Triangulation and the importance of establishing valid methods for food safety culture evaluation

Author names and affiliations

Lone Jespersen, [University of Guelph (1] and Prof. Carol A Wallace, [University of Central Lancashire (2)].

(1) University of Guelph, Department of Food Science, 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, Department of Food Science, 50 Stone Road East, Guelph, ON, N1G 2M7, Canada. +41792460807. lone@uoguelph.ca

Abstract

The research evaluates maturity of food safety culture in five multi-national food companies using method triangulation, specifically self-assessment scale, performance documents, and semi-structured interviews. Weaknesses associated with each individual method are known but there are few studies in food safety where a method triangulation approach is used for both data collection and data analysis. Significantly, this research shows that individual results taken in isolation can lead to wrong conclusions, resulting in potentially failing tactics and wasted investments. However, by applying method triangulation and reviewing results from

a range of culture measurement tools it is possible to better direct investments and interventions. The findings add to the food safety culture paradigm beyond a single evaluation of food safety culture using generic culture surveys.

Keywords

Method triangulation, food safety culture evaluation, maturity profiling culture scale, content analysis, semi-structured interview.

Highlights

- Establishes importance of triangulation for valid food safety culture evaluation
- Compares data from scale, performance documents, and semi-structured interviews
- Confirms need for multiple methods for trustworthy evaluation of food safety culture
- Applies culture coding framework to interview transcripts and performance documents
- Inter-coder and construct validity, and discrimination in food safety culture profiles

1.0 Introduction

The understanding of culture to enable organizational effectiveness has been studied at length since 1970 and before. (Hofstede, 1980, 2001, 2013) studied national culture through his cross-cultural organizational studies research, starting with the international (IBM) survey in 1966, and showed predictive validity of his ‘Values Survey Module’ instrument to dimensions of national culture. D. R. Denison (1997) developed a model for corporate culture and organizational effectiveness through his research on organizational culture evaluation methods with predictive validity of two measures of organizational effectiveness: behavioral data and financial data (D. Denison, Hooijberg, Lane, & Lief, 2012; D. R. Denison, 1997; D. R. Denison & Mishra, 1995). These types of evaluations appeal to leaders in organizations as they quantify areas of strength and weakness in an accessible and validated form. Culture researchers, in all domains, must take seriously these lessons from early front-runners, like Hofstede and Denison, to understand the dichotomy of fulfilling leaders needs for aggregated, leading indicators of culture change progress and developing meaningful and trustworthy measurement tools. (Guldenmund, 2000) discusses this dichotomy specific to the people safety culture domain. He postulates that assumptions are often made that organizations are homogeneous and can be evaluated using an organization-wide, generic questionnaire survey but that this approach can be risky and virtually meaningless as organizations are highly heterogeneous and made up of formal and informal working groups (Guldenmund, 2000). This suggests that other approaches are needed to understand the heterogeneity of organizations which are typically made up of sub-groups and macro-cultures (Schein & Schein, 2017).

1.1 Theoretical framework

To link the food safety domain with existing models for organizational culture, safety climate/culture, and food safety climate/culture, Jespersen et al (2017) developed a theoretical framework based on eight existing cultural evaluation models (Ball, Wilcock, & Aung, 2009; De Boeck, Jacxsens, Bollaerts, Uyttendaele, & Vlerick, 2016; De Boeck, Mortier, Jacxsens, Dequidt, & Vlerick, 2017; Denison et al., 2012; Denison, 1997; Denison & Mishra, 1995; Jespersen, Griffiths, Maclaurin, Chapman, & Wallace, 2016; Srinivasan & Kurey, 2014; Taylor, 2015; Wilcock, Ball, & Fajumo, 2011; Wright, 2013). The framework was developed through content analysis of eight culture or food safety culture evaluation systems. Each of the systems had been applied to evaluate culture in food companies by applying mostly self-assessment surveys. Content analysis was completed in NVivo 11 [Computer Software] QSR International, Doncaster, Australia] by importing textual material into NVivo and coding content to nodes deduced from literature review. The researchers deduced the dimensions from the coded material by comparing the details of the specific dimensions from each system. Although these had been named differently by each author, i.e., dimensions, traits, capability areas, categories, elements, Jespersen et al (2017) aligned the descriptors in this framework under the title "dimensions." Together the five dimensions (Figure 1) encompass all the individual dimensions in the eight culture evaluation systems, although none of the eight systems covers all five dimensions. The framework (Jespersen et al, 2017) was the first work to compare and contrast culture evaluation systems with the goal of developing one theoretical framework. Its development is an attempt to bring consensus to the theory of food safety culture and the framework has been applied by the Global Food Safety Initiative (GFSI) in its work to provide guidance to its stakeholders on food safety culture (pers. comm. Robach[1], 2016).

[1] Mike Robach, Chair of Global Food Safety Initiative Board.

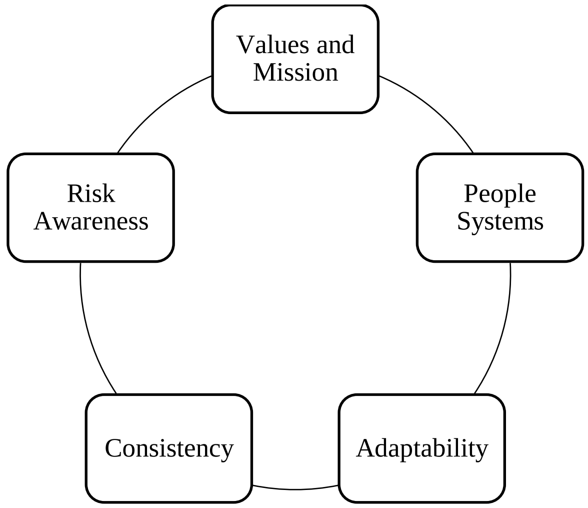


Figure 1: Food safety culture – dimensional framework (Jespersen, Griffiths, and Wallace, 2017)

1.2 Food safety culture evaluation systems

Jespersen et al (2017) report that it is necessary to determine the trustworthiness of culture evaluation system results to assess their validity and reliabililty and this is particularly important where cultural evaluation is being used as part of consumer protection measures in the food safety domain. However, current systems for evaluating culture are fragmented and built on disparate scientific theories (De Boeck, Jacxsens, Bollaerts, & Vlerick, 2015; Guldenmund, 2000), and many make use of single evaluation methods, e.g. a self-assessment scale or audit (Jespersen et al, 2017), an approach not without its limitations (Guldenmund, 2000). Thus it is important to consider whether food safety culture evaluation systems could be strengthened by extension with additional evaluation methods and whether this can give richer information about the heterogeneious organisations in the global food supply chain.

