Evaluating and Predicting Maturity of Food Safety Culture in Food Manufacturing.

A Thesis
Presented to
The Faculty of Graduate Studies

The University of Guelph

by

Lone Jespersen

In partial fulfilment of requirements
for the degree, of
Ph.D.
in
Food Science

Guelph, Ontario, Canada
© Lone Jespersen, May, 2017
ABSTRACT

EVALUATING AND PREDICTING MATURITY OF FOOD SAFETY CULTURE IN FOOD MANUFACTURING.

Lone Jespersen
University of Guelph, 2017

Advisor: Professor T. MacLaurin

Food safety culture might well be the most discussed item in modern food safety. As such, many offer opinions on what food safety culture is without offering science-based arguments and subsequently create more confusion than resolution. The research presented seeks to further the conversation through development of a valid and predictable system for evaluating food safety culture in food manufacturing. The researcher proposes five dimensions that can help segment the topic of food safety culture. These were derived from a comparative analysis of eight existing systems designed to evaluate either organizational culture or food safety culture. The dimensions were applied through method triangulation to evaluate maturity of food safety culture in five North American food companies. Analysis of the data proved the scale’s predictive validity as well as the validity and necessity of assessing social desirability in the food safety context. The research concludes with an updated food safety maturity model that provides researchers and practitioners alike with a roadmap to maturing food safety culture.
along four of the five suggested dimensions. Finally, the researcher proposes a culture model for defining critical actions between food safety cultural building blocks and calls for more research in quantifying business impact of the maturity of these critical actions.
Dedication

To Jakob for your never wavering belief in me, to Henrik for showing the world that out of pain grows but strength, and to morfar, inspired by your memory in my heart for always.

Acknowledgment

I would like to acknowledge Bush Brothers & Company, Cargill Protein Group, Cargill Corn Milling, Land O’ Frost, and Land O’ Lakes, Inc. for their leadership and openness to allow this research to take place.

I wish to thank my advisory committee and co-authors, Mansel Griffiths, Tanya MacLaurin, Ben Chapman, Carol A. Wallace and my co-authors Peter Vlerick, Michelle Edwards, Jeff Taylor, Greg Holler, John Butts, Sara Mortimore, and Dave Harlan for all your support and guidance throughout this amazing journey.
# Table of Content

1. **Introduction** .................................................................................................................. 1

   1.1. Organizational culture .................................................................................................. 1

   1.2. Failure of organizational cultures in food related recalls ........................................ 3

   1.3. Food safety culture literature ....................................................................................... 3

   1.4. Maturity and capability areas of food safety culture ................................................... 4

   1.5. Research description and hypotheses ......................................................................... 12

   1.6. Thesis structure ........................................................................................................... 13

2. **Comparative analysis of existing food safety culture evaluation systems** ................. 14

   2.1. Abstract ....................................................................................................................... 14

      2.1.1. Purpose .................................................................................................................. 14

   2.2. Keywords .................................................................................................................... 14

   2.3. Highlights ................................................................................................................... 14

   2.4. Introduction .................................................................................................................. 16

      2.4.1. Research quality ................................................................................................. 18

      2.4.2. Research trustworthiness .................................................................................... 19

      2.4.3. Valid and reliable research .................................................................................. 19

   2.5. Material and method .................................................................................................... 22

      2.5.1. Sample ................................................................................................................ 22

      2.5.2. Method ............................................................................................................... 22

      2.5.3. Research trustworthiness .................................................................................... 26

   2.6. Results .......................................................................................................................... 27
## 2.6.1. Summary of in-scope culture evaluation systems

Page 27

## 2.6.2. Differences in validation strategies

Page 32

## 2.6.3. System content comparison

Page 34

## 2.6.4. Suggested framework to unify the research field

Page 37

## 2.7. Discussion

Page 38

### 3. Development and validation of a scale to capture social desirability in food safety

Page 42

#### 3.1. Abstract

Page 42

#### 3.2. Keywords

Page 42

#### 3.3. Highlights

Page 42

#### 3.4. Introduction

Page 44

##### 3.4.1. Organizational culture and food safety failures leading to contamination

Page 44

##### 3.4.2. Social desirability

Page 47

##### 3.4.3. Research objectives

Page 49

#### 3.5. Material and method

Page 50

##### 3.5.1. The food safety desirability response scale (FSDRS)

Page 50

##### 3.5.2. Ethical considerations

Page 51

##### 3.5.3. Data collection and sample

Page 51

##### 3.5.4. Analysis

Page 52

#### 3.6. Results

Page 53

##### 3.6.1. Factorial validity and reliability

Page 53

##### 3.6.2. Construct validity

Page 57

##### 3.6.3. Discriminant validity

Page 58

#### 3.7. Discussion

Page 59
4. Predictive attributes of food safety culture and climate.................................62
   4.1. Abstract........................................................................................................62
   4.2. Keywords.......................................................................................................62
   4.3. Highlights .....................................................................................................62
   4.4. Introduction ..................................................................................................63
   4.5. Material and method.....................................................................................68
       4.5.1. Participants. .........................................................................................68
       4.5.2. Scale. .....................................................................................................69
       4.5.3. Analysis suitability. ..............................................................................69
   4.6. Results...........................................................................................................70
       4.6.1. Factor analysis. ....................................................................................71
       4.6.2. Predictive validity. ................................................................................76
       4.6.3. Suggested food safety culture and climate model based on predictive validity....77
   4.7. Discussion......................................................................................................78

5. Triangulation and the importance of establishing valid methods for food safety
   culture evaluation ...............................................................................................81
   5.1. Abstract........................................................................................................81
   5.2. Keywords.......................................................................................................81
   5.3. Highlights .....................................................................................................82
   5.4. Introduction ..................................................................................................83
       5.4.1. Theoretical framework. ..........................................................................84
       5.4.2. Method triangulation. ............................................................................85
   5.5. Material and method.....................................................................................87
       5.5.1. Methods strengths and weaknesses. ......................................................89
5.5.2. Response analysis of self-assessment scale ................................................................. 90
5.5.3. Content analysis of performance documents ............................................................. 91
5.5.4. Content analysis of semi-structured interviews ......................................................... 91
5.5.5. Content coding ............................................................................................................ 93
5.5.6. Data triangulation ........................................................................................................ 96
5.6. Results ............................................................................................................................ 97
  5.6.1. Self-assessment results ............................................................................................. 97
  5.6.2. Coding comparisons ............................................................................................... 98
  5.6.3. Coding discrimination and cluster analysis ............................................................... 100
  5.6.4. Content analysis comparison – performance documents and interviews ............ 104
  5.6.5. Plant discrimination – method triangulation ............................................................. 105
5.7. Discussion ....................................................................................................................... 107
6. A path to maturing culture for realizing financial gain ..................................................... 110
  6.1. Abstract ......................................................................................................................... 110
  6.2. Keywords ..................................................................................................................... 110
  6.3. Highlights ..................................................................................................................... 110
  6.4. Introduction ................................................................................................................ 111
    6.4.1. Cultural impact on organizational performance and effectiveness ..................... 111
    6.4.2. Organizational culture and quality maturity ....................................................... 113
    6.4.3. Organizational culture and food safety ............................................................... 114
  6.5. Material and method .................................................................................................... 115
    6.5.1. Response analysis of self-assessment scale .......................................................... 116
    6.5.2. Content analysis of performance documents ....................................................... 117
    6.5.3. Content analysis of semi-structured interviews ................................................... 118
6.6. Results .............................................................................................................................................. 118

6.6.1. Company aggregated results and baseline. .................................................................................. 123

6.6.2. Path to food safety maturity. ......................................................................................................... 127

6.7. Discussion ........................................................................................................................................... 133

7. Discussion of findings ............................................................................................................................ 136

7.1. Comparative analysis of existing food safety culture evaluating systems .................. 136

7.2. Triangulation and the importance of establishing valid methods for food safety culture evaluations .......................................................................................................................... 141

7.2.1. Scale validity and predictive attributes of food safety culture and climate . . . . . . . 141

7.2.2. Socially desirable responding .................................................................................................... 143

7.2.3. Triangulating food safety culture evaluations ........................................................................ 145

7.3. A path to realizing value by maturity food safety culture ...................................................... 147

7.4. Evaluating and predicting effectiveness and maturity of food safety culture in food manufacturing 148

7.5. Limitations of this study and recommendations for future research. .......................... 151

8. Conclusion ............................................................................................................................................. 152

9. References ............................................................................................................................................. 154
List of Tables

Table 1: Cultural dimensions and components of organizations, adapted from (E.H. Schein, 2004). ................................................................................................................................................. 2

Table 2: Food safety maturity model. .................................................................................................................. 5

Table 3: Capability areas, cultural dimensions, and food safety descriptors used to profile food safety culture. ................................................................................................................................. 11

Table 4: Content coding framework. .................................................................................................................... 24

Table 5: Differences between the validation strategies applied in the eight culture evaluation models. .................................................................................................................................................. 32

Table 6: Content comparison of the eight culture evaluation systems. ................................................................. 35

Table 7: The food safety desirability response statements (FSDR-scale). ......................................................... 50

Table 8: Output from parallel analysis.................................................................................................................. 53

Table 9: Results of the oblimin rotated principal factor analysis of FSDRS-items ............................................ 56

Table 10: Correlations among the food safety desirability response subscales ................................................... 57

Table 11: Sub-scale factor analysis for FSQ leaders in a high level of maturity .................................................. 71

Table 12: Sub-scale factor analysis for FSQ leaders in a low level of maturity ..................................................... 72

Table 13: Sub-scale factor analysis for manufacturing leaders in a high level of maturity ............................... 73

Table 14: Sub-scale factor analysis for manufacturing leaders in a low level of maturity ............................... 73

Table 15: Sub-scale factor analysis for manufacturing supervisors in a high level of maturity ....................... 74

Table 16: Sub-scale factor analysis for manufacturing supervisors in a low level of maturity ....................... 74

Table 17: Sub-scale factor analysis for FSQ supervisors in a high level of maturity ....................................... 75

Table 18: Sub-scale factor analysis for FSQ supervisors in a low level of maturity ....................................... 75

Table 19: Sources by plant and data type for the five plants in the triangulation validation ........................... 88
Table 20: Coding framework used in the content and textual analysis. Adapted from Jespersen, Griffiths, and Wallace (2017). ................................................................. 92

Table 21: Plant results from maturity self-assessment scale, total and by dimension. .................. 97

Table 22: Financial performance differences between companies who invested in a performance-enhancing culture and those that did not (Kotter, 1992). ........................................ 112

Table 23: Data collected from the five participating companies. .................................................. 115

Table 24: Aggregated company demographics, culture performance results, and baseline (mean and total). ......................................................................................... 123

Table 25: Maturity model glossary. ......................................................................................... 127

Table 26: Food safety culture - maturity model version 2.0 ...................................................... 130

Table 27: Culture evaluation systems comparative analysis. .................................................... 138
# Table of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Thesis structure</td>
<td>13</td>
</tr>
<tr>
<td>2</td>
<td>Six step process applied in the comparative analysis</td>
<td>23</td>
</tr>
<tr>
<td>3</td>
<td>Food safety culture - dimensional framework</td>
<td>38</td>
</tr>
<tr>
<td>4</td>
<td>Predictive attributes of food safety culture and climate</td>
<td>78</td>
</tr>
<tr>
<td>5</td>
<td>Five dimensions to the culture evaluation system (Jespersen, Griffiths, and Wallace, 2017)</td>
<td>85</td>
</tr>
<tr>
<td>6</td>
<td>Methods and data triangulation applied to evaluate of food safety culture</td>
<td>88</td>
</tr>
<tr>
<td>7</td>
<td>Coding process applied to deriving data through content analysis</td>
<td>95</td>
</tr>
<tr>
<td>8</td>
<td>Codes by dimension with pairwise comparison and difference by coder</td>
<td>98</td>
</tr>
<tr>
<td>9</td>
<td>‘Values and mission’ dimension by sub-node and by coder</td>
<td>99</td>
</tr>
<tr>
<td>10</td>
<td>‘Risk awareness’ dimension by coder and sub-nodes</td>
<td>100</td>
</tr>
<tr>
<td>11</td>
<td>Nodes clustered by word similarity</td>
<td>102</td>
</tr>
<tr>
<td>12</td>
<td>Revised dimension framework and sub-nodes based on cluster analysis</td>
<td>103</td>
</tr>
<tr>
<td>13</td>
<td>Coding by document type by dimension</td>
<td>105</td>
</tr>
<tr>
<td>14</td>
<td>Plant maturity - mean values as per method triangulation</td>
<td>106</td>
</tr>
<tr>
<td>15</td>
<td>Methods and data triangulation applied to evaluate of food safety culture</td>
<td>116</td>
</tr>
<tr>
<td>16</td>
<td>Culture model: Organizational effectiveness, culture building blocks, and actions</td>
<td>121</td>
</tr>
<tr>
<td>17</td>
<td>Annualized sales per company and COPQ based on evaluation result (bar), one maturity stage up (dot), and one stage down (diamond)</td>
<td>127</td>
</tr>
<tr>
<td>18</td>
<td>Food safety culture theoretical framework</td>
<td>139</td>
</tr>
<tr>
<td>19</td>
<td>Food safety culture and climate model</td>
<td>143</td>
</tr>
<tr>
<td>20</td>
<td>Food safety culture evaluation – method triangulation</td>
<td>145</td>
</tr>
</tbody>
</table>
Figure 21: Organizational effectiveness with cultural building blocks and actions. .................... 149
1. Introduction

