersen, L., & Huffman, R. (2014). Building food safety into the company culture: a look at Maple Leaf Foods. *Perspectives in Public Health*(May 8, 2014). doi:10.1177/1757913914532620

Louis Cohen, L. M. a. K. M. (2007). *Research Methods in Education* (Vol. Sixth edition).

Meyrick, J. (2006). What is good qualitative research? . *Journal of Health Psychology, 11(5), 799-808.*

NSF. (2016). www.nsf-foodeurope.com/service.asp?service_id=consulting&subservice_id=9&servicepage_id=127.

Nyarugwe, S. P., Linnemann, A., Hofstede, G. J., Fogliano, V., & Luning, P. A. (2016). Determinants for conducting food safety culture research. *Trends in Food Science & Technology, 56*, 77-87. doi:http://dx.doi.org/10.1016/j.tifs.2016.07.015

Pennington, H. (2009). *The Public Inquiry into the September 2005 Outbreak of E.coli O157 in South Wales*. Retrieved from

Pennington, T. H. (2014). E. coli O157 outbreaks in the United Kingdom: past, present, and future. *Infection and Drug Resistance, 7, 211+*.

Powell, D. A., Jacob, C. J., & Chapman, B. J. (2011). Enhancing food safety culture to reduce rates of foodborne illness. *Food Control, 22*(6), 817-822. doi:10.1016/j.foodcont.2010.12.009

Richard J. Shavelson and Lisa Towne. (2002). *Scientific Research in Education Research*: National Research Council and The National Academies Press.

Richie, J. L., J. (2003). *Qualitative Research Practice - A Guide for Social Science Students and Researchers*: SAGE Publications.

Roberts, P., & Priest, H. . (2006). Reliability and validity in research. . *Nursing Standard, 20(44), 41+. Retrieved from* http://go.galegroup.com.subzero.lib.uoguelph.ca/ps/i.do?p=AONE&sw=w&u=guel77241&v=2.1&it=r&id=GALE%7CA149022548&asid=467ae8ef2228c61c7546df411b0db083.

Robson, C., & Robson, C. (2011). *Real world research : a resource for users of social research methods in applied settings*. Chichester, West Sussex: Chichester, West Sussex : Wiley.

Schein, E. H. (2004). *Organizational culture and leadership*. San Francisco: San Francisco : Jossey-Bass.

Seward, S. (2012). Assessing the food safety culture of a manufacturing facility. *Food technology, 66*(1), 44.

Shepard, L. A. (1993). Evaluating Test Validity. *Review of Research in Education, 19*(Jauary 1993), 405-450.

Shepard, L. A. (2016). Evaluating test validity: reprise and progress. *Assessment in Education: Principles, Policy & Practice, 23*(2), 268-280. doi:10.1080/0969594X.2016.1141168

Srinivasan, A., & Kurey, B. (2014). Creating a Culture of Quality. *Harvard Business Review, 92*(4), 23-25.

T. Long, M. J. (2000). Rigour, reliability and validity in qualitative research. *Clinical Effectiveness in Nursing, 4*(1), 30-37.

Taylor, E. A., & Taylor, J. Z. (2004). Using qualitative psychology to investigate HACCP implementation barriers. *International journal of environmental health research, 14*(1), 53-63. doi:10.1080/09603120310001633877

Taylor, J., Garat, J. P., Simreen, S., & Sarieddine, G. (2015). An industry perspective: A new model of Food Safety Culture Excellence and the impact of audit on food safety standards. *Worldwide Hospitality and Tourism Themes, 7(1), 78-89. Retrieved from* http://sfx.scholarsportal.info/guelph/docview/1648846697?accountid=11233.

Trochim, W. M. K. (2006). Research Methods Knowledge Base. *Reliability and Validity.* Retrieved from http://www.socialresearchmethods.net/kb/relandval.php

Weiser, A. A., Thöns, C., Filter, M., Falenski, A., Appel, B., & Käsbohrer, A. (2016). FoodChain-Lab: A Trace-Back and Trace-Forward Tool Developed and Applied during Food-Borne Disease Outbreak Investigations in Germany and Europe. *PLoS ONE, 11*(3), 1-11. doi:10.1371/journal.pone.0151977

Wilcock, A., Ball, B., & Fajumo, A. (2011). Effective implementation of food safety initiatives: Managers', food safety coordinators' and production workers' perspectives. *Food Control, 22*(1), 27-33. doi:10.1016/j.foodcont.2010.06.005

Wright, M., & Leach, P. (2013). Achieving an effective food hygiene culture: the next step in assuring excellence. *International Food Hygiene, 24,* 21-23.

Yiannas, F. (2009). *Food safety culture creating a behavior-based food safety management system*: New York : Springer, c2009.