1.3 Method Triangulation

Triangulation has for more than 75 years been an accepted method to confirm that the variance of a phenomenon is tested and not the variance of the method(s) used (Campbell, 1959; Denzin, 1970; Denzin, 2012; Miles, 1994). These and other authors have defined six types of triangulation including the one applied in this research – method triangulation. Method triangulation means to *gather information pertaining to the same phenomenon through more than one method, primarily to determine if there is a convergence and hence, increased validity in the findings* (Carugi, 2016; Kopinak, 1999). Triangulation enables examination of similarities and discrepancies in a research topic, and the assessment of socially desirable responding in sensitive and complex topics (Bauwens, 2010). In addition, it allows researchers to strive for completeness and confirmation of research findings (Yeasmin & Rahman, 2012) as weaknesses in one method can be counterbalanced by the strength in others (Carugi, 2016; Kopinak, 1999). Given both the inner and outer influences that can significantly influence the strength of organizational and -food safety culture, as in other social science domains e.g., health (Carugi, 2016; Kopinak, 1999), it is reasonable to assume that combining or triangulating methods in the investigation process can provide a more comprehensive evaluation of cultural strength. Social realities, such as those existing in organizational and food safety cultures, are inherently complex and therefor difficult to evaluate with one method (Yeasmin & Rahman, 2012). Triangulation can lead to an elaboration and enrichment of findings e.g., by providing more detail, multilayered and multi-dimensional perspectives of the phenomenon being studied (Carugi, 2016; Kopinak, 1999) and increase credibility of scientific knowledge by improving both internal consistency and generalizability (Yeasmin & Rahman, 2012). Quoting McKinlay (1992), “rigid adherence to one approach at the expense or to the exclusion of the other, is destructively parochial and results in often incomplete or even inaccurate explanations and by extension, wrongly focused research. In the data analysis phase triangulation offers several benefits: verification of overlapping results, validation of quantitatively generated constructs through comparison, opportunity to probe and investigate potential causes of discrepancies due to instruments or misrepresentation of data, and clarity of ambiguous and provocative replies or questions (Floyd, 1993). There are difficulties related to the application of method triangulation. There must be consistent and clear foci between the different methods and, in advance of the research, the researcher must have clear prior understanding of the main ontological and epistemological position of the phenomenon under investigation without which the findings and conclusions might be meaningless (Norman K Denzin & Lincoln, 2011). Also, triangulation is time consuming and will increase the time needed to complete a study; however, the authors would argue that this approach is essential in establishment of new evaluation methods. Lastly triangulation is carried out with complex research designs and there are limited guidelines available to researchers as for how to meaningfully combine different data types, interpret divergent results, decide what to do with overlapping concepts, and how to weigh different sources of information (Carugi, 2016; Kopinak, 1999). Further literature discussion would be beneficial to overcome gaps in guidance; however, discussion of potential approaches with other researchers to reach consensus in triangulation plans would seem to be a good way forward and was applied in this research. The objective of this research was to develop and apply method triangulation to increase validity of food safety culture evaluation results.

2.0 Materials and methods

This research was part of a large study of food safety culture performance conducted in collaboration with five multi-national North American-based food manufacturing companies from October 2015 to March 2016. The five companies volunteered to participate in the research and provided the researcher access to total 21 plants. The companies varied in sizes from total three manufacturing sites to over 100 per company. Products manufactured by the companies varied as well from prepared meats, canned vegetables, milk power, and cheese. To reach saturation in qualitative research there are various guidelines regarding sample sizes (Creswell, 1998; Denzin & Lincoln, 2011). For this triangulation study, one plant from each company was sampled and three data sets were collected from each plant (Table 1).

Table 1: Sources by plant and data type

Plant ID	1	2	3	4	5
Self-assessment responses	63	14	10	15	71
Performance documents	5	1	6	5	3
Semi-structured interviews	2	2	2	2	2

The authors believe this sample size to be large enough to obtain a result that could help test the hypothesis that triangulation provides a more comprehensive evaluation of culture than relying on a single method. Three data sets were; food safety culture maturity self-assessment responses, food safety documents, and semi-structured interviews with plant leaders (Figure 2).

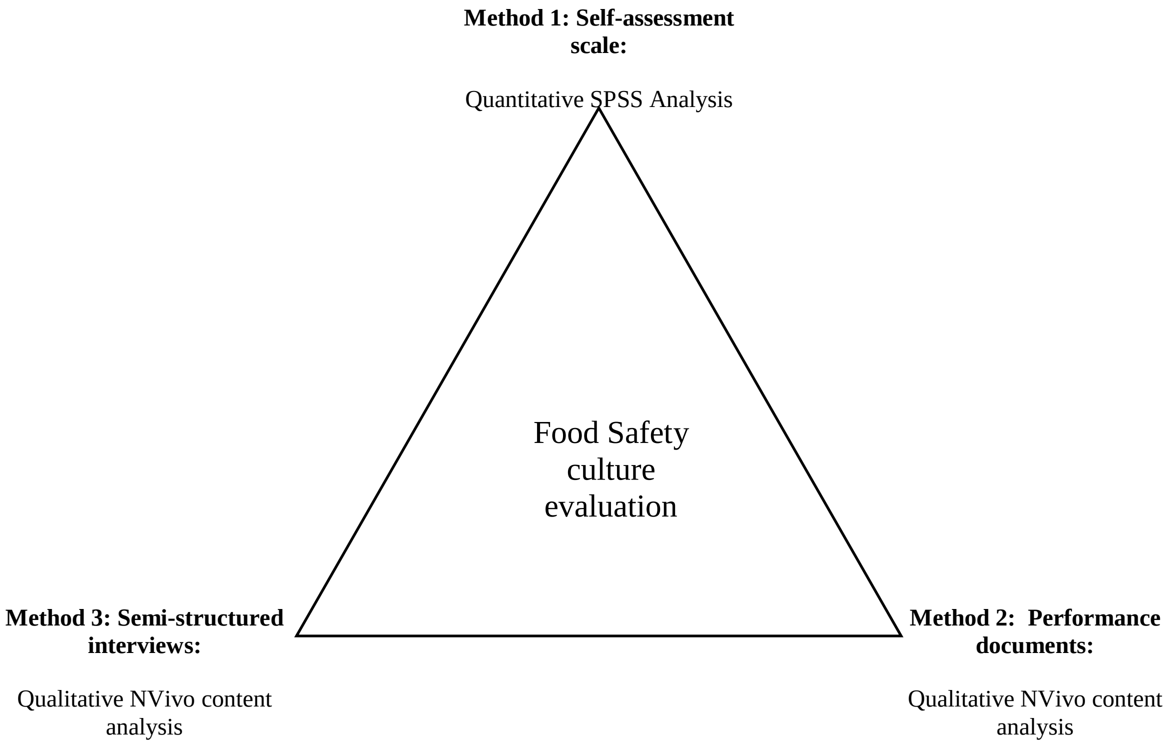


Figure 2: Methods and data triangulation applied to evaluate of food safety culture.

Each method was selected to provide as much data possible on the same phenomenon – food safety culture – to counter weaknesses in each other method, to gain depth of understanding and to make use of already existing data e.g., food safety documents.

2.1 Methods strengths and weaknesses

Three methods were selected for the study of triangulation (Figure 2). These three were selected as they were believed to collectively minimize the method weaknesses of the individual methods and provide complementary data from the plants under investigation based on the strengths and practicalities of each. Strength and weaknesses of each of the three methods are discussed to illustrate how each method can mitigate weaknesses in others through method triangulation. Method 1- Scale: The strengths of scales or survey are that they are simple and straightforward methods for respondents to share knowledge, they provide generalizable information, and maintain respondent anonymity. The weaknesses are that data are affected by the characteristics of the respondents, there can be a gap between respondents' actual beliefs and attitudes to the responses, low response rates that can make it difficult to know if the results are representatives of all groups, and insincere responses can be hard to detect (Denzin, 1970; Robson, 2011). Method 2 – Performance document content analysis: Strengths of content analysis are data gathering is virtually unobtrusive, low cost, can be used non-reactively, and data can relatively easy be generated for longitudinal analysis. The weaknesses of this method are potential difficulty in locating content relevant to the research questions, that it is limited to analyzing records and information that others have decided were worth preserving, and it is ineffective for testing causality as such content analysis can be used to say what is present but not why (Berg, 2012; Robson, 2011). Method 3 – Semi-structured interviews: Strengths of semi-structured interviews are the ability to follow up on leads, providing a moving trail of investigation based on the respondents answer. They are especially suitable for collecting data of sensitive topics because of interviewers ability to investigate underlying motivations, and capture non-verbal clues that can help better understand the verbal responses. The weaknesses are quality of data is highly dependent on the skills and experience of the interviewer, internal consistency can be difficult to demonstrate due to lack of standardization, interviews are time consuming, it can be difficult to penetrate a groups language and mechanisms of symbolisms, and there can be a resistance for the interviewee to "tell it all" (Berg, 2012; Brinkmann, 2015;

Holstein, 1995; Robson, 2011). As such, the weaknesses of each method are countered by either one or both the other methods. For example, survey and interviews can help assign causation, survey can help mitigate impact of interviewer skill and experience, content can help penetrate the group language and symbol mechanisms, content and survey can get data to close the attitude to behaviour gap, survey social desirability and interviews can help identify insincere respondents.