By 2035, the number of people over the age of 60 will have doubled from 11% to 22% of the total global population (United Nations & Social, 2012). Although this increase in the elderly population is generally seen as an indicator of global health, it is also indicative of a growing number of people vulnerable to infectious diseases, such as foodborne infections or intoxications (International Union of Food Science and Technology, 2015). This, along with other disease trends, e.g., a 1.5-fold increase in the number of cases of diabetes expected during the same period (International Diabetes, 2014) and continued foodborne illness outbreaks and recalls will maintain food safety as paramount for the near future. It is known that older and immunocompromised members of society are more susceptible to foodborne illness (Centers for Disease & Prevention, 2014; Public Health Agency of, 2014; World Health, 2014). As such, the issue remains how to minimize population exposure to foodborne pathogens. Many outbreaks of foodborne illness are linked to a breakdown in hygiene because of improper human behaviours or a lack of an appropriate food safety culture (Baur, Getz, & Sowerwine, 2017; Canadian Food Inspection, 2013b; Jespersen & Huffman, 2014).

1.1. Organizational culture

Definition and measurement of food safety culture can be rooted in the definition, dimensions, and characteristics of organizational culture. Schein (2004) defines organizational culture as,

“a pattern of shared basic assumptions that was learned by a group as it solved its problems. The group found these assumptions to work 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.”

Cultural dimensions and characteristics (Table 1) adapted from Schein’s work serve as a theoretical background to characterize an individual organization’s food safety culture (Schein, 2004).

Table 1: Cultural dimensions and components of organizations, adapted from (E.H. Schein, 2004).

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>External adaptation</td>
<td>Mission and goals, means (e.g., day-to-day behaviours, skills, knowledge, time and technology) to reach goals, degree of autonomy, how the organization decide what to measure, measures (what and how), how to judge success, remediate and repair processes, and crisis history</td>
</tr>
<tr>
<td>Internal integration</td>
<td>System of communication, common language, group selection and exclusion criteria, allocation systems (e.g., influence, power and authority), rules for relationships and systems for rewards and punishment</td>
</tr>
<tr>
<td>Reality and truth</td>
<td>High vs. low context, definition of truth, information, data, and knowledge needs; training and competencies; systems (e.g., sign-off), continuous improvement</td>
</tr>
<tr>
<td>Time and space</td>
<td>Four different dimensions for characterizing time orientation, assumptions around time management</td>
</tr>
<tr>
<td>Human nature, activity and relationship</td>
<td>Theory x/y managers, the doing/being/being-in-becoming orientation, and four basic problems solved in a group: identity and role; power and influence; needs and goals; acceptance and intimacy, individualism/groupism, power distance and accepted behaviours and practices</td>
</tr>
</tbody>
</table>
1.2. Failure of organizational cultures in food related recalls

A review of four recalls between 2008 and 2013 was completed to investigate connections between failing organizational cultures and consumer and business impacts. These four recalls were selected for this review, as food safety culture was specifically identified in the investigation reports as one of the reasons for the recall. The review of each recall, Fonterra, XL Foods, Jensen Farms, and Maple Leaf Foods, revealed failures specific to a range of cultural dimensions and similarities were found across the four recalls related to organizational culture. Based on Schein’s (2004) culture dimensions, failures were mostly found in the four of the five cultural dimensions; ‘External adaptation’, ‘Internal integration’, ‘Reality and truth’ and ‘Human nature, activity, and relationship. First, related to ‘External adaptation’ and ‘Internal integration’, assumptions were made around communication between stakeholders (e.g., inspectors and employees) and between functions (e.g., levels and silos). Secondly, related to ‘Reality and truth’, failure to derive information from collected data, such as identifying trends and focus areas from data, were not in place as a measure for how to judge success, remediate, or continuously improve. Thirdly, ‘Human nature, activity, and relationship’, acceptance of bypassing written procedures and practices was found in all cases. Lastly, although not directly related to a cultural dimension, it was also noted that regulators, independent of country, acknowledged that the outbreaks highlighted the need to improve their regulatory systems and practices.

1.3. Food safety culture literature

A review of the literature was conducted to identify gaps in current research and to contrast and compare content and measurement systems in current literature. Authors like Chris
Griffith, Doug Powell, and Frank Yiannas have described food safety culture in general terms in their respective publications and give insights into different ways of characterizing food safety culture and its impact on food manufacturers (Griffith, 2010; Powell, Jacob, & Chapman, 2011; Yiannas, 2009). These authors have been instrumental in bringing attention to the topic. Griffith (2010) and Yiannas (2009) offer an introduction to components that make up food safety culture in food service organizations. Other literature suggests practical ways to improve food safety culture specifically in food manufacturing (Clayton, Griffith, Price, & Peters, 2002; Hanacek, 2010; Jespersen & Huffman, 2014; Seward, 2012). Common across all these research sources were gaps in standardization and validation of methods to measure food safety culture and general recognition of the importance of group development.

There is limited empirical and confirmatory research completed in the field of food safety culture and food manufacturing with few exceptions (Ball, Wilcock, & Aung, 2009; De Boeck, Jacxsens, Bollaerts, Uyttendaele, & Vlerick, 2016; Hinsz & Nickell, 2015; Jespersen & Huffman, 2014; Wilcock, Ball, & Fajumo, 2011) this indicate that further research is needed to establish effective tools for food safety culture prediction. The collaboration with industry partners is important and further tools are needed to establish effective tools for food safety culture prediction.

1.4. Maturity and capability areas of food safety culture

Organizational cultural dimensions and characteristics can be applied to food safety in the context of a food safety maturity model (Jespersen, Griffiths, Maclaurin, Chapman, & Wallace, 2016) (Table 2).
Table 2: Food safety maturity model.
<table>
<thead>
<tr>
<th>Capability Area</th>
<th>Stage 1</th>
<th>Stage 2</th>
<th>Stage 3</th>
<th>Stage 4</th>
<th>Stage 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived Value</td>
<td>Doubt</td>
<td>React to</td>
<td>Know of</td>
<td>Predict</td>
<td>Internalize</td>
</tr>
<tr>
<td>Completing tasks because regulators make us do so.</td>
<td>Little to no investment in systems (people and processes) to prevent food safety firefighting.</td>
<td>Food safety issues are solved one at a time, getting to the root of the issue, to protect the business.</td>
<td>Reoccurrence of food safety issues is prevented by used of knowledge and leading indicators.</td>
<td>Ongoing business improvement and growth is enabled by food safety.</td>
<td></td>
</tr>
<tr>
<td>Capability Area</td>
<td>Stage name</td>
<td>Stage 1</td>
<td>Stage 2</td>
<td>Stage 3</td>
<td>Stage 4</td>
</tr>
<tr>
<td>-----------------</td>
<td>------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>People System</td>
<td>Doubt</td>
<td>React to</td>
<td>Know of</td>
<td>Predict</td>
<td>Internalize</td>
</tr>
<tr>
<td></td>
<td>Tasks are only completed when senior leader’s demand, without understanding responsibility, the task, or why it is important.</td>
<td>Responsibilities for problems are established as the problems are discovered and solved mostly by use of negative consequences.</td>
<td>Deeper understanding for the importance of foods safety systems, where responsibilities are clearly defined and communicated, is gained one issue at a time.</td>
<td>Develop and assess tools for improving processes through knowledge and data. communication and assessment.</td>
<td>Strategic direction is set across the complete organization with defined accountabilities, responsibilities, and food safety as one of the business enablers.</td>
</tr>
<tr>
<td>Capability Area</td>
<td>Stage 1</td>
<td>Stage 2</td>
<td>Stage 3</td>
<td>Stage 4</td>
<td>Stage 5</td>
</tr>
<tr>
<td>-----------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td><strong>Doubt</strong></td>
<td>React to</td>
<td>Know of</td>
<td>Predict</td>
<td>Internalize</td>
<td></td>
</tr>
<tr>
<td><strong>Process</strong></td>
<td>Unstructured problem solving to remove the immediate pain.</td>
<td>&quot;Plan, Do, Check, Act&quot; with emphasis on control in the check phase and expectation of an immediate 100% perfect solution.</td>
<td>Structure problem solving with significant risk of over analysing.</td>
<td>&quot;Plan, Do, Study, Act&quot; with emphasis on study and not control. Problem solving is accepted as an iterative process.</td>
<td>Horizon scanning and continuous improvement are used to identify risks. Risks inform the development and/or improvement of mitigation plans. Mitigation plans are integrated in the global business management system.</td>
</tr>
<tr>
<td><strong>Thinking</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Capability Area</td>
<td>Stage 1</td>
<td>Stage 2</td>
<td>Stage 3</td>
<td>Stage 4</td>
<td>Stage 5</td>
</tr>
<tr>
<td>-----------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>--------</td>
</tr>
<tr>
<td><strong>Technology</strong></td>
<td><strong>Doubt</strong></td>
<td><strong>React to</strong></td>
<td><strong>Know of</strong></td>
<td><strong>Predict</strong></td>
<td><strong>Internalize</strong></td>
</tr>
<tr>
<td><strong>Enabled</strong></td>
<td>Little technology being adopted and few see this to be an issue.</td>
<td>Responsibility is left to the individual to identify data needed and there is a high reliance on the individual to derive information from the data.</td>
<td>Standard technology is adopted on going and standardized training provided to individuals as needed.</td>
<td>Data is collected in a precise and accurate manner to constantly improve processes.</td>
<td>Integrated, global information systems (e.g., ERP) are in place in the organization making it quick to adapt, improve, and use automated workflows.</td>
</tr>
<tr>
<td>Capability Area</td>
<td>Stage 1</td>
<td>Stage 2</td>
<td>Stage 3</td>
<td>Stage 4</td>
<td>Stage 5</td>
</tr>
<tr>
<td>-----------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>Doubt</td>
<td>React to</td>
<td>Know of</td>
<td>Predict</td>
<td>Internalize</td>
<td></td>
</tr>
<tr>
<td>Tools and</td>
<td>Minimal tools in the hands of few individuals.</td>
<td>It takes a problem to get the right tools. This often leads to findings the right tools in a hurry and resulting in rework.</td>
<td>The organization invests readily in the right tools and infrastructure when solving a problem calls for it.</td>
<td>Food safety tools and infrastructures are in place and are continuously improved for ease of use and cost of the organization.</td>
<td>Investment in tools and infrastructure is evaluated long-term and prioritized along with other business investments.</td>
</tr>
<tr>
<td>Infrastructure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The structure and content of the food safety maturity model was developed based on existing literature on food safety culture, input from an industry expert panel, and learnings from maturity models in other areas e.g., quality and health care (Crosby, 1972; Goonan, Muzikowski, & Stoltz, 2009). The food safety maturity model categorizes generic organizational cultural dimensions into five capability areas central to characterize and measuring food safety culture. The capability areas in the model enable food safety professionals to measure and profile current food safety culture and develop a plan for how to strengthen food safety culture. The five areas are, ‘Perceived value’, ‘People systems’, ‘Process thinking’, ‘Technology enabled’, and ‘Tools and infrastructure’ and each capability area maps directly to one of Schein’s (2004) organizational dimensions (Table 3).

Table 3: Capability areas, cultural dimensions, and food safety descriptors used to profile food safety culture.