2.2 Response analysis of self-assessment scale.

All salaried staff in each manufacturing plant were invited to participate in an online survey between November 2015 and March 2016. The survey invitation was sent via email with a letter of invitation and purpose of the study for which the data were to be used. The participants were also informed of the confidential nature of their individual responses and encouraged through total three contact points (i.e., invitation, reminder, final reminder) to participate in the study. The scale was developed by (Jespersen et al., 2016) and included questions pertaining to four areas to measure food safety culture maturity; social norms, behavioral intent, motivation, and social desirability. Response data were imported into SPSS [Computer Software] IBM Corporation, New York, U.S.A. from Qualtrics [Computer Software] Qualtrics, Provo, Utah, USA and readied (e.g., removal of incomplete data sets, reversal of negative scales) for analysis. An aggregated maturity score (mean and standard deviation) as well as maturity level by dimension (mean and standard deviation) were calculated for each plant with control for social desirability score (Jespersen, Maclaurin & Vlerick, 2017) amended with the findings from (Jespersen & Edwards, Under review)

2.3 Content analysis of performance documents.

The content analysis of food safety performance documents provides an insight into the documented food safety culture e.g., level of consistency, adaptability, and perceived value of food safety. Each of the manufacturing plants were asked to share food safety documents dating back 12-months from November 2015. Food safety documents such as food safety audit reports, food safety meeting minutes, inspection reports, and Good Manufacturing Practice (GMP) records were obtained from each plant. Content analysis was applied to generate textual data from these documents using a predefined coding framework deduced from literature review and analysis of food safety culture and organizational culture evaluation tools. The coding framework (Table 2) was defined using the theoretical framework (Figure 1) of food safety culture and translated into nodes in NVivo [Computer Software] QSR International, Doncaster, Australia. Sub-nodes were deduced through literature review and induced throughout the coding process. Each document was imported into NVivo and all documents were coded by two researchers.

2.4 Content analysis of semi-structured interviews.

Semi-structured interviews with senior plant leader and senior food safety leader were arranged through the participating company sponsors. Invitation to the interview was sent via email from the lead researcher and logistical detail arranged directly with the plant leader. Interview questions were shared in advance with the interviewees and informed consent obtained for each interview. All interviews were recorded and each audio file transcribed and codified to ensure anonymity of the interview and uploaded to NVivo for content analysis. The same coding framework was used for the interview files as the food safety documents (Table 2)

Table 2: Coding framework used in the content and textual analysis'. Adapted from Jespersen, Griffith, and Wallace (2017).

Node	Sub-Nodes
Values and Mission	Compliance.
	Measures/metrics/KPIs.
	Mission, vision, goals.
	Ownership/owning.
	Plan/roadmap, direction.
	Recall/recalls/withdrawals.
	Responsibility, accountability, commitment.
	Direction, setting expectations, corporate direction.
	Financials, budgets, and prioritizing.
People Systems	Any reference to persons' role/education/job and group or team and references to individuals.
	Behaviour/practice, work routine.
	Communication and dialog.
	Involvement.
	Consequence, escalation.
	Pride.
	Rewards and celebration.
	Training, education, learning, proficiency.
	Cross-functional.
	Unionized.
	Rotation and retention.
	"Making choices…"
Consistency	Actions, tasks, action due date.
	Non-conformance, reoccurring.
	Technology.
	Tools, infrastructure, and policies/procedures.
	References to third party standards.
	Problems, breakdowns, and issues.
Adaptability	Change readiness, open to change, change ready.
	Improvement, must improve, continuous improvement, improvement process, improvement system, continuous improvement, Six Sigma, Lean manufacturing.
Risks and Hazards	Leaders risk awareness and perception.
	Operator risk awareness and perception.
	Risks, hazards.

2.5 Content coding.

The content was coded using practices already applied in the food safety domain (Wallace, 2009). The process for coding content (Figure 3) was used by two independent coders to ensure validity of data. The process consists of two checks for consistency evaluated through calculation of percentage pairwise agreement. (Neuendorf, 2002) argues that the goal for pairwise agreement in social sciences often are .8 but that .9 levels are most appropriate. This higher threshold level has also been suggested to account for some weaknesses in this method (Lombard, Snyder-Duch, & Bracken, 2002). Based on these references the standard for this research for pairwise agreement level was set to .9 (90% agreement). Detailed research questions were defined (step 1) and a coding framework was deduced (step 2) and translated into NVivo nodes and sub-nodes (step 3). The framework was an important component as it connects the coded data to the theoretical framework and the research domain. Following this, coders were trained (step 4) and two documents coded by same coders (step 5). The results were analyzed by detailed review of verbatim data to look for similarities and differences between coders. A decision was made to go back to the coding framework and update with addition of sub-nodes and to go back to the test documents for recoding (step 6). Following this loop, the decision was made to carry on with the full document coding as coders were considered “consistent” based on another detailed verbatim review (step 7). Midway discussions between coders allowed comparison of experience, and discussion of coding difficulties and issues. These results led to another rework of the two selected documents and finalization of the 30 documents (step 8). Finally, the data was analyzed to derive information to answer the RQs (step 9).