<table>
<thead>
<tr>
<th>Capability area</th>
<th>Cultural dimension</th>
<th>Food safety descriptor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived value</td>
<td>External adaptation</td>
<td>Describes the extent to which food safety is a regulatory check (Stage 1) or as critical to business performance (Stage 5).</td>
</tr>
<tr>
<td>People systems</td>
<td>Internal integration and Human nature, activity, and relationship</td>
<td>Describes a task-based organization with signs of miscommunication accountabilities (Stage 1) or clear, accepted accountability set in behaviour-based working groups (Stage 5).</td>
</tr>
<tr>
<td>Process thinking</td>
<td>Human nature, activity, and relationship</td>
<td>Describes how problems are solved (e.g., as independent tasks) (Stage 1) or problem solving is seen as an iterative process built on critical thinking and data (Stage 5).</td>
</tr>
<tr>
<td>Technology enabled</td>
<td>Reality and truth</td>
<td>Describes how the organization turns data into information; manual and independent (Stage 1) compared too automatically and integrated into a company-wide information system (Stage 5).</td>
</tr>
</tbody>
</table>
### Capability area  
### Cultural dimension  
### Food safety descriptor

<table>
<thead>
<tr>
<th>Capability area</th>
<th>Cultural dimension</th>
<th>Food safety descriptor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tools and infrastructure</td>
<td>Reality and truth</td>
<td>Describes how organizations make decisions on food safety investments. Independent cases as a need occur (Stage 1) or planed and based on a holistic, long-term analysis (Stage 5). Can also be illustrated by whether an employee needs to walk far to a sink (Stage 1) or sinks are conveniently located (Stage 5).</td>
</tr>
</tbody>
</table>

Schein’s dimension ‘time and space’ was not included in the maturity model as an independent dimension included in the ‘Perceived value’ capability area.

**1.5. Research description and hypotheses**

This research builds on existing research in the field of organizational culture, food science, and social cognitive science and tests the food safety culture measurement system developed by the researcher in the master’s thesis “Measuring and predicting food safety culture in food manufacturing.” The research aims to validate measurement system across a range of global manufacturing companies through method triangulation by investigating four hypotheses:

**H1:** The food safety culture maturity model and measurement system is a valid and reliable method to evaluate and predict food safety culture.

**H2:** The food safety culture maturity model and measurement system adds additional information in comparison to other measurement systems enabling food manufacturers to profile and predict food safety culture over time.

**H3:** Triangulation of methods, in contrast to a single method, is required to provide food manufacturers with a comprehensive evaluation of food safety culture including insights into the variables that drive food safety behaviours.
H4: Capturing social desirability in food safety and evaluating impact on food safety culture evaluations is required for food safety culture evaluation results to be valid.

1.6. Thesis structure

The format of the thesis is a ‘manuscript’ format. As such, it consists of five published or submitted peer-reviewed papers – one in each chapter (Figure 1). Each chapter has a unique introduction with literature review, material, results, and discussion sections. Publication status is captured at the beginning of each chapter. This overarching introduction and the concluding discussion chapter connects the different chapters to collectively seek answers to the thesis hypotheses.

![Figure 1: Thesis structure.](image)
2. Comparative analysis of existing food safety culture evaluation systems


2.1. Abstract

2.1.1. Purpose

The purpose of the research was first to, analyse 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 varying degrees in the food industry was conducted. The systems were found to vary considerably 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 increased commonality of approach to validation strategy and adoption of an overarching structural framework.

2.2. Keywords

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

2.3. 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.
• Results demonstrate that common validation techniques were 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 were proposed to unify the research field.
2.4. Introduction

The problem of food safety culture – what is it and how do you know how good yours measures up – is probably one of the main issues in modern thinking about food safety (Griffith, 2010; Jespersen et al., 2016; Jespersen & Huffman, 2014; Nyarugwe, Linnemann, Hofstede, Fogliano, & Luning, 2016; Powell et al., 2011; Yiannas, 2009). Although food safety culture is now subject to much discussion, the concept 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, 2013a), melamine poisoning 2007 – China (Gossner et al., 2009; Ingelfinger 2008), enterohemorrhagic *Escherichia coli* (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. (Pennington, 2009; 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, invest their resources based on attitudes and assumptions, and interpret 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. Often a breakdown of learned and shared attitudes, values, and beliefs cause deadly failures in food safety management systems. These failures often become embedded 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 prevent consumers from being exposed to foodborne hazards and for brand protection and employment security (Cameron, 2006; Jespersen & Huffman, 2014; Seward, 2012). Since 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 reliability.

Current systems for evaluating culture have been referred to as fragmented and built on disparate scientific theories (Boeck, Jacxsens, Bollaerts, & Vlerick, 2015; 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 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 (Guldenmund, 2000; Howe & Eisenhart, 1990).
The objectives of this research were to analyse existing culture evaluation systems for commonalities and differences in research quality, applied validation strategies, and content; and, to suggest a simple structure of food safety cultural dimensions to help unify the culture evaluation field.

2.4.1. 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; Shavelson & Towne, 2002).

2.4.2. 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 & 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. Krippendorff (2004) reflects on the trustworthiness of scientifically-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.”

2.4.3. 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 (Cohen, Manion, & Morrison, 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 while avoiding untrustworthy, invalid research (Cohen et al., 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 (Cohen et al., 2007). However, there is some disagreement with these statements as Robson (2011) states “unless a measure is reliable, it cannot be valid.” 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, 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 (Cohen et al., 2007; Long & Johnson, 2000; Meyrick, 2006; Robson, 2011).

### 2.4.3.1. 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 chapter 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 (2004) discuss this and highlight how, in some cases, outright mistakes are 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. 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.

2.5. Material and method

2.5.1. Sample.

The sample consisted of eight culture evaluations systems that had been used by companies in the food industry: five systems, referred to in this paper by the name of the authors who first described them i.e., Ball, Denison, De Boeck, Jespersen, and Wright, as well as three systems referred to in this paper by their commercial names, i.e., 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. Content from the commercial systems was included in the analysis, however, details regarding validation strategies either do not exist or are not publically available and therefore not included in the comparison of applied validation strategies. At the time of the research these eight evaluation systems were found through an online search and discussion with practitioners in the food industry.

2.5.2. Method.

A six-step process was used for the comparative analysis (Figure 2). 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. This information was 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 4) was developed based on initial
reading and descriptions of the eight systems. The framework consists of three levels of textual analysis: global, organizing, and basic themes (Attride-Stirling, 2001).

1. Identify culture evaluation systems in scope of the analysis.
2. Review publically available material related to research quality, system validation, and content.
3. Develop content analysis framework and textual data analysis in NVivo 11.
4. Summarize NRC compliance, validation strategy, and content for each system, review with experts.
5. Compare and contrast research quality and trustworthiness.
6. Compare and contrast content against global themes.

Figure 2: Six step process applied in the comparative analysis.
Table 4: Content coding framework.

<table>
<thead>
<tr>
<th>Global</th>
<th>Organizing</th>
<th>Basic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Values and Mission</td>
<td>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…”</td>
<td>Compliance</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Measures/metrics/KPIs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mission, vision, goals</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ownership/owning</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Plan/roadmap, direction</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Recall/recalls/withdrawals</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Responsibility, accountability, and commitment</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Direction</td>
</tr>
<tr>
<td>People Systems</td>
<td>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.</td>
<td>Any reference to person’s role/education/job and group or team (e.g., name of team, established teams, established groups)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Behaviour/practice, work routine</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Communications (e.g., written, spoken, and dialog (e.g., interview)) and involvement</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Consequence, escalation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pride</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rewards and celebration</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Training, education, learning, proficiency</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cross-functional</td>
</tr>
<tr>
<td>Consistency</td>
<td>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, performance reports, performance, decisions and use of technology.</td>
<td>Actioned data and performance metrics. Actions, tasks, action due date.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Non-conformance, reoccurring,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Technology</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tools and infrastructure (missing infrastructure, appropriate/right tools, appropriate infrastructure, missing tools)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>References to third party standards</td>
</tr>
<tr>
<td>Global</td>
<td>Organizing</td>
<td>Basic</td>
</tr>
<tr>
<td>---------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Adaptability</td>
<td>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.</td>
<td>Change readiness, open to change, change ready Improvement, must improve, continuous improvement, improvement process, improvement system, continuous improvement, Six Sigma, Lean manufacturing</td>
</tr>
<tr>
<td>Risk awareness</td>
<td>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.</td>
<td>Leaders risk awareness and perception Operator risk awareness and perception Risks, hazards.</td>
</tr>
</tbody>
</table>
The aggregation of basic and organizing themes was used to define the global themes. The global themes were later translated into suggested food safety culture dimensions. The degree of agreement or disagreement in content of each system was evaluated using the coverage of each system for each of the global themes. Coverage was evaluated by quantifying organizing codes per system per global theme and qualitatively by looking for latent meaning of detailed system content. A summary was written for each system, focusing on structure, compliance to the National Research Council (NRC) guidelines and validation strategies. By summarizing the global and organizing codes per system a comparative figure was developed to visually analyse for agreement or disagreement across the five suggested food safety cultural dimensions.

2.5.3. 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 was achieved through a literature review, expert consultation, analysis and inclusion of language found in content analysis and currently applied in the food safety domain. Experts were chosen based on their knowledge of culture and independent of the research and/or the models. The experts were trained in the NRC guidelines and reviewed samples of documentation against these.

(2) Internal face and construct validity to ensure the content reflects what it was intended to and that the evaluation construct was robust. Peer review was completed and representative quotes added to link findings from each evaluation system to the global themes (Graneheim, 2004).
(3) Internal validity to ensure replicability 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 was also evaluated through technical accuracy. Evaluation through discussions with practitioners and academicians to ensure that the construct and results were representative of the concept they intended to measure.

2.6. Results

2.6.1. 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.

2.6.1.1. Ball system.

This system was developed by Ball in 2009 as part of her doctorate work at the University of Guelph (Ball et al., 2009; Wilcock et al., 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 no reference to a definition of food safety culture in the papers and this is likely 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 self-reported food safety behaviours of food handlers. The system research meets 11 of the 12 NRC guidelines as it is not clear from publically available material how alternative explanations of the findings were explored.

### 2.6.1.2. De Boeck system.

This system was developed by De Boeck in 2015 as part of her doctorate work at the University of Gent (Boeck et al., 2015; De Boeck et al., 2016). 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 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.”

However, the authors themselves state that the concepts remain vague and with no unanimous definitions, therefore, for this analysis no differentiation is made between the terms climate and culture 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. As noted, the authors define culture and climate and it is not apparent how precisely these definitions are based on existing research (Griffith et al., 2010a; Guldenmund, 2007; Schein, 2004) in the domains of culture and climate. The research meets nine of the 12 NRC guidelines as it is not clear from publically available material how comprehensive the literature review was that led to the model, how alternative explanations of the findings are explored, and if there is a potential impact of systematic bias.

2.6.1.3. CEB system.

The system was developed by CEB (retrieved from cebglobal.com and personal conversations) 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 evaluation is made through 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 evaluation. Details regarding validation strategies for this model were not published.

2.6.1.4. Denison system.

This system was developed by Denison in 1989 and applied extensively with global organizations since as the Denison Model (Denison, Hooijberg, Lane, & Lief, 2012; Denison, 1997; Denison & Mishra, 1995). 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.

2.6.1.5. Jespersen system.

This system was developed by the researcher in 2014 at the University of Guelph (Jespersen et al., 2016; Jespersen & Huffman, 2014). The authors 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 tested with one global food manufacturing company in North America. The evaluation was conducted using triangulation between self-assessment surveys; behavioural observations and interviews; and performance assessments and made use of combined deductive and inductive content analysis and quantitative self-assessment data. The triangulation 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 and predictability the model needs to be tested by multiple organizations. The system research meets eight of the 12 NRC guidelines 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.

2.6.1.6. NSF model.

The system was developed by NSF (retrieved from nsf.org and personal conversations) 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 behavioural science (Fone, 2012). 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 and 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).

2.6.1.7. **TSI system.**

This system was developed by TSI in 2015 and applied to food service in Dubai and small food manufacturers in the U.K. (Taylor & Taylor, 2004; Taylor, 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 around HACCP application conducted in U.K. small and medium size companies and food service restaurants (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.

2.6.1.8. Wright system.

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 (C. J. Griffith, K. M. Livesey, & D. A. Clayton, 2010b) 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 behavioural 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.

2.6.2. Differences in validation strategies.

All eight tools were reviewed for validation strategies but information was only available for the scientifically-based tools. Each of the five 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 construct validation but only two show predictive validation. Reliability testing is shared in two culture evaluation systems (Ball and Denison) and not clear in the remaining (Table 5).

Table 5: Differences between the validation strategies applied in the eight culture evaluation models.