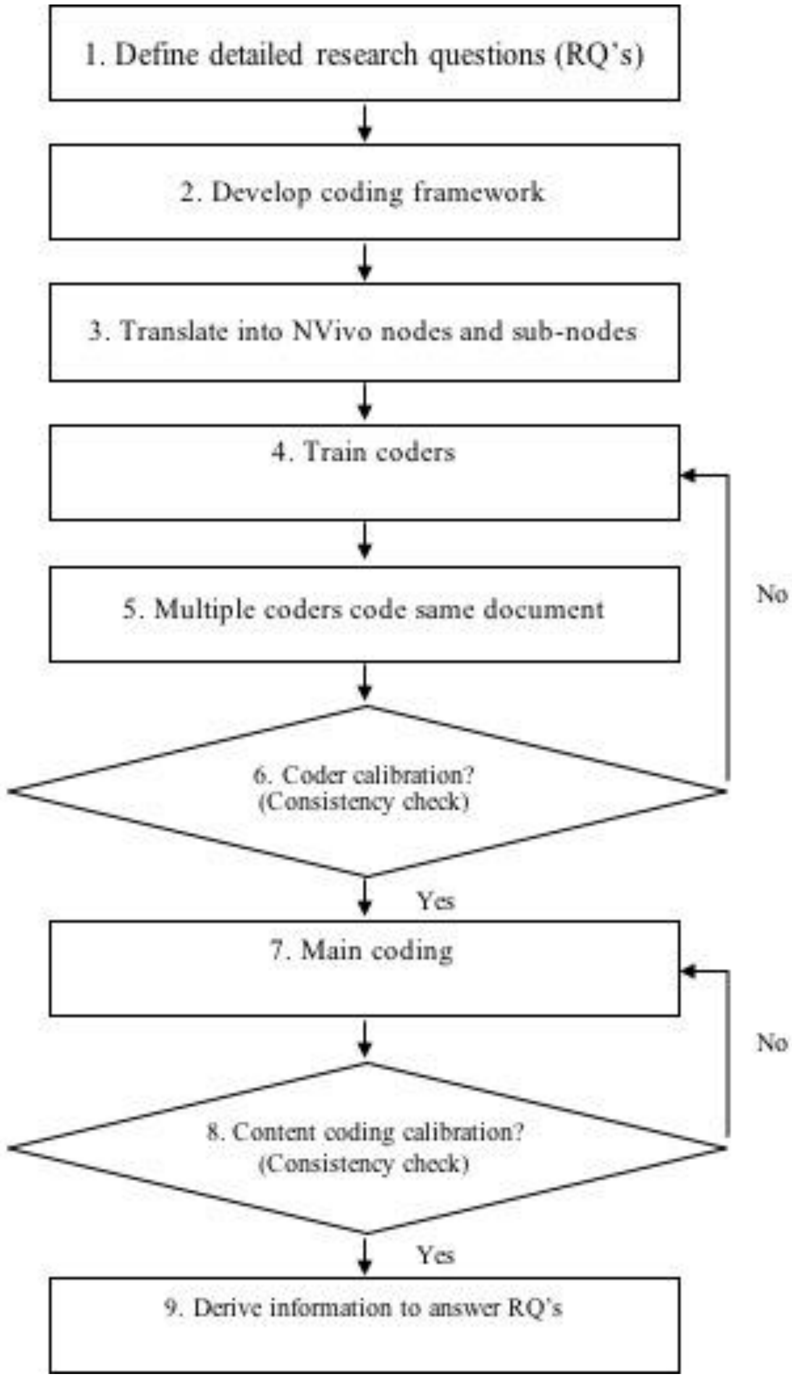


Figure 3: Coding process applied to deriving data through content analysis

2.6 Data triangulation.

An updated version of the food safety maturity model (Jespersen et al., 2016) was used to plot maturity by plant by cultural dimension based on the theoretical framework and scale analysis (Jespersen and Edwards, 2017, under review). Three data points were plotted for each plant, (1) quantitative results from the self-assessment scale were plotted directly on the model's scale from stage one to stage five, (2) qualitative data based on the results from the file analysis was grouped by plant by dimension and each cluster was plotted on the stage of maturity with best fit to maturity model descriptors and behaviours, and (3) qualitative data based on the results from the semi-structured interview analysis was grouped by plant by dimension and each group was plotted on the stage of maturity with best fit to maturity model descriptors and behaviours. By reviewing coded material for both (2) and (3) and comparing verbatim samples to the definition of each maturity stage an individual score for (2) and (3) was assigned. For example, "…yes, so we have some proactive and mainly reactive plethora of data, all manual…everything is manual, right" this verbatim sample would be tagged as a stage 3 statement "knowing." Taking another example, "…this company has never had a recall. I can't be the one that lets that happen…" this verbatim sample would be tagged as a stage 2 "reactive" statement. In this way, all codes were reviewed and placed in stage of maturity with best fit and an aggregated mean score calculated from proportions of coded results in each stage. The triangulation allowed for interpretation of findings for similarities, differences, identifying relationships, extracting themes, and creating generalizations and to ensure that strengths and weaknesses of each method were offset.

3.0 Results

3.1 Self-assessment results.

Differences in overall, aggregated maturity ratings through the self-assessment scale for the five plants in the sub-set are not statistically significant for the overall maturity F (4,182) = .273, p =*.895* (Table 3).

Table 3: Sample size and mean maturity score from self-assessment scale. Total and by individual dimension by plant. Lowest maturity score = 1; highest maturity score = 5.

Maturity	Plant 1	2	3	4	5
N (Response rate)	63 (82%)	14 (78%)	10 (43%)	15 (58%)	71 (41%)
Overall, aggregated score	3.14	3.18	3.17	3.06	3.15
Values and Mission	3.10	3.39	2.82	2.79	3.29
People	3.41	3.41	3.46	3.44	3.29
Consistency	2.93	2.76	3.22	2.97	2.87

The dimensions of Risk Awareness and Adaptability emerged from the food safety culture dimensional framework developed by assessing 8 culture evaluation systems (Jespersen et al, 2017); however, these dimensions did not form part of the earlier Jespersen *et al* (2016) tool and the subsequent evaluation scale which was tested through this research. As such, these two dimensions could not be part of the method triangulation validation of the self-assessment scale.

3.2 Coding comparisons.

A comparison of Coders by dimension is shown in Figure 4. Total 4,522 references were coded in 10 interview transcripts and 20 performance documents. Coders are considered similar if within the set standard of 90% agreement. Agreement between coders was calculated for each dimension and lowest level of pairwise agreement was calculated to 90.4%. This result was obtained after coding and recoding as per Figure 3. As such, content from two dimensions needed to be recoded; Values and Mission and Risk Perception. The bar chart (Figure 4) shows that coders are within 90% agreement on scoring except for Values and Mission (69% agreement) and Risk Awareness (79% agreement).

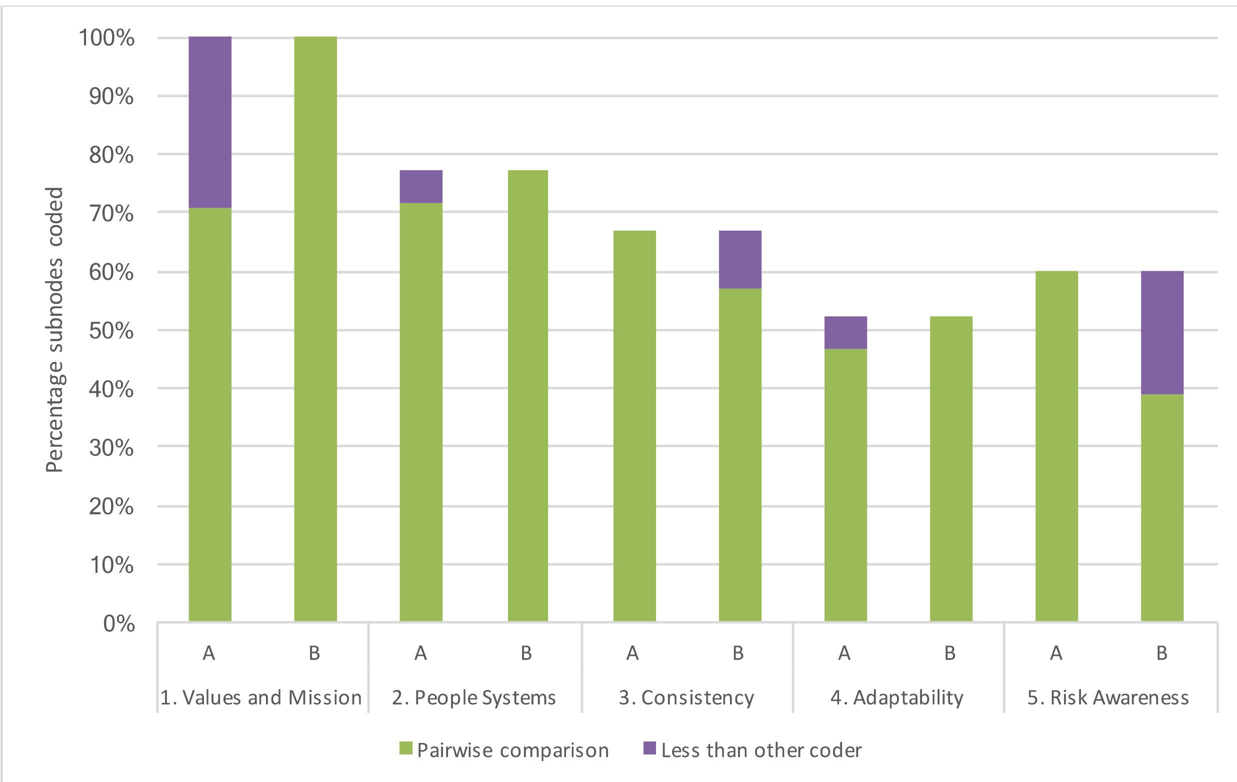


Figure 4: Codes by dimension with pairwise comparison and difference by coder (A and B = two different coders).