<table>
<thead>
<tr>
<th>Culture evaluation model</th>
<th>Validity methodology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ball</td>
<td>Internal face and construct validation through expert solicitation.</td>
</tr>
<tr>
<td></td>
<td>External population validation through focus groups and peer</td>
</tr>
<tr>
<td>Culture evaluation model</td>
<td>Validity methodology</td>
</tr>
<tr>
<td>--------------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td></td>
<td>review. Respondent validation through interview transcript review. Internal consistency through Cronbach’s Alpha. Predictive validation through behavioural model and triangulation.</td>
</tr>
<tr>
<td>De Boeck</td>
<td>External population validation through peer review. Internal, face, and construct validation through twenty experts, cross-sector, from Belgium.</td>
</tr>
<tr>
<td>Denison</td>
<td>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.</td>
</tr>
<tr>
<td>Jespersen</td>
<td>External population and ecological validation through review of existing food safety performance data and adoption language from existing food safety standards. Internal construct and face validity through an 18-member expert panel cross-sectional, from United States (US), U.K., and Canada, consisting of academicians and practitioners. Predictive validation through behavioural model and triangulation.</td>
</tr>
<tr>
<td>Wright</td>
<td>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.</td>
</tr>
</tbody>
</table>
2.6.3. System content comparison.

The findings from the content analysis provided data for comparison of the content in all eight culture evaluation systems. It should be noted that not all systems apply the food safety culture definition introduced earlier. The content was grouped to provide a graphical representation of dimension coverage by each culture evaluation system (Table 6).
Table 6: Content comparison of the eight culture evaluation systems.

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

Ledger: 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 dimensions are ‘Values and mission’, ‘People systems’, ‘Consistency’, ‘Adaptability’, and ‘Risk awareness’.

**2.6.3.1. Values and mission.**

This dimension covers cultural content related to (1) management and employee commitment to food safety (Ball, De Boeck, TSI), (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.

2.6.3.2. 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, TSI), (3) leaders’ and employees’ communication of food safety (De Boeck, Wright), (4) actual and expected involvement, autonomy, degree of membership input (Denison, TSI, 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 these drive food handler food safety behaviours.

2.6.3.3. 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.
2.6.3.4. **Adaptability.**

Dimension covers cultural content related to; (1) how the organization embraces or resists change (Denison, TSI), (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.

2.6.3.5. **Risk awareness.**

Dimension covers cultural content related to; 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.

2.6.4. **Suggested framework to unify the research field.**

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

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. The five scientific systems was evaluated for compliance to the National Research Council (NRC) guidelines (Shavelson & 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 principles. 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 this most comprehensively through discussion of focus group discoveries and missing evidence around impact on culture assessment through organizational levels.

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

Content was compared through content analysis of the textual data. It was found that almost all culture evaluation systems contain some content related to an organization’s values and mission. The Wright system dedicates four of five elements to this dimension. All eight culture evaluation systems cover content related to ‘People systems.’ Four of the eight cultural evaluation systems cover ‘Consistency’ and Jespersen was found to cover this in three of five capability areas. ‘Adaptability’ was covered by the models of Ball and Denison, and the Ball system dedicated two of six constructs to this dimension. ‘Risk awareness’ was only covered in detail by the De Boeck and Wright systems. The content findings suggest that the proposed five dimensions cover all of what each system independently cover and it suggests some areas that could be strengthened in some systems e.g., ‘Adaptability’, ‘Consistency,’ and ‘Risk awareness.’ By reviewing the detailed basic, organizing, and global themes from all systems a strong picture
of not simply organizational culture but food safety culture emerges. It is concluded that the five proposed dimensions could be used to unify research in the food safety culture domain and 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 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 CEB, NSF, and TSI culture evaluation systems. More documented details for each of these three 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 (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 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 (Griffith et al., 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 needed 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.
3. Development and validation of a scale to capture social desirability in food safety

Publication status: under production with Food Control.

3.1. Abstract

The evaluation of food safety culture in a food company is influenced by human factors such as the employees’ tendency to respond to social desirability – a reflection of respondents’ tendency to answer questions in a manner that will be viewed favourably by others. Building on previous research, a self-assessment scale consisting of 18-statements, the food safety desirable responding scale (FSDRS), was developed to capture desirable responding to food safety issues. Statistical analyses of data collected from 816 North-American food manufacturing professionals revealed a shortened 14-item version of the FSDRS scale which provides a reliable and valid measurement of the extent to which employees deceive themselves. It is concluded that the proposed FSDRS will enrich food safety culture measurement and food safety performance.

3.2. Keywords

food safety, social desirability, scale validity

3.3. Highlights

- Impact of the human factor – social desirability - a respondent’s tendency to answer questions in a manner that was viewed favourably by others – on food safety culture evaluations.
- An ecological social desirability scale was adapted to the food safety domain.
- Empirical study captured magnitude of social desirability in five North American food companies.
- Considered social desirability as a control factor in food safety research.
• Contextualized food safety by measurement of social desirability.
3.4. Introduction

The World Health Organization (WHO) estimates that 33 million healthy life years are lost annually due to food and drink related contamination ("Global burden of foodborne diseases," 2015). A significant portion of these contamination events originate from actions taken while processing food and drink products. Recall of processed food and drink products is a regulated practice institutionalized to minimize the loss of health and life by removing contaminated products from consumers’ reach. In the United States (U.S.) and Canada alone, 626 recalls were conducted in 2015 to remove products suspected of presenting a risk to the consumer (Maberry, 2016). Some reasons for actions resulting in the contamination of food and drink products are found in the organizational culture of food and drink processors (Griffith, 2010; Powell et al., 2011) where assumptions are made about the appropriateness of actions taken or decided ‘simply because it is the way we do things around here’ (Schein, 2004). As such, theories and findings from empirical studies regarding organizational culture at large (Cameron, 2006; Denison, 1997; Schein, 2004) and specifically within the context of the food industry e.g., Jespersen and Huffman (2014) can provide insight into some of the reasons for these actions leading to the loss of quality adjusted life years (QALYs).

3.4.1. Organizational culture and food safety failures leading to contamination.

Through a review of four recalls, between 2008 and 2013, the connection between failing organizational cultures and contamination actions was investigated by (Jespersen et al., 2016). These four recalls were selected for review, since food safety culture was specifically identified in the investigation reports as one of the factors causing the recall. Similar analyses were completed by (Powell et al., 2011) and (Wright, 2016) with similar findings of cultural failures.
In addition to the findings from these studies, the following provides a mapping of food safety failures to established dimensions of organizational culture.

The review included the following recalls; Fonterra (Incident, 2014), Jensen Farms (Centers for Disease & Prevention, 2012), Maple Leaf Foods (Canadian Food Inspection, 2013a), and XL Foods (Canadian Food Inspection, 2013b) and identified some common failures related to established dimensions of organizational culture (Schein, 2004). These were, ‘External adaptation’, ‘Internal integration’, ‘Reality and truth’ and ‘Human nature, activity, and relationship’. First, related to the two dimensions ‘External adaptation’ and ‘Internal integration,’ it was assumed that effective communication between stakeholders was taking place, e.g., between inspectors and employees in different functions and between different levels of the organization, whereas post event evidence showed this not to be the case. Evidence was also found of employees working within functions creating a silo effect where assumptions were made about actions taken outside their own functional silo. Secondly, related to ‘Reality and truth,’ it was assumed that somebody was deriving and analysing information from collected food safety data when this was not the case. This created a false sense of security around actual food safety performance. Thirdly, related to ‘Human nature, activity, and relationship’ it was assumed to be acceptable in all cases for employees to by-pass or not follow written procedures and this had become a common and accepted practice. These findings were remarkable given that in publically available material, such as information on the company websites, annual reports, and press releases, all four companies mentioned explicit written statements of commitment to and stressed the strategic importance of food safety for their organisation. This overall noted attitude-to-behaviour disparity or varying degree of ‘walking the food safety talk’ within the four companies has been attributed to weaknesses in their organizational culture.
before the recalls, which impacted negatively on their food safety performance (Jespersen, Griffiths, & Wallace, 2017).

Other empirical research and publications in the popular press have also connected food safety culture to food safety performance. Impact of food safety culture attributes (e.g., food safety perceived value; food safety climate, level of the implemented food safety management system) have been suggested to impact food safety performance (Ball et al., 2009; Boeck et al., 2015; Griffith et al., 2010a; Jespersen et al., 2016; Powell et al., 2011; Taylor, 2015; Wright; Yiannas, 2009). Common to all these studies is the claim that, if measured, these attributes could provide companies with a valid indication of the strength of their food safety culture.

Next to the sometimes observed disparity between organizational attitudes and real organizational food safety decision-making and behaviour, incongruent perceptions regarding food safety culture might exist within an organization. For instance, De Boeck et al. (2015) found that managers of affiliates in a large scale meat company in Belgium perceived some attributes of the food safety culture in their organization differently than their subordinates. They called for further research to demonstrate to what extent self-assessments (e.g., by an individual employee) versus assessment by others (e.g., by auditors, inspectors, colleagues) of food safety climate correlate or are shared within a work unit/organization on the one hand and how disagreement in food safety culture perceptions between organizational stakeholders might be explained (e.g., social desirability).

Studying the attitude-to-behaviour disparity at the organizational level and unravelling deviating human perceptions in the context of food safety is important as both might impact the validity of an evaluation system, bias the measurement of food safety performance and provide companies visibility to potential cultural blind spots. (Kaiser, 1998) describes some key features
of ecological behaviours and points out that ecological behaviours are impacted by sociocultural and personal factors causing ecological behaviours to appear inconsistent. Features of social desirability, as researched in ecology and hypothesized in food safety, can therefore be difficult to assess. A valid scale for evaluating social desirability can help enhance understanding of the impact of these features on food safety culture. As a recent comparative analysis of systems and tools assessing food safety culture and its attributes revealed that none of the existing evaluation systems assessed the potential impact of social desirability in the food safety domain (Jespersen et al., 2017).

### 3.4.2. Social desirability.

A social desirability bias is a social science research issue that describes the tendency of survey respondents to answer questions in such a certain way to be viewed favourably by others. It can take the form of over-reporting good behaviour or under-reporting undesirable behaviour. The concept of social desirability has developed significantly over the past 60 years. Seminal work has been done by (Crowne, 1960; Edwards, 1953; Paulhus, 1984; Paulhus, 1991). For instance, (Crowne, 1960) increased the social desirability discussions with a review of scales like the Minnesota Multiphasic Personality Inventory (MMPI) \( K \) scale and its inability to assess ‘social desirability’ responding since many of its items have pathological implications, such as, ‘worry quite a bit over possible misfortunes.’ The MMPI \( K \) scale was designed to measure the absence or presence of such signs not if the response is attributable to social desirability. (Crowne, 1960) developed another social desirability scale, the M-C SDS, defined by behaviours which were culturally sanctioned and approved but improbable to occur. Said differently, social desirability was defined more broadly to refer to the need of subjects to obtain approval by responding in a culturally appropriate and acceptable manner to statements, such as, “I never
intensely dislike anyone.” The authors found that Edwards’ SDS scale correlated higher with MMPI $K$ scales than M-C SDS and the authors questioned if the high degree of correlation did not in fact mean that the Edwards SDS and MMPI $K$ scales were functionally unitary. The M-C SDS scale was validated through the fact that the most pathological of the clinical scales correlated negatively with the M-C SDS scale. The M-C SDS scale consists of 33 items in the Personal Reaction Inventory and measures two concepts, (a) very socially desirable but untrue for most people, and (b) very socially undesirable but very common. The Balanced Inventory of Desirable Responding (BIDR), developed by (Paulhus, 1991), also measures two constructs, self-deceptive positivity i.e., the tendency to give self-reports that are honest but positively biased, and impression management. Impression management is a term applied when respondents to surveys or interview questions deliberately alter their answer to convey or portray a chosen image (Ewert, 2009). The 40 BIDR items are stated as affirmative propositions and respondents rate their agreement on a seven-point scale. Compared to other methods such as the Responding Desirably on Attitudes and Opinions (RD-16) by (Schuessler, Hittle, & Cardascia, 1978) the BIDR was found to be the best validated and most applied of the methods reviewed (Robinson, 2013). (Paulhus, 1991) investigated the difference between enhancement and denial compared to self-deception and impression management. Three studies were conducted, first a study of the importance of the keying direction, a second to replicate the results from the first, and the third on the influence of valence characteristics, positive and negative was studied. The studies showed good correlation between the M-C SDS scale and self-deception and impression management. Denial sub-factors did not make a difference to the results and replicated the results from the 1988 study. Also, self-deception correlated with factors in standard personality.
instruments such as, extraversion, perspective taking, and empathic concern. Impression management also correlated with perspective taking.

Next Ewert (2009) suggested a broad array of potential moderators or mediators of the attitude-to-behaviour disparity or the human’s tendency to social desirability responding in an environmental context such as belief systems, personal values, empathy, environmental identity, obstacles in practicing environmental behaviours, levels of connectivity with the environment, the method used to measure behaviour, thinking about environment as a function of culture and tradition, and methods of coping with environmental problems. The authors developed a domain specific ecological social desirability scale (EDRS). It is upon this validated and documented research history that our research on social desirability in the food safety domain stands.

3.4.3. Research objectives.

As environmental and organizational risks are both partly experienced as more distal then proximal by employees, the researchers assume that an individual’s perception of risk in the ecological domain has similarities with an individual’s risk perception in the food safety domain. The purpose of this research was to adapt the EDRS scale and validate it explicitly for the food safety domain. This adapted scale aims to capture the degree of social desirability responding by humans when evaluating food safety in food manufacturing organizations and results might be interpreted in the context of food safety culture.

The EDRS scale for measuring social desirability in the ecological domain was modified with specific content from the food safety domain and administered to 1,273 participants from five North American food manufacturing companies and 21 manufacturing plants between October 2015 and March 2016.
3.5. Material and method

3.5.1. The food safety desirability response scale (FSDRS).

The food safety desirability response scale (FSDRS) was developed by modifying the EDRS scale of Ewert and Galloway (Ewert, 2009). The EDRS consists of 18 statements all framed within the ecological context. The EDRS-items are grouped into three subscales, ‘Self-deception - assertion of positives’ nine items, ‘Image management’ five items, and ‘Self-deception - Denial of negatives’ four items (Table 7).

In line with Ewert (2009), and after interdisciplinary discussions among the authors and with seven experts with expertise in food safety and seven people with expertise in survey methodology, each of the ecological references in the 18 original EDRS-items were replaced, by consensus, with specific food safety references (Table 7). All statements in the FSDRS were rated by participants on a five-point Likert scale, (1) ‘Not at all like me’, (2) ‘Not like me’, (3) ‘Neutral’, ‘Like me’ (4), and ‘Just like me’ (5).

Table 7: The food safety desirability response statements (FSDR-scale).

<table>
<thead>
<tr>
<th>Self-deception – Assertion of positives (SD1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. My behaviour is consistent with my beliefs about food safety issues.</td>
</tr>
<tr>
<td>2. I know what actions I should take regarding how best to protect food safety.</td>
</tr>
<tr>
<td>3. I am always honest with myself about how I really feel about food safety.</td>
</tr>
<tr>
<td>4. I do not regret my decisions about food safety issues.</td>
</tr>
<tr>
<td>5. I have very definite views about what government policy should be regarding food safety.</td>
</tr>
<tr>
<td>6. I do not know the reasons why I feel the way I do about food safety. *</td>
</tr>
<tr>
<td>7. I appreciate other people’s opinions regarding food safety.</td>
</tr>
<tr>
<td>8. I try to understand other people’s views about food safety, particularly when they differ from my own.</td>
</tr>
</tbody>
</table>
9. I am not concerned about food safety issues. *

Image management (IM)

1. I never say bad things about people who disagree with my views about food safety.
2. I never say anything to hurt the feelings of someone who disagrees with me about a food safety issue.
3. I never get upset when people express opinions about food safety which differ from my own.
4. I am not interested in trying to influence people’s thinking about food safety.
5. I will not disagree about food safety issues with new people I meet.

Self-deception – Denial of negatives (SD2)

1. I try to cover up mistakes I make in conversations about food safety issues. *
2. I feel resentful when I don’t get my own way in a discussion about food safety issues. *
3. It bothers me if people dislike me because of my views about food safety. *
4. I form opinions about food safety issues without always thinking about the issues thoroughly. *

NOTE: Items with * are reversed scored.

3.5.2. Ethical considerations.

This study was approved by the Research Ethics Board at University of Guelph and the Operations Leaders and Food Safety Leaders in the five participating companies. A covering letter was attached to the electronic invitation to participate and informed all leaders of the scientific goal of the study and stated that filling in the survey implied consent. Participants were informed that participation was not mandatory and that confidentiality was guaranteed.

3.5.3. Data collection and sample.

To test reliability and validity of the proposed FSDRS scale (Table 7), data were collected through an online survey among all leaders, broadly defined as a person with responsibility for managing the work activities of others (e.g., plant manager, quality manager,
maintenance supervisor, and sanitation lead) of five food manufacturing companies and 21 manufacturing sites in U.S. and Canada. Social desirability questions were part of a broader study to measure and predict maturity of food safety culture that five companies had volunteered to participate in. As (Kaiser, 1998) found that demographical variables such as age impacted degree of social desirability responding, for example younger people tend to respond lower than older people on desirability scales, some demographic data (i.e., company, job, age, and tenure) were measured as well.

In total, 1,273 leaders were invited to participate in the study, 816 returned a usable questionnaire, yielding a response rate of 64.10%. The sample was comprised of 398 respondents in manufacturing, 125 in food safety and quality, and 293 in other functions. Managerial levels represented were 42 respondents in executive positions, 271 in middle-management positions, and 503 in frontline positions. The median age was 34-44 years ($SD = 1.134$), median seniority in the food industry was 15-19 years ($SD = 1.597$), median seniority in current company was 10-14 years ($SD = 1.674$), and median seniority in current job was 5-9 years ($SD = 1.561$).

3.5.4. Analysis.

Principal component analysis (PCA) was used to examine the structure of the FSDRS. Confirmatory factor analysis was not utilized at this stage of development of the FSDRS given it is still unclear whether its structure will reflect the factor structure of the EDRS. Factor analysis was completed after evaluating criteria for good factor analysis and factor extraction. Distribution was tested to ensure normality, followed by a oblimin rotation to investigate factor loading.

Reliability of FSDRS was evaluated through calculation of Cronbach’s Alpha for internal consistency. Pearson product-moment correlations and ANOVA were used to analyse construct
and discriminant validity. All statistical analyses of the data were conducted in SPSS version 23 (IBM Corporation, New York, U.S.A.).

3.6. Results

3.6.1. Factorial validity and reliability.

The 18 items from the FSDRS were subjected to principal component analysis (PCA). Prior to performing PCA, the suitability of data for factor analysis was assessed using criteria identified by Tabachnick and Fidell (Tabachnick, 2012). This included Kaiser-Meyer-Olkin (KMO) value greater than 0.60, this was exceeded by a value of 0.853 in this study (Kaiser, 1998) and Bartlett’s Test of Sphericity (Bartlett, 1954) reached statistical significance (p < 0.05).

Initially, the PCA revealed the presence of five components with Eigen values exceeding 1. An inspection of the scree plot revealed a break after the second component. To better select the most accurate solution a Parallel Analysis was completed using Monte Carlo simulation (Watkins, 2006). Only components with Eigen values from the SPSS analysis larger than the Monte Carlo Eigen values were accepted (Table 8).

Table 8: Output from parallel analysis.

<table>
<thead>
<tr>
<th>Component</th>
<th>Actual eigenvalue from PCA</th>
<th>Actual eigenvalue from parallel analysis</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.309</td>
<td>1.2693</td>
<td>Accept</td>
</tr>
<tr>
<td>2</td>
<td>2.420</td>
<td>1.2136</td>
<td>Accept</td>
</tr>
<tr>
<td>3</td>
<td>1.744</td>
<td>1.1770</td>
<td>Accept</td>
</tr>
<tr>
<td>4</td>
<td>1.362</td>
<td>1.1449</td>
<td>Accept</td>
</tr>
</tbody>
</table>
The oblimin rotated PCA of the 18 FSDRS-items resulted in three factors with acceptable Cronbach’s Alpha values (> 0.7) and a fourth factor with four items with lower than acceptable internal consistency value. The oblimin rotation was redone using the three factor solution (Table 9) and this simple structure of three factors was found suitable for use. Each of the 14 retained FSDRS-items exceeded the chosen factor loading cut-off value of 0.30 to decide which loadings were significantly associated with a given factor. Moreover, none of the 14 retained FSDRS-items had a meaningful cross-loading and each of the 14 items were loading on their expected factor. Finally, each factor was interpretable and comprised items in line with social desirability theory and earlier research (Ewert, 2009).

As items 6 and 9 from the original SD1-scale (both mentioned in Table 7) did not load substantially on their expected factor and instead were cross-loading on the SD2-factor, these two items were removed from the SD1-scale. The researchers speculate that the similar item wording ‘I am not...’ in both items elicited confusion in respondents. Indeed, both items had to be reverse-scored and were the only two items in the initial SD1-scale which are formulated negatively, by the absence of a behaviour. Also, items 4 and 5 from the original IM-scale (both mentioned in Table 7) did not load substantially on their expected factor. Moreover, both had a bi-modal distribution. Consequently, both were also removed from the original IM-scale. Remarkably, also these two items were the only items in the initial IM-scale which contained the negative self-relevant word ‘...not’, respectively ‘I am not....’ and ‘I will not...’ which might have caused confusion in respondents. In conclusion, 14 of the original 18 FSDRS-statements were retained (see Table 7 for an overview).
The first factor associated with seven items, was labelled ‘Self-deception – assertion of positives’ (SD1), and accounts for 33.42% of the variance. The second factor was labelled ‘Image management’ (IM), which is in line with Ewert and Galloway (2009) and was chosen over the more commonly used term ‘Impression Management’ because ‘Image’ seems to be more closely tied to how an individual intrinsically views him- or herself. Three items loaded on that factor and accounted for 16.52% of the variance. The third factor, which was labelled ‘Self-deception – denial of negatives’ (SD2), is characterized by four items and accounts for 11.89% of the variance. Collectively the three factors explained 61.83% of the variation.

Table 9 also contains the Cronbach’s Alpha estimates of reliability for the items loading significantly on each factor. The Alpha values include the following, SD1 (Factor 1) = 0.863, IM (Factor 2) = 0.837 and SD2 (Factor 3) = 0.767. It is widely assumed that alpha values greater than 0.7 are required for acceptable reliability (Nunnally, 1967).
Table 9: Results of the oblimin rotated principal factor analysis of FSDRS-items.

<table>
<thead>
<tr>
<th>FSDRS-items</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Self-Deception</td>
<td>Image</td>
<td>Self-Deception</td>
</tr>
<tr>
<td></td>
<td>Positives</td>
<td>management</td>
<td>Negatives</td>
</tr>
<tr>
<td>SD1_3  Always honest with myself about how I really feel about food safety</td>
<td>.856</td>
<td>.049</td>
<td>.347</td>
</tr>
<tr>
<td>SD1_1 Behaviour is consistent with my beliefs about food safety</td>
<td>.824</td>
<td>.084</td>
<td>.321</td>
</tr>
<tr>
<td>SD1_2  Know what actions to take regarding how best to protect food safety</td>
<td>.811</td>
<td>.015</td>
<td>.316</td>
</tr>
<tr>
<td>SD1_4  Do not regret my decisions about food safety issues</td>
<td>.757</td>
<td>.113</td>
<td>.335</td>
</tr>
<tr>
<td>SD1_8  Try to understand other people’s opinions about food safety,</td>
<td>.704</td>
<td>.130</td>
<td>.302</td>
</tr>
<tr>
<td>particular when they differ from my own</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD1_7  Appreciate other people’s opinions regarding food safety</td>
<td>.702</td>
<td>.081</td>
<td>.281</td>
</tr>
<tr>
<td>SD1_5  Have very definite views about what government policy should be</td>
<td>.577</td>
<td>-.013</td>
<td>.137</td>
</tr>
<tr>
<td>regarding food safety</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IM_1   Never say bad things about people who disagree with my views on</td>
<td>.057</td>
<td>.895</td>
<td>.001</td>
</tr>
<tr>
<td>food safety</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IM_2   Never say anything to hurt the feelings of someone who disagrees</td>
<td>.046</td>
<td>.869</td>
<td>-.044</td>
</tr>
<tr>
<td>with me about a food safety issue</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IM_3   Never get upset when people express opinions about food safety</td>
<td>.097</td>
<td>.842</td>
<td>-.016</td>
</tr>
<tr>
<td>which differ from my own</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD2_2  Feel resentful when I don’t get my own way in a discussion about</td>
<td>.317</td>
<td>.019</td>
<td>.823</td>
</tr>
<tr>
<td>food safety issues</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD2_1  Try to cover up mistakes I make in conversations about food safety</td>
<td>.361</td>
<td>-.020</td>
<td>.763</td>
</tr>
<tr>
<td>SD2_3  Bothers me if people dislike me because of my views about food</td>
<td>.203</td>
<td>-.054</td>
<td>.762</td>
</tr>
<tr>
<td>safety</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD2_4  Form opinions about food safety issues without always thinking</td>
<td>.315</td>
<td>.002</td>
<td>.731</td>
</tr>
<tr>
<td>about issues thoroughly</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cronbach’s Alpha</td>
<td>.863</td>
<td>.837</td>
<td>.767</td>
</tr>
<tr>
<td>Eigenvalues</td>
<td>4.679</td>
<td>2.312</td>
<td>1.665</td>
</tr>
<tr>
<td>% variance</td>
<td>33.42%</td>
<td>16.52%</td>
<td>11.89%</td>
</tr>
<tr>
<td>Total variance accounted for by the rotated factors = 61.83%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The results in Table 9 show how applying the modified 18 item EDRS scale to the food safety domain resulted in a shortened 14 item FSDRS scale with higher internal reliability, using Cronbach’s Alpha, compared to the original EDRS scale developed for the ecological domain. Moreover, the three factorial structure measured by the shortened 14 item FSDRS scale explained in total a higher degree of variance (61.83%) compared to 33.20% found by (Ewert, 2009).