In looking at the sub-nodes for Values and Mission (Figure 5) most of this difference comes from differences in scoring of sub-nodes “Measures, metrics, and KPIs” and “Mission, Vision, and Goals”. Coder B coded 52.1% more in the “Measures” than coder A and Coder A coded 40.3% more in “Mission” than Coder B. In addition, in “Recall, recalls, withdrawals” Coder B coded 32.5% more than Coder A, the sub-node “Measures”, where verbatim data show that Coder B coded any “metric” e.g., LM Product 0%, whereas Coder A was looking for measures taken to improve. Sub-node “Mission” verbatim shows that Coder A coded any paragraph or statement leading to direction or priority of the organization. Coder A also included any reference to “policy” which Coder B did not. Sub-node “Recall” verbatim show that Coder A coded any paragraph with the word “recall” whereas Coder B coded paragraphs that indicate recall as a potential outcome of a situation or environment. The differences between coders were reviewed by both coders, discussed, and where needed, amendments were made to increase clarity of application of the coding framework.

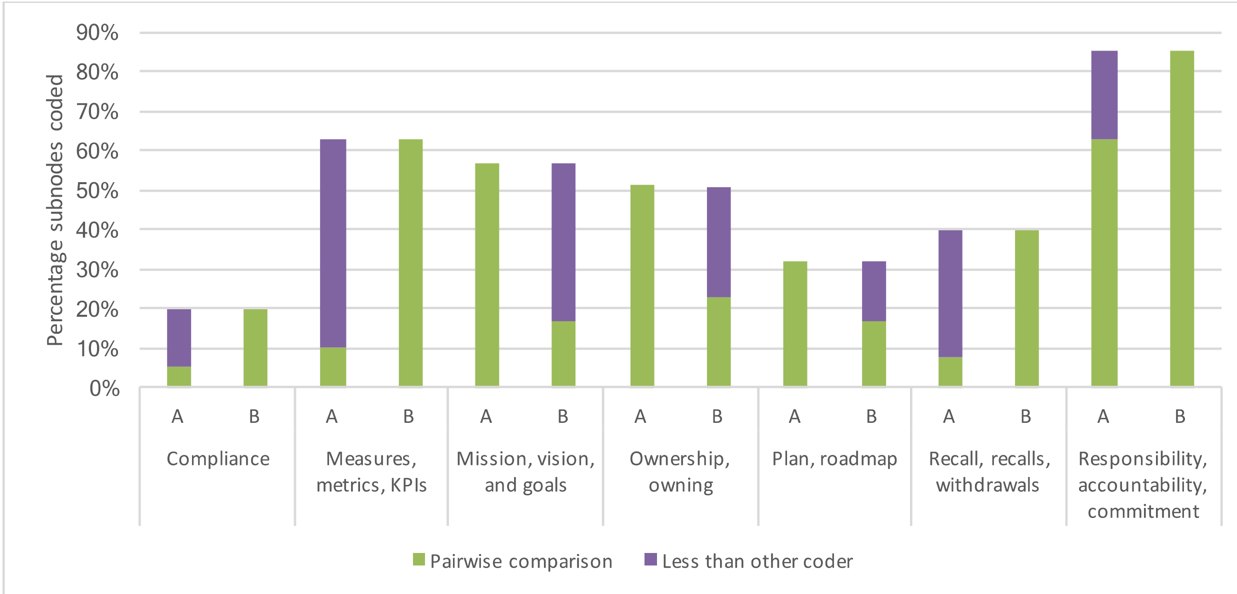


Figure 5: Values and Mission by sub-node and by coder (A and B = two different coders).

For Risk Awareness (Figure 6), most of the difference comes from the sub-node "Risks and Hazards." Coder A coded 29.75% more in this sub-node than coder B. In looking at the verbatim, it shows that HACCP, risk assessment, contamination, foreign material, CCP, specific foreign material findings, food security were examples of words and phrases being coded. Generally, Coder A has more detailed word coding on hazards and risks and Coder B coded specific bacteria references and risks and hazards more generally.

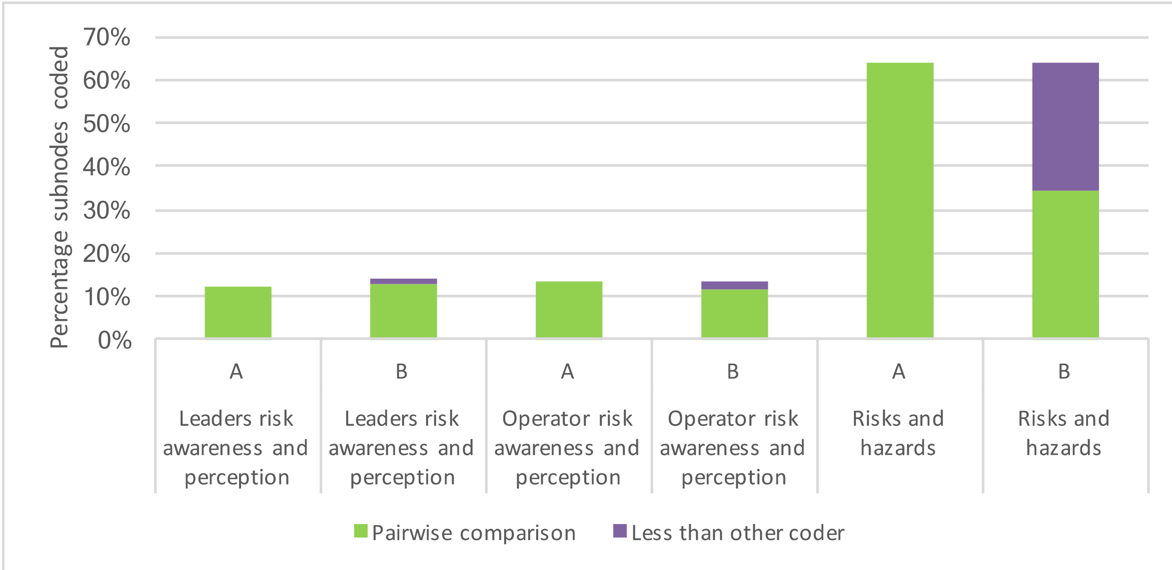


Figure 6: Risk by coder and sub-nodes (A and B = two different coders).

3.3 Coding Discrimination and Cluster Analysis

To investigate if data from the coding framework and process can discriminate between the food safety culture dimensions a cluster analysis of the coded sections of the verbatim content was completed (Figure 7). The Pearson's coefficient shows values at or equal to 0.5 or above for similar items and values less than of 0.5 or less for items distinctly different. The distinctly different items were discussed by the coders and the coding framework was updated. As such, eight major "stems' of similar word content were identified, (1) Rewards and Celebration, (2) Technology and Data, (3) Risks and Hazards, (4) Actions/NCs, (5) Training, education, learning proficiency (6) A group of items related to, vision, mission, values, improvements, consequences, awareness, and ownership (7) Team, and (8) Pride and Recall. The eight "stems" can be directly aligned to the five dimensions but also add more structure to the sub-nodes. This suggested dimensional framework (Figure 8) raises interesting questions that can be useful in the assessment of maturity e.g., what is the connection between pride and recall? What is driving similarity between leaders and employee risk awareness and change, communication, and responsibility? The revised sub-nodes help get closer to some of the manifest data in the texts analyzed. For example, original sub-node was worded as ‘mission, vison, and goals’ this lead to significant discrepancy between coder A and B (figure 5). By revising this sub-node to two sub-nodes ‘direction’ and ‘goal’ the coders were able to meet the standard of 90% agreement and the content coded provided more clarity as for how the organization set both direction and goals or not. In other words, more accuracy in coding by individual coders was gained using these revised sub-nodes and this allowed not only better consistency between the coders but also more detail to be identified from the data, thereby adding to the overall analysis of an organizations food safety culture maturity.