### 3.6.2. Construct validity.

Construct validity is the degree to which a test measures what it claims to be measuring. The statements in the EDSR were adjusted to include reference to food safety and the FSDSR was reviewed by food safety experts (n=7) through conversations to ensure relevance specific to the food safety domain. Variation caused by the data collection method could be a threat to construct validity and warrants attention. The data collection technique used in the present investigation, self-assessment from the individual participant, is a potential source of such variance (Podsakoff, 2003). However, as the interrelations between the three FDSDR-subcales are relatively weak (< 0.4) (Table 10), we can distinguish the three expected constructs (SD1, SD2, and IM) by means of the FSDSR.

<table>
<thead>
<tr>
<th>Component</th>
<th>SD1</th>
<th>SD2</th>
<th>IM</th>
</tr>
</thead>
<tbody>
<tr>
<td>SD1</td>
<td>1.000</td>
<td>.080</td>
<td>.374</td>
</tr>
<tr>
<td>SD2</td>
<td>.080</td>
<td>1.000</td>
<td>-.012</td>
</tr>
<tr>
<td>IM</td>
<td>.374</td>
<td>-.012</td>
<td>1.000</td>
</tr>
</tbody>
</table>

**NOTE**: SD1 = Self-deception – assertion of positives, SD2 = Self-deception – denial of negatives, and IM = Image management.
3.6.3. Discriminant validity.

Discriminant validity refers to a scale’s ability to distinguish between two or more groups that it should be able to distinguish between (Cohen et al., 2007; Tabachnick, 2012). One-way analysis of variance was conducted to explore the association between company and the degree of social desirability responding as measured through the FSDRS scale.

In line with Paulhus and Douglas (1991), a total score (maximum 14) for each respondent was calculated based on a dichotomous scoring of the 14 retained items. Per se, responds to any of the 14 statements of 1 to 4 was coded ‘0’ and responds of 5 coded ‘1.’. This scoring procedure provides some assurance that respondent style and not scale content is being tapped. This scoring also ensures that high scores are attained only by subjects who give exaggeratedly social desirable responses. All 14 items were added to give an overall measure of FSDSR per participant. The data were found to be normally distributed by use of the Kolmogorov-Smirnov statistic (p = 0.000). The overall mean total social desirability score (N=813) was 4.94 (SD = 4.23). A one-way analysis of variance was conducted to explore social desirability score differences among the five participating companies and revealed a statistically significant difference at the p < 0.05 level in the social desirability score for the five companies: F (4, 460) = 10.079, p < 0.000. The effect size is large and supports the difference between companies (Eta squared = 0.0805). Same analysis was completed among the participating 21 manufacturing plants and it also revealed a statistically significant difference at p < 0.05 in social desirability score for the 21 plants: F (19, 435) = 3.166, p < 0.000. The effect size is large and reflects statistically meaningful differences between plants (Eta squared = 0.121).
3.7. Discussion

The purpose of this research was to evaluate degree of social desirable responding and interpret these results in the context of food safety culture. This was done by investigating the applicability of an established ecological scale for measuring social desirability to the domain of food safety.

The results showed that a validated scale for quantifying social desirability in the ecological domain when modified to fit the food safety domain can provide a valid measure of social desirability in food safety or, said differently, capture the degree to which employees ‘walk the food safety talk.’ The FSDSR scale was found to be both reliable and valid using data collected from five food manufacturing companies in North America. Through factor extraction it was shown that 61.83% of the variation in the social desirability score across the participants could be explained through three social desirability components; Self-deception – assertion of positives (SD1), Self-deception – denial of negatives (SD2), and Image management (IM). The proposed FSDSR scale consists of 14 items and three subscales and respondents were asked to report their extent of agreement with each statement on a five-point answer scale. By dichotomous scoring a total score (maximum score 14) and three subscale scores can be calculated such that higher scores on the SD1, SD2 and IM scales reflects higher social desirability responding when dealing with food safety issues through respectively high self-deception (asserting positive behaviours and denial of negative behaviours) and more management of their self-image. These three indicators of social desirability are in line with social desirability theory and previous empirical research on social desirability in other domains.

The data and analysis showed that social desirability can be captured and evaluated in the food industry by the proposed 14 statements of the FSDSR-scale. It is important to note,
however, that the companies who participated in the study are all based in North America. Given this geographical focus, further research might investigate the potential impact of national cultures on social desirability in food safety. The modified scale – FSDSR – was found valid and reliable based on data collected from leaders at all levels in the companies. An opportunity exists to apply the scale among other stakeholders and groups e.g., associates/operatives, inspectors to compare social desirability between groups. The researchers also acknowledge the need for predictive validity and recommend further psychometrical research to determine this.

In conclusion, as food companies recognize more and more the strategic importance of their food safety culture, its reliable and valid evaluation gains importance. As this research, might be biased or influenced by human factors such as employees’ food safety motivation and their willingness to participate and fill in self-assessment tools (Boeck et al., 2015; Clayton et al., 2002; Guldenmund, 2007), we propose that the employee’s tendency to respond in a social desirability way regarding food safety issues should be captured as well. The FSDSR allows quantification of this tendency by estimating the extent to which employees deceive themselves by asserting positive behaviours and denying negative behaviours on the one hand and manage their image on the other hand. The current research is significant as a new validated scale can now be applied. Further food safety research might benefit from considering social desirability, statistically speaking, as a control variable. By measuring human tendency to answer food safety related questions in a manner that will be viewed favourably by others, the food industry might get a more authentic, valid, and contextualized picture of the food safety culture in their organizations. This input can facilitate the designing of interventions aimed at changing the food safety culture in an organization and the food safety behaviour of employees. In other words, by
taking social desirability into account food safety management and control activities might be strengthened.
4. Predictive attributes of food safety culture and climate

Publication status: under review with Food Research International.

4.1. Abstract

Impact of organizational culture and macro-cultures such as food safety continues to be of great importance to maximize organizational effectiveness and minimize food safety risks. The objective of this research was to propose a scale for evaluating food safety culture and demonstrate its predictive validity. The scale was found to explain up to 72.0% of variation with good internal consistency. A predictive model was derived with the ‘Values and mission’, ‘People systems’, and ‘tenure in company’ explaining 82.5% of the variation in the model. Based on research led by three other scientists as well as findings of this work, a collaborative model of cultural attributes is proposed. Common for all attributes is that they are derived using predictive validation methods and therefore seen as going beyond the simple content and face validity stages.

4.2. Keywords

Food safety culture predictability and performance.

4.3. Highlights

- Food safety culture evaluation
- Predictive cultural attributes
- Collaborative model
4.4. Introduction

The impact of organizational culture on organizational performance and effectiveness has been discussed for 40+ years (Denison, 1997; Denison & Mishra, 1995; Hofstede, Hofstede, & Minkow, 2010; Kotter, 1992; Schein, 2004; Schein & Schein, 2017). (Denison & Mishra, 1995) sought to answer the question “what can the cultural characteristics of an organization tell us about effectiveness?” and demonstrated the connection between four organizational traits: Involvement, Consistency, Adaptability, and Mission and organizational effectiveness. Denison measured organizational effectiveness through behavioural performance using an established scale Survey of Organizations (SOO) and financial performance through income/sales ratio and income/investment ratio. He found a valid connection of these cultural traits to both behavioural performance and financial effectiveness. (Kotter, 1992) conducted research to answer a slightly different question “what kind of corporate cultures enhance long-term economic performance?” Like the work completed by Denison, the researchers connected culture strength, fit, and adaptability to financial performance through net income growth, return on invested capital, annual growth of stock price, and market value growth. They demonstrated that both cultural fit and organizational adaptability impact organizational performance not unlike the findings around Mission and Adaptability traits in Denison’s work. An additional point of importance is brought to the discussion by (Brown, 1998) who demonstrated a connection between organizational culture and business context. That business context exerts influence on what organizational culture characteristics are functional or not. Organizational performance and effectiveness measured through the collection and analysis of economic measures e.g., (Brown, 1998), (Denison, 1997), and (Kotter, 1992), all demonstrated a quantifiable connection between culture
and organizational performance. However, there are still some challenges replicating and proving causality between organizational culture and economic performance.

To establish causality between organizational culture and performance it has been anecdotally stated that “an organization’s culture is set from the top” (Hermalin, 2013) and permeates through functions and groups, each introducing their own sub–or macro-cultures (Schein, 2004; Schein & Schein, 2017). Across the food supply chain (i.e., primary producers, processors, distributors, retailers, and caterers), there are countless companies and organizational cultures (Jespersen, 2017) and within each a food safety culture. Empirical research in the domain of people-safety macro-culture show similar structures and researchers have argued that this has led to fragmentation and confusion about what people safety culture is about (Guldenmund, 2000). As a less mature domain, food safety culture has an opportunity to learn from organizational culture and the people safety domain and move towards more knowledge and action founded in science and predictive validity not simply face and content validity. Few studies of food safety culture go to the extent of confirmatory studies or predictive validity but three researchers have demonstrated that select socio-psychological factors can predict food safety performance.

Hinsz lead or participated in five research studies between 2004 – 2015 (Betts, 2014; Hinsz & Nickell, 2015; Hinsz, Nickell, & Park, 2007; Hinsz & S., 2004; Nickell & Hinsz, 2011). All studies were focused on exploring predictive validity of select socio-psychological factors and their impact on the individual food worker’s intent and self-reported food safety behaviours. The theoretical frameworks for the research were the The Reasoned Action Approach (Ajzen, 2011) and the Theory of Planned Behaviour (Fishbein & Ajzen, 2009). Path analyses were completed based on an integrative framework for motivated behaviour (Betts, 2014; Hinsz &
Nickell, 2015; Hinsz et al., 2007; Hinsz & S., 2004; Nickell & Hinsz, 2011). Regression analyses were conducted in all these studies to test the ability of social norms, perceived control, work habits, and work routines to predict self-reported food safety behaviours. The authors of the four studies found that behavioural intention (i.e., social norms, attitude, and perceived control) ($r^2=0.57$), work habits ($r^2=0.56$), and work routines ($r^2=0.59$) significantly contribute to the prediction of food safety behaviours but with work routines significantly mediating the variation captured through work habits. Social desirability was assessed through the Balanced Inventory of Desirable Responding (Paulhus, 1988) and it was found, by controlling for this variable, that results were not significantly impacted by participant social desirable responding. (Nickell & Hinsz, 2011) investigated the predictive ability of food safety climate and conscientiousness on food safety behaviours and found a significant relationship ($r^2=0.45$). Further research by (Betts, 2014) demonstrated predictability of behavioural intent and self-reported behaviours by mindfulness and food safety knowledge ($r^2$ not reported but model was significant ($p < 0.01$)). As with all scientific studies there are limitations and some of these were highlighted by the authors. The study made use of self-reports only and not observational data or other methods. Also, data were collected from employees at a single manufacturing site and from students employed in food service. Thus, generalizability across countries, industries, companies, and product categories must be investigated. These studies were the first of their kind and ground breaking for food safety culture and bring much needed rigor and depth to the domain.

Ball lead the completion of another study and subsequent path analysis of factors influencing worker food safety behaviours (Ball et al., 2009; Wilcock et al., 2011). By collecting data from both large and small food companies through surveys, interviews, and focus groups, the authors investigated factors that affected worker food safety behaviours. The research led to
the development of a model describing the factors influencing implementation of food safety systems. The model was a further development of the Van der Wende (2006) model and the path analysis applied principles from the Model for Planned Behaviour (Ajzen, 2011). The researchers demonstrate in the path analysis the predictive ability \( r^2 = 0.65 \) of the model especially related to impact of the work unit commitment to worker food safety behaviour. The model also showed indirect effects from management commitment, infrastructure to support food safety, and food safety training. Again, there are limitations, some of which the authors discuss. Samples were drawn from meat processing facilities in Southern Ontario, Canada only. This could lead to a problem of generalizability for other sectors of the food supply chain or other countries. Also, none of the plants in the study were unionized which could impact the factors found as significant. Having said that, the study is built on proven and applied principles from social psychology and work completed around food safety management systems and considered a significant contribution towards the food safety culture domain; specifically, around the influence of the work group.

De Boeck led the investigation of mediators’ and moderators’ impact on an individual’s food safety climate assessment (De Boeck, Mortier, Jacxsens, Dequidt, & Vlerick, 2017). The proposed model was built on survey data collected at two small Belgian produce companies and was an extension of a model built on data from small meat shops and processing plants in Belgium (De Boeck et al., 2016) with input from work completed by (Luning et al., 2015) on the technology and managerial route. By applying a multi-regression procedure, the authors show a relationship between food safety climate and food safety behaviours \( r^2 = 0.70 \), specifically behaviours related to compliance and participation of the individual employee in food safety. Again, there are a few limitations in the study, as discussed by the authors. The analysis was
completed using principles and empirical findings from the people safety domain. This can be a weakness as there are differences in employee’s risk perception between safety and food safety (Jespersen, T., & Vlerick, In press). Further, it is postulated that food safety climate is made up of leadership, communication, commitment, risk awareness, and resource and this is different from the climate definition used by other authors (Denison & Mishra, 1995; Guldenmund, 2000; Nickell & Hinsz, 2011). Although studied in a limited context, it is not clear that there is consensus of this climate definition in the food safety domain or in the study of organizational culture. Overall, the work completed adds great content and much needed rigour to the food safety culture domain specifically related to the individual employee and food safety climate as defined by the authors.

The research summarized is ground breaking and reduces the risk for the food safety domain to become as fragmented and contradictory as the people safety domain. A few general gaps are associated with all these studies and the authors call for action to address these in future research. There is a lack of global comparative studies and it is important to note that the research studies reviewed here are conducted in a narrow geographical area (i.e., U.S., Canada, and Belgium) giving detailed insights into macro-cultures but lacking the broader geographical perspective. This could bring the generalizability of the results into question. Also, the studies mostly focus on frontline employees and less on management staff. Further, the studies also assume homogeneity across all employee groups and functional areas in the cultures studied and did not explore differences between potential macro-climates or cultures. Awareness of these gaps are important to avoid repeating the fragmentation and confusion found in the organizational culture domain (Schein & Schein, 2017) and people safety domain (Guldenmund, 2000).
The objective of this research was to evaluate a scale for assessing food safety culture and propose a combined and collective model for food safety climate and culture based on existing research with predictive validity. Answers to the following five research questions (RQs) are sought to meet this objective,

- **RQ1:** Is the proposed scale reliable and valid to evaluate food safety culture?
- **RQ2:** How well does each dimension predict overall rating of maturity? Also, how much variation in the overall maturity assessment can be explained by the maturity rating of Values and mission?
- **RQ3:** If we control for the possible effect of tenure, age, and social desirability, can dimensions still predict a significant amount of variation in food safety maturity rating?
- **RQ4:** How well do respondents span of control predict overall rating of maturity?
- **RQ5:** How well does functional responsibility rating of maturity predict overall rating of maturity? How much variation in food safety rating can be explained by manufacturing rating?

### 4.5. Material and method

#### 4.5.1. Participants.

To test reliability and validity of the proposed scale and its predictive validity, data were collected through an online survey distributed to leaders, broadly defined as persons with responsibility for managing the work activities of others (e.g., plant manager, quality manager, maintenance supervisor, and sanitation lead) in five multinational food manufacturing companies and a sub-set of 21 manufacturing sites. As (Kaiser, 1998) found that demographical variables, such as age, impacted degree of social desirability responding, specifically that younger people tends to respond lower than older people on desirability scales, some demographic data (i.e., company, job, age, and tenure) were collected. In total, 1,273 leaders were invited to participate in the study, 816 returned a usable questionnaire, yielding a response rate of 64.10%. The sample was comprised of 398 respondents in manufacturing, 125 in food
safety and quality, and 293 in other functions. Managerial levels represented were 42 respondents in executive positions, 271 in middle-management positions, and 503 in frontline positions. The median age was 34-44 years ($SD = 1.134$), median seniority in the food industry was 15-19 years ($SD = 1.597$), median seniority in current company was 10-14 years ($SD = 1.674$), and median seniority in current job was 5-9 years ($SD = 1.561$).

4.5.2. Scale.

The scale was developed on the basis of the Reasoned Action Model for planned behaviour (Fishbein & Ajzen, 2009). The specific behaviours in the maturity model were defined at two extreme stages of maturity; ‘Doubt’ and ‘Internalized’. By defining pinpointed behaviours at the extreme endpoints of the maturity scale it was possible to create a self-assessment survey with fewer questions and, by use of a five-point Likert scale, measure across the entire maturity model. The objective of the questionnaire was to gather participant’s self-assessment results against the specific behaviours and collect demographic data pertaining to plant, function group, and work role. Each participant was asked to rate their own behaviour against a series of questions and statements. The scale consisted of 126 individual statements divided into eight sub-scales. Each sub-scale was defined around specific role and functional behaviours e.g., leader in food safety and quality (FSQ) or Manufacturing Supervisor (Jespersen et al., 2016). This dichotomous scale was designed to increase the likelihood that the participant’s perspective to the topic rather than content (Paulhus, 1988) was evaluated.

4.5.3. Analysis suitability.

To answer RQ1 a factor analysis was conducted using principal component analysis (PCA). This was followed by a regression analysis to determine strength of predictive validity to answer RQ2 to RQ5. Data were reviewed and analyzed for suitability using guidelines described by
Tabachnick and Fidell (Tabachnick, 2012). For factor analysis, sample size and respondent ratio were evaluated and the guideline suggests at least 300 respondents or 150 if solutions have several high loading variables should be obtained (i.e., > 0.80). With 816 useful responses and five variables the dataset meets all the before mentioned criteria. Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and Bartlett’s test of sphericity should be calculated and analysed to assess the factorability of the data. KMO should be above 0.6 and Bartlett’s test should be significant ($p < 0.05$). The analysis KMO ranges from 0.618 - 0.831 and all Bartlett’s tests are significant ($p < 0.05$) For the regression analysis, the data set was investigated to ensure generalisability through sample size to predictor ratio and the guideline $N > 50 + 8m$ ($m$=number of independent variables) applied. With five independent variables, it is suggested that $N=90$ and with 816 useful datasets the criteria are meet. There is no dependency between the independent variables and singularity is thereby ensured. Correlation is not consistently above $r=.09$ and the absence of multicollinearity is thereby ensured. Both independent and dependent variables were screened for outliers using residual plots. The guideline stipulates outliers as those with residual values above 3.3. or less than -3.3, no outliers were discovered in the data set. Tests for normality, linearity, and homoscedasticity were conducted and residuals were found to be normally distributed. Regression analysis was conducted using factors extracted through PCA and internal consistency was above 0.5 (Schmitt, 1996).

4.6. Results

The results from the factor and regression analysis provide evidence that the scale is valid to measure maturity across the food safety culture dimensions. Some limitations of the scale and its predictive ability were discovered and will be reviewed individually in the following sections.
4.6.1. Factor analysis.

Analysis of the self-reported data was completed using factor analysis to answer RQ2 to RQ5. Sub-scale reliability was evaluated using Cronbach Alpha’s and factor analysis completed on each sub-scale to identify key statements and factors for each sub-scale (Tables 11 - 18). The analysis showed that six sub-scales are internally reliable with Cronbach’s Alpha scores ranging from 0.554 to 0.830 (Loewenthal, 2004) and explains 53.05 to 72.10% of the variance, respectively. Factor analysis completed on each sub-scale revealed an overall scale that is both reliable and valid to evaluate three of the five food safety culture dimensions (Jespersen et al., 2017): Values and mission; People systems, and Consistency. Two dimensions; Adaptability and Risks and Hazards were not supported by the self-assessment scale and this result is aligned to the findings of existing research (Jespersen et al., 2016) The sub-scales for FSQ Supervisor were not found consistent and unable to differentiate between cultural dimensions. Three of the manufacturing leaders’ statements were likely seen as low maturity statements instead of high maturity statements (i.e., B21_8, 9, and 11 in Table 13 and Table 14). This is an important finding to discuss as it speaks to the potential perception of maturity in that specific sub-group. Aside from the groups and dimensions it was also confirmed that the dichotomous scale could differentiate level of maturity for each of the role and functional sub-groups.

Table 11: Sub-scale factor analysis for FSQ leaders in a high level of maturity.

<table>
<thead>
<tr>
<th>Statement</th>
<th>Factor 1 (Values and Mission)</th>
<th>Factor 2 (Consistency)</th>
<th>Factor 3 (People)</th>
</tr>
</thead>
<tbody>
<tr>
<td>B11_8 I make sure monthly that my team understands how food safety decisions are made from data in the IT system.</td>
<td>0.871</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B11_6 I analyse and summarize food safety trends monthly for senior leaders.</td>
<td></td>
<td>0.824</td>
<td></td>
</tr>
<tr>
<td>B11_5 I meet monthly with manufacturing leaders to plan and execute preventive actions.</td>
<td></td>
<td></td>
<td>0.804</td>
</tr>
<tr>
<td>Item</td>
<td>Factor 1 (Values and Mission)</td>
<td>Factor 2 (Consistency)</td>
<td>Factor 3 (People)</td>
</tr>
<tr>
<td>------------</td>
<td>-------------------------------</td>
<td>------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>B11_7 I review the preventive control plan(s) quarterly to verify effectiveness.</td>
<td>0.781</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B11_9 I approve investment in food safety tools and technology as requested by employees in food safety at least once a year.</td>
<td>0.697</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B11_16 I circle back after each corrective action has been implemented.</td>
<td>0.441</td>
<td>-0.385</td>
<td></td>
</tr>
<tr>
<td>B11_17 I look to the IT department to set direction for tools used to gather food safety data.</td>
<td></td>
<td>0.952</td>
<td></td>
</tr>
<tr>
<td>B11_18 I review every plant's food safety spending monthly.</td>
<td></td>
<td>0.766</td>
<td></td>
</tr>
<tr>
<td>B11_1 I acknowledge leaders in other functions when they act on food safety.</td>
<td></td>
<td></td>
<td>-0.906</td>
</tr>
<tr>
<td>B11_2 I acknowledge manufacturing leaders who make good food safety decisions.</td>
<td></td>
<td></td>
<td>-0.905</td>
</tr>
<tr>
<td>B11_4 I acknowledge everyone who makes improvements for food safety.</td>
<td></td>
<td></td>
<td>-0.895</td>
</tr>
<tr>
<td>B11_3 I make sure daily that peers and teams have the power to make food safety decisions.</td>
<td></td>
<td></td>
<td>-0.758</td>
</tr>
<tr>
<td>B11_11 I check regularly if my teams have the needed food safety knowledge, skills, or ability to do their job.</td>
<td></td>
<td></td>
<td>-0.432</td>
</tr>
<tr>
<td>% variance</td>
<td>49.32%</td>
<td>13.12%</td>
<td>9.66%</td>
</tr>
<tr>
<td>Total variance accounted for by the rotated factors = 72.10%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 12: Sub-scale factor analysis for FSQ leaders in a low level of maturity.

| Item       | Factor 1 (People) | Factor 2 (Vision and Mission) | Factor 3 ...
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>B11_10 I communicate negative consequences when we have a food safety problem.</td>
<td>0.811</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B11_15 I regularly ask for new food safety problems to be solved by a plant team.</td>
<td>0.768</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B11_12 I direct my employees to consult with others before making food safety decisions.</td>
<td>0.764</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B11_14 I base my food safety decisions on discussions rather than data.</td>
<td></td>
<td>-0.938</td>
<td></td>
</tr>
<tr>
<td>B11_13 I only act when there is a regulatory or customer food safety nonconformity.</td>
<td></td>
<td>-0.902</td>
<td></td>
</tr>
<tr>
<td>% variance</td>
<td>44.40%</td>
<td>26.66%</td>
<td></td>
</tr>
<tr>
<td>Total variance accounted for by the rotated factors = 71.06%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 13: Sub-scale factor analysis for manufacturing leaders in a high level of maturity.