Figure 7: Nodes clustered by word similarity

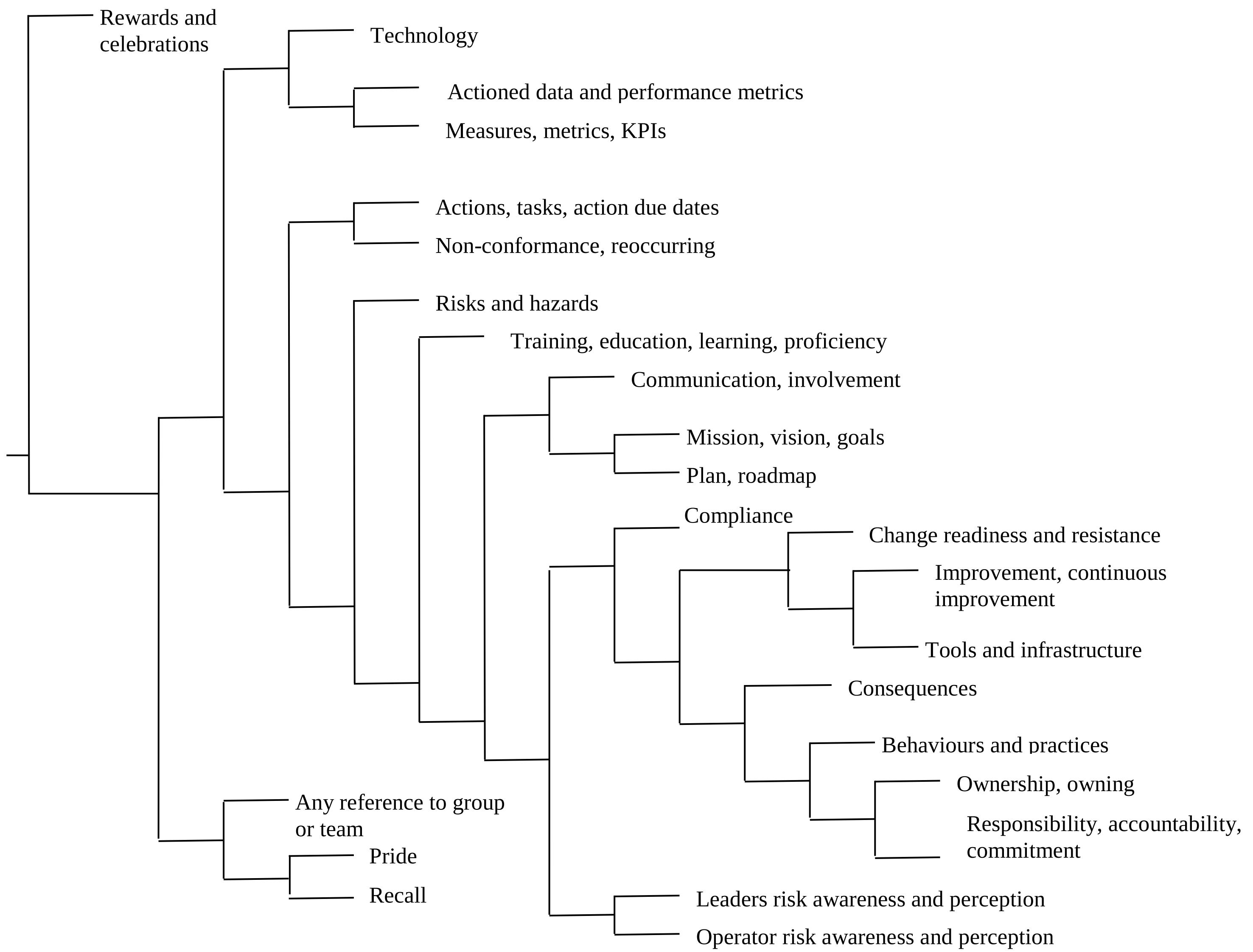
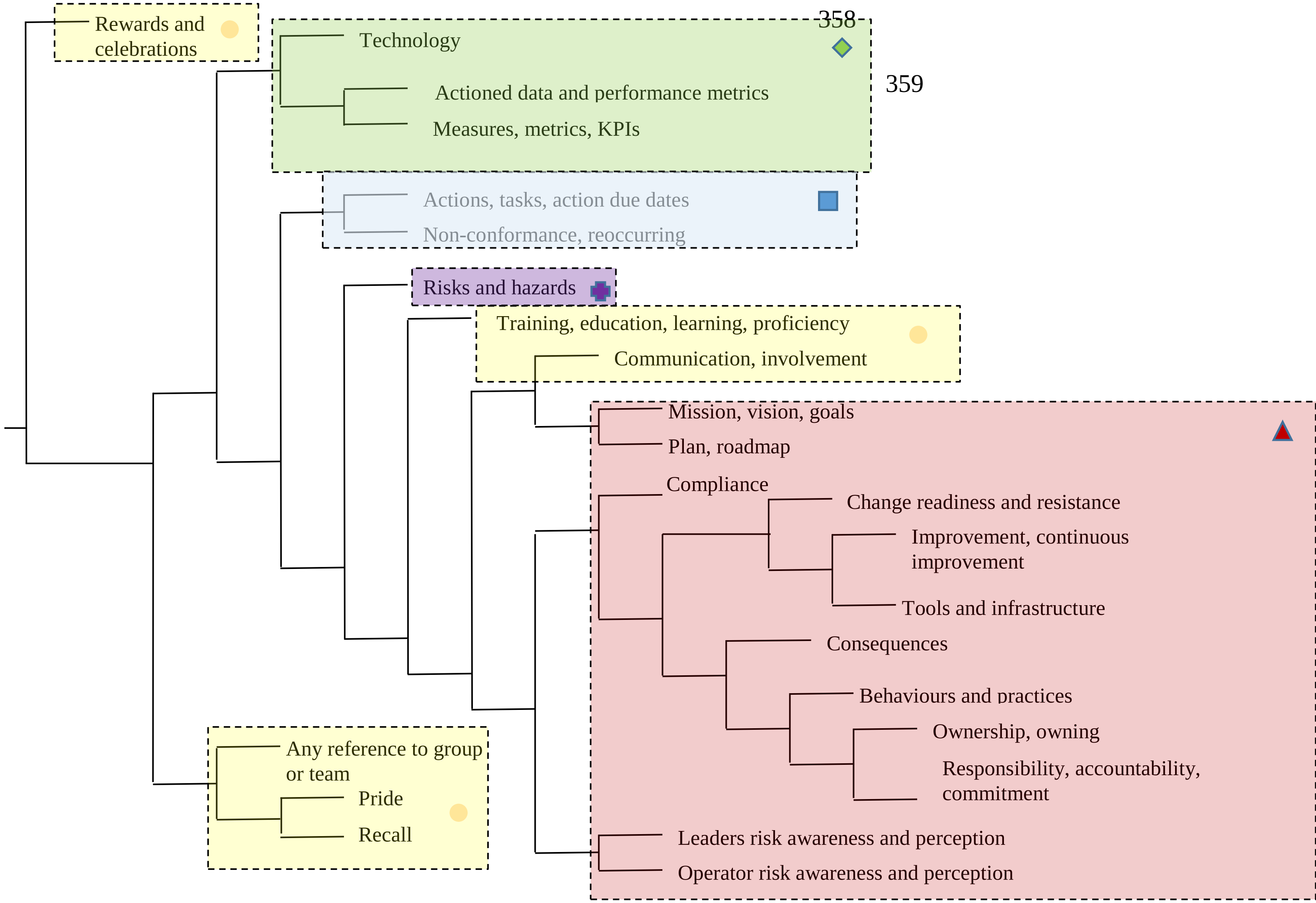

Rewards and celebrations
Technology
Actioned data and performance metrics
Measures, metrics, KPIs
Actions, tasks, action due dates
Non-conformance, reoccurring
Risks and hazards
Training, education, learning, proficiency
Communication, involvement
Mission, vision, goals
Plan, roadmap
Compliance
Change readiness and resistance
Improvement, continuous improvement
Tools and infrastructure
Consequences
Behaviours and practices
Ownership, owning
Responsibility, accountability, commitment
Any reference to group or team
Pride
Recall
Leaders risk awareness and perception
Operator risk awareness and perception

Figure 8: Revised dimension framework and sub-nodes based on cluster analysis. Ledger: Red (▲) = Vision and Mission, Yellow (●) = People, Green (◆) = Consistency, Blue (■) = Adaptability, and Purple (✚) = Risks and Hazards.



Rewards and celebrations
Technology
Actioned data and performance metrics
Measures, metrics, KPIs
Actions, tasks, action due dates
Non-conformance, reoccurring
Risks and hazards
Training, education, learning, proficiency
Communication, involvement
Mission, vision, goals
Plan, roadmap
Compliance
Change readiness and resistance
Improvement, continuous improvement
Tools and infrastructure
Consequences
Behaviours and practices
Ownership, owning
Responsibility, accountability, commitment
Leaders risk awareness and perception
Operator risk awareness and perception
Any reference to group or team
Pride
Recall

358

359

3.4 Content Analysis comparison – performance documents and interviews

A comparison of data from the performance documents and interview transcripts was completed to investigate if method triangulation increases the validity and quality/trustworthiness of food safety culture evaluation (Figure 9). Except for audit reports which include reproduction of requirements from respective standards, performance documents, mean word count ranges between 767 – 1,986 per document depending on document type compared to interview transcripts mean word count between 4,601 – 7,369 per transcript depending on function. Food safety and Quality interviews were generally longer than Manufacturing. As such, it was to be expected that content of the interview transcripts was more detailed and targeted for the purpose. The chart shows that more content was coded in the interviews than in the performance documents except for the dimension “people systems.” This is interesting as most of the documents submitted for analysis were technical in nature e.g., audit reports, meeting minutes, and inspection reports. Still these documents provide valuable data related to people systems, specifically rewards and celebrations, teams, knowledge, and learning.

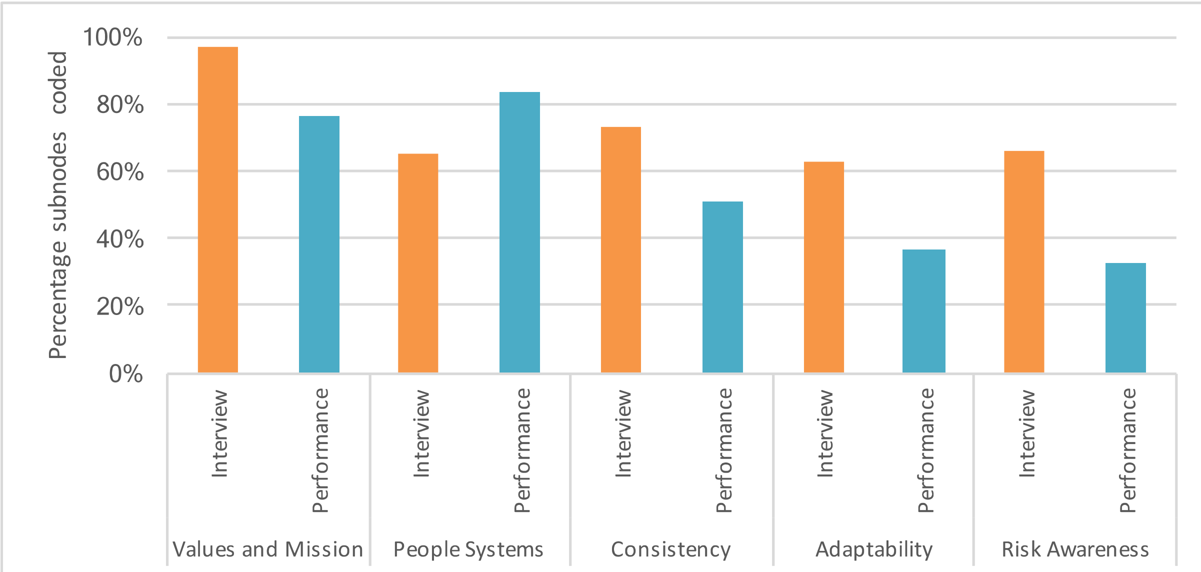


Figure 9: Coding by document type by dimension

3.5 Plant discrimination – method triangulation.

The triangulation analysis revealed a difference between and within plants. Based on the coding consistency and discrimination it was concluded that the coding process is a valid method for evaluating food safety culture. Based on this conclusion three scores per plant were plotted on the maturity model (Figure 10). This shows some disparity both within and between plants. The results for P2 and P5 have the least difference between methods. This means that the individuals rating of food safety maturity, the documented performance, and what was said by leaders in conversation are telling similar stories. In a reevaluation situation, it could be considered to only apply one of the three methods to save time and effort. P3 shows the greatest difference between methods. This means that individuals rate the plants food safety maturity significantly higher than what was found in documented data and what was being said by leaders. In follow up, it would be important to schedule more interviews and focus groups to better understand this difference as a scale does not provide a complete picture to help the plant change. P1 and P4 have comparatively low scores for the documented performance compared with their other measures and it might be interesting to look at the purpose of the submitted documents and if there is an opportunity to better used these; however, what was evaluated by the individual and said by leaders are relatively close, particularly in P1, P5 and, to a lesser extent, P2. P1 is especially interesting as leaders appear to evaluate maturity directionally higher than all employees. This reflects the findings in earlier study with a significant difference between leaders and supervisor (Jespersen et al., 2016)

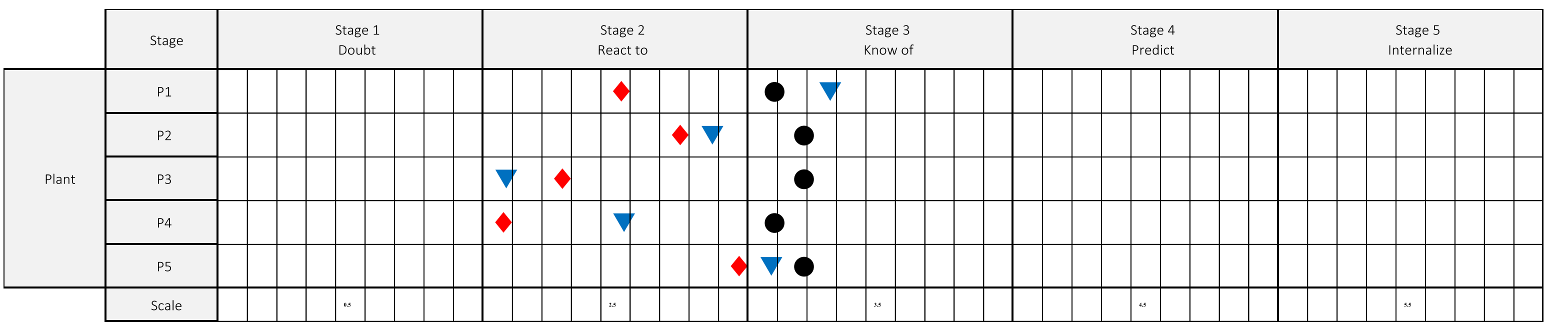


Figure 10: Plant Maturity - Plot of mean values as per method triangulation. Ledger: Dot = Self-assessment scale result, Diamond – Performance document coding result, and Triangle = Interview coding result.

4.0 Discussion and conclusion

The objective of this research was to develop and apply method triangulation to increase validity of food safety culture evaluation results. Data from multiple sources were collected and evaluation results from each plotted on a food safety culture maturity model. Data were analyzed for inter-coder and construct validity, and capability of discrimination within a food safety culture maturity profiling system. Results from analysis of data from three methods, self-assessment scale, document content analysis, and semi-structured interviews, were aggregated and plotted on a food safety culture maturity scale. The dispersion between the mean results per method per plant confirms the need to apply triangulation to get an accurate and trustworthy evaluation of food safety culture. With use of just one of the methods applied in this research the stage of maturity would have been evaluated either too low or too high and subsequent tactical interventions would not have been as effective as intended. For example, a learning program for frontline supervisors in stage 2 "reactive" is largely about creating a personal connection to build a strong foundation of "why food safety is important to you?" A program in stage 3 "knowing" is mostly about increasing cognitive capacity for solving problems, finding root causes, and removing issues permanently. These are two very different objectives that, if applied to the wrong stage, would likely fail and be seen as not valuable to business results. The results showed that mean maturity for all plants was generally higher when assessed through the self-assessment scale ranging from 3.06 – 3.18. The results from the semi-structured interviews were closer to the self-assessment scale for two plants and lower than the self-assessment scores for the other three plants. It was also found that results from the food safety and quality leader interviews generally rated maturity higher than that for manufacturing leaders. The findings from the two functions were found to be significantly different both in maturity assessments and amount of textual data. Mean maturity scores derived from the textual data were the lowest of the three measures except for one plant. In general, more action content (e.g., tasks, follow up) was captured in the textual data and this was to be expected given the original purposes of the documents e.g., meeting minutes and inspection reports.

A coding framework was applied to derive data via content and textual analysis. The framework was consistently applied by two researchers within 90% agreement except for two dimensions; Values and Mission and Risks and Hazards. This difference called for clarification and better definition of the sub-nodes e.g., “mission” this sub-node is better defined as “direction” and can include content related to mission, vision, strategies and generally where a specific direction for food safety is documented. In the Risks and Hazards dimension it was found that one coder coded very specific words e.g., hazards, CCP. It is worth noting that this coder has a long and detailed background in defining hazard and risk management strategies and was likely influenced by this in the coding. This underlines the importance of the iterative coding process with the two checks for consistency; however, it also questions if Risks and Hazards is, in fact, a stand-alone dimension. Is content related to “hazards” and “CCPs” relevant for evaluating culture? Because of this issue and the fact that only two systems (De Boeck et al., 2017; Wright, 2013) have separated out Risks as a stand-alone dimension (Jespersen, Griffith, and Wallace, 2017), it is worth discussing if this dimension should remain in the food safety culture theoretical framework (Figure 1) or if is best considered in the evaluation of food safety management systems.

This study was conducted as part of a larger study with 21 plants but this analysis was completed with data from a sub-set of five. This was done both to ensure that there was enough time to execute the coding process fully on 10 interview transcripts and 20 performance documents by two researchers and to analyze a sufficiently large sample for triangulation purposes. It is recommended that more work is done with more researchers to promulgate content analysis as a method for evaluating both food safety performance and food safety culture maturity. It was unexpected that such similarity would be found in the five plants, where performance ranged from stage 2 maturity “reactive” to stage 3 “know” (Jespersen & Edwards, Under review; Jespersen et al., 2016) or all plants and documents. This could be due to the geographical dispersion of the plants, this subset all being in North America, and therefor under similar North American legal systems and customer expectations. It could also be a case of selection bias as the participating companies were not gathered via randomization or quasi-random assignment, rather through senior leader interest and board willingness to participate in the research. In this research, selection would be present if those who participated in the study and responded to the survey are those that have internalized the importance of culture and/or those that engage in “cheap talk" about culture. It is reasonable to assume some sampling bias due to the voluntary nature of the participants.

In summary, the research adds information and knowledge, derived through a transparent and rigorous process, to the food safety culture domain. Specifically, it adds proof that reliance on a single method for evaluation food safety culture can give inaccurate results and should be treated with caution. This has practical significance for companies who invest, not just in such results, but in subsequent improvement tactics.

Acknowledgements

The authors wish to thank representatives from the five companies who participated in the study. For taking the time to provide data and guidance in this important endeavor.

References

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
Bauwens, A. (2010). The Use of Method Triangulation in Probation Research. *European Journal of Probation, 2*(2), 39-52. doi:doi:10.1177/206622031000200204
Berg, B. L., Howard. (2012). *Qualitative Research Methods for the Social Sciences* (Vol. 8th): Pearson Education Inc.
Brinkmann, S. K., Steinar. (2015). *InterViews: learning the craft of qualitative research interviewing*: Sage Publications.
Campbell, D. a. F., D. . (1959). Convergent and discriminant validation by the multitrait - multi method matrix. *Psychological bulletin, 56*, 81-105.
Carugi, C. (2016). Experiences with systematic triangulation at the Global Environment Facility. *Evaluation and Program Planning, 55*, 55-66. doi:http://dx.doi.org/10.1016/j.evalprogplan.2015.12.001
Creswell, J. W. (1998). *Qualitative Inquiry and Research Design : Choosing Among Five Traditions*. Thousand Oaks, Calif: Sage Publications.
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
De Boeck, E., Mortier, A. V., Jacxsens, L., Dequidt, L., & Vlerick, P. (2017). Towards an extended food safety culture model: Studying the moderating role of burnout and jobstress, the mediating role of food safety knowledge and motivation in the relation between food safety climate and food safety behavior. *Trends in Food Science & Technology*, 1-13.
Denison, D., Hooijberg, R., Lane, N., & Lief, C. (2012). *Leading Culture Change in Global Organizations: Aligning Culture and Strategy*. John Wiley & Sons.
Denison, D. R. (1997). *Corporate Culture and Organizational Effectiveness* (Vol. 2nd edition (May 1997)): Aviat.
Denison, D. R., & Mishra, A. K. (1995). Toward a Theory of Organizational Culture and Effectiveness. *Organization Science, 6*(2), 204-223.
Denzin, N. K. (1970). *The Research Act: A Theoretical Introduction to Sociological Methods*: Transaction.
Denzin, N. K. (2012). Triangulation 2.0. *Journal of Mixed Methods Research, 6*(2), 80-88. doi:10.1177/1558689812437186

Denzin, N. K., & Lincoln, Y. S. (2011). *The Sage handbook of qualitative research* (4th ed.).
Floyd, J. (1993). The use of across-method triangulation in the study of sleep concerns in healthy older adults. *ANS Adv Nurs Sci., December*(16), 70-80.
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
Hofstede, G. (1980). *Culture's Consequences – International differences in work-related values.* : Sage Publications, London.
Hofstede, G. (2001). *Culture's Consequences – Comparing Values, Behaviours, Institutions and Organizations Across Nations*: Thousand Oaks, Sage Publications.
Hofstede, G. (2013). Values Survey Module 2013, Manual and VSM Questionnaire – English Version. Retrieved from http://geerthofstede.com/research-and-vsm/vsm-2013/
Holstein, J. A. G., J. F. (1995). *Qualitative Research Methods: The active interview*: SAGE Publications Ltd.
Jespersen, L., & Edwards, A. M. (Under review). Predictive Attributes of Food Safety Culture and Climate. *Food Research International*.
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., T., M., & Vlerick, P. (2017). Development and validation of a scale to capture social desirability in food safety culture. *Food Control*.
Kopinak, J. K. (1999). The Use of Triangulation in a Study of Refugee Well-Being. *Quality and Quantity, 33*(2), 169-183. doi:10.1023/a:1026447822732
Lombard, M., Snyder-Duch, J., & Bracken, C. C. (2002). Content Analysis in Mass Communication: Assessment and Reporting of Intercoder Reliability. *Human Communication Research, 28*(4), 587-604. doi:10.1111/j.1468-2958.2002.tb00826.x
Miles, M. B. H., A.M. (1994). *Qualitative data analysis: an expanded sourcebook* (2nd ed.): Thousand Oaks: Sage Publications.
Neuendorf, K. A. (2002). *The content analysis guidebook*: Thousand Oaks : Sage Publications.
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., & Schein, P. (2017). *Organizational Culture and Leadership*: Wiley.
Srinivasan, A., & Kurey, B. (2014). Creating a Culture of Quality. *Harvard Business Review, 92*(4), 23-25.
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.
Wallace, C. A. (2009). *The impact of personnel, training, culture and organisational factors on the application of the HACCP system for food safety management in a multinational organisation, .* (PhD Thesis).
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. (2013). A tool to diagnose culture in food business operators, Food Standards Agency Research Report.
Yeasmin, S., & Rahman, K. F. (2012). 'Triangulation' Research Method as the Tool of Socical Science Research. *BUP Journal, 1*(1), 154-163.