<table>
<thead>
<tr>
<th>Item Description</th>
<th>Factor 1 People</th>
<th>Factor 2 Consistency</th>
<th>Factor 3 Values and Mission</th>
</tr>
</thead>
<tbody>
<tr>
<td>B21_1 I acknowledge leaders in other functions when they act on food safety.</td>
<td>0.889</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B21_2 I acknowledge manufacturing leaders who make good food safety decisions.</td>
<td>0.923</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B21_3 I verbalize that peers and teams have the power to make food safety decisions.</td>
<td>0.716</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B21_4 I acknowledge everyone who makes improvements for food safety.</td>
<td>0.829</td>
<td></td>
<td>-0.847</td>
</tr>
<tr>
<td>B21_5 I gather longer term tools and assess technology needs from plants and bring these for approval by executives once a quarter.</td>
<td></td>
<td>-0.843</td>
<td></td>
</tr>
<tr>
<td>B21_18 I review at least monthly every plant's food safety spending.</td>
<td>0.843</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B21_7 I review the preventive control plan(s) quarterly to verify effectiveness.</td>
<td>0.670</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B21_8 I make sure every month that my team understands how food safety decisions are made from data in the information technology system.</td>
<td>-0.342</td>
<td>0.592</td>
<td></td>
</tr>
<tr>
<td>B21_6 I discuss monthly food safety trends derived from plant data with my peers in the Quality Department.</td>
<td></td>
<td></td>
<td>-0.650</td>
</tr>
<tr>
<td>% variance</td>
<td>49.819%</td>
<td>17.43%</td>
<td></td>
</tr>
<tr>
<td>Total variance accounted for by the rotated factors = 67.25%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 14: Sub-scale factor analysis for manufacturing leaders in a low level of maturity.

<table>
<thead>
<tr>
<th>Item Description</th>
<th>Factor 1 People</th>
<th>Factor 2 Consistency</th>
<th>Factor 3 Values and Mission</th>
</tr>
</thead>
<tbody>
<tr>
<td>B21_15 I ask at least weekly for new food safety problems to be solved by a plant team.</td>
<td>0.904</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B21_8 I make sure every month that my team understands how food safety decisions are made from data in the information technology system.</td>
<td>0.824</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B21_9 I approve at least annually investment in food safety tools and technology as requested by manufacturing leaders.</td>
<td>0.503</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B21_13 I mostly act when there is a regulatory or customer nonconformity.</td>
<td></td>
<td></td>
<td>0.894</td>
</tr>
<tr>
<td>B21_10 I communicate negative consequences when we have food safety problem.</td>
<td></td>
<td></td>
<td>0.767</td>
</tr>
<tr>
<td>Factor 1</td>
<td>Factor 2</td>
<td>Factor 3</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>----------</td>
<td>----------</td>
<td></td>
</tr>
<tr>
<td><strong>People</strong></td>
<td><strong>Consistency</strong></td>
<td><strong>Values and Mission</strong></td>
<td></td>
</tr>
<tr>
<td>B21_12</td>
<td>I direct my employees to consult with others before making food safety decisions.</td>
<td>0.692</td>
<td></td>
</tr>
<tr>
<td>B21_11</td>
<td>I check if my teams have the needed food safety knowledge, skills or ability to do their job.</td>
<td>-0.370</td>
<td></td>
</tr>
<tr>
<td>B21_16</td>
<td>I circle back after corrective actions have been implemented</td>
<td>0.592</td>
<td></td>
</tr>
<tr>
<td>% variance</td>
<td></td>
<td>36.25%</td>
<td></td>
</tr>
<tr>
<td>Total variance accounted for by the rotated factors</td>
<td></td>
<td>62.71%</td>
<td></td>
</tr>
</tbody>
</table>

Table 15: Sub-scale factor analysis for manufacturing supervisors in a high level of maturity.

<table>
<thead>
<tr>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>People</strong></td>
<td><strong>Consistency</strong></td>
<td></td>
</tr>
<tr>
<td>A21_11</td>
<td>I correct food safety behaviours on the spot every time I see an opportunity.</td>
<td>0.815</td>
</tr>
<tr>
<td>A21_9</td>
<td>I acknowledge anyone who goes above and beyond for food safety.</td>
<td>0.785</td>
</tr>
<tr>
<td>A21_15</td>
<td>I use tools and technology to demonstrate correct food safety practices.</td>
<td>0.776</td>
</tr>
<tr>
<td>A21_13</td>
<td>I address food safety challenges immediately based on data collected during production.</td>
<td>0.736</td>
</tr>
<tr>
<td>A21_12</td>
<td>I improve food safety processes every day e.g., take out process steps, reduce resource needs.</td>
<td>0.399</td>
</tr>
<tr>
<td>A21_14</td>
<td>I enter and report food safety performance daily exclusively in the information technology system and nowhere else.</td>
<td>0.761</td>
</tr>
<tr>
<td>A21_10</td>
<td>I solve food safety problems with my team and only involve the Quality department if I need coaching.</td>
<td>0.742</td>
</tr>
<tr>
<td>A21_5</td>
<td>I collect food safety data at least weekly for others to act upon.</td>
<td>-0.491</td>
</tr>
<tr>
<td>% variance</td>
<td></td>
<td>38.66%</td>
</tr>
<tr>
<td>Total variance accounted for by the rotated factors</td>
<td></td>
<td>53.67%</td>
</tr>
</tbody>
</table>

Table 16: Sub-scale factor analysis for manufacturing supervisors in a low level of maturity.

A21_7 | I let the Quality department analyse food safety data. | 0.812 |
<table>
<thead>
<tr>
<th>A21_3</th>
<th>I regularly consult with others before making food safety decisions.</th>
<th>0.767</th>
</tr>
</thead>
<tbody>
<tr>
<td>A21_6</td>
<td>I rarely have time to identify the root cause of food safety problems.</td>
<td>0.748</td>
</tr>
<tr>
<td>A21_4</td>
<td>I mostly act when there is a regulatory or customer nonconformity.</td>
<td>0.730</td>
</tr>
<tr>
<td>A21_8</td>
<td>I improvise when my team does not have the right tools to perform food safety tasks.</td>
<td>0.667</td>
</tr>
<tr>
<td>A21_1</td>
<td>I eliminate food safety issues to avoid others or myself to get penalized.</td>
<td>0.834</td>
</tr>
<tr>
<td>A21_2</td>
<td>I learn how to solve food safety issues as they happen.</td>
<td>0.573</td>
</tr>
</tbody>
</table>

% variance: 28.63%, 20.07%, 14.57%

Total variance accounted for by the rotated factors = 63.28%

---

Table 17: Sub-scale factor analysis for FSQ supervisors in a high level of maturity.

<table>
<thead>
<tr>
<th>A11_3</th>
<th>I immediately congratulate anyone who goes above and beyond for food safety.</th>
<th>0.806</th>
</tr>
</thead>
<tbody>
<tr>
<td>A11_6</td>
<td>I emphasize correct food safety practices every day by use of tools and technology.</td>
<td>0.787</td>
</tr>
<tr>
<td>A11_11</td>
<td>I collect, analyse, and report food safety data daily to others at the plant.</td>
<td>0.664</td>
</tr>
</tbody>
</table>

% variance: 57.01%

Total variance accounted for by the rotated factors = 57.01%

---

Table 18: Sub-scale factor analysis for FSQ supervisors in a low level of maturity.

<table>
<thead>
<tr>
<th>A11_12</th>
<th>I mostly act when there is a regulatory or customer nonconformity.</th>
<th>0.804</th>
</tr>
</thead>
<tbody>
<tr>
<td>A11_4</td>
<td>I solve food safety problems mostly when they cannot be ignored and must be dealt with.</td>
<td>0.703</td>
</tr>
<tr>
<td>A11_8</td>
<td>I design my own tools e.g. spreadsheets and forms, to gather food safety data.</td>
<td>0.671</td>
</tr>
</tbody>
</table>

% variance: 53.05%

Total variance accounted for by the rotated factors = 53.05%
4.6.2. Predictive validity.

Hierarchical multiple regression was used to assess the ability of individual cultural dimensions to predict food safety culture maturity after controlling for ‘social desirability’, ‘tenure’, and ‘age.’ ‘Social desirability’ and ‘age’ were entered at step 1, ‘tenure’ in step 2, and ‘cultural dimensions’ in step 3. Only ‘tenure in company’ contributed significantly of the demographic variables and only two cultural dimensions contributed significantly: ‘Values and mission’ and ‘People systems.’ The total variance explained by the model was 82.5%, $F (3,925) = 1457.68$, $p < 0.001$. Thereby answering RQ2 and RQ3 that two dimensions can predict overall maturity and that controlling for ‘social desirability’, ‘age’, ‘tenure in the food industry’, and ‘tenure in role’ did not contribute to this prediction. However, ‘tenure in company’ did contribute, although small compared to the ‘Values and mission’ dimension, this contribution was still significant ($p = 0.096$).

Hierarchical multiple regression was also used to assess the ability of ‘roles’ and ‘functions’ to predict food safety cultural maturity after controlling for ‘tenure in role.’ ‘Tenure in role’ was entered at step 1, ‘roles’ in step 2, and ‘function’ in step 3. ‘Tenure in role’ and two ‘roles’; ‘leader’ and ‘supervisors’ and two ‘functions’; ‘food safety and quality’ and ‘other’ contributed significantly. The total variance explained by the model was 36.0%, $F (3, 781) = 148.24$, $p < 0.001$. Thereby answering RQ4 and RQ5 that a combination of ‘roles’ and ‘function’ can predict overall maturity and that ‘tenure in role’ does contribute, although small compared to the ‘roles’ and ‘functions’, this contribution is still significant ($p = 0.047$). So, a combination of functional responsibility and span of control can predict food safety culture maturity.
4.6.3. Suggested food safety culture and climate model based on predictive validity.

Based on findings from the existing studies by Ball, De Boeck, and Hinsz and the findings in this research, a model of attributes with predictive validity is suggested to describe, evaluate, and impact the strength of food safety culture (Figure 4). Since, culture and climate are defined differently by the four researchers the model is labelled ‘food safety culture and climate.’ The dependent variables from the individual studies were food safety behavioural intention, self-reported food safety behaviours, keeping food clean and uncontaminated, compliance and participation, execution of food safety practices, and evaluation of food safety culture maturity.

The figure shows cultural attributes connected to food safety performance through predictive analysis and validation. The attributes were segmented by the researcher into three levels; organization, working group, and individual. The attributes were established through empirical research in the food industry conducted individually by four researchers, Ball, De Boeck, Hinsz, and Jespersen.
4.7. Discussion

The objective of this research was to evaluate a scale for evaluating food safety culture and propose a combined and collective model for food safety climate and culture based on existing research with predictive validity. The results show that the proposed scale defined by six sub-scales is reliable and valid to evaluate maturity of food safety culture. The valid scale was developed specifically to differentiate between functions (i.e., manufacturing, FSQ, and other) and an individual’s span of control (i.e., leaders and supervisors) and the statistical results
confirmed this ability through a factor analysis. The degree of factor variance explained by the sub-scales ranged from 53.05 – 72.10%. A hierarchical regression analysis suggests that overall maturity can be predicted by the ‘Values and mission’ dimension in combination with the ‘People system’ dimension while controlling for ‘tenure in role.’ This highlights the scales predictive validity and also the importance of the ‘Values and mission’ dimension similar to the findings suggested by Denison and Kotter in their work on organizational culture (Denison, 1997; Kotter, 1992). A slightly weaker model was found to predict overall maturity by functions and roles. This shows the potential of the macro-cultures and working groups (e.g., manufacturing vs. FSQ and leaders vs. supervisors) as highlighted by both (Ball et al., 2009; Schein, 2004; Schein & Schein, 2017). A model was proposed based on detailed findings derived through predictive studies lead by four researchers; Ball, De Boeck, Hinsz, and Jespersen. Each researcher has led work resulting in individual factors predicting food safety performance. The proposed model is a combination of their individual factors into a three-level model: organizational, working group, and individual. By focusing on this collective set of factors, future researchers can continue the positive trajectory for the food safety culture domain to better understand what cultural attributes contribute to food safety effectiveness.

The weak impact of social desirability on the predictive results was unexpected. Although this is in line with findings from a different scale applied by Hinsz, Nickell, and Park (Hinsz et al., 2007) it was expected to have had a larger impact on food safety results as food safety is a domain, similar to people safety and ecology, that could present some disparity between attitude and actual behaviours. It was also unexpected that the sub-scales for FSQ supervisors were not reliable and valid. This is a significant result and suggests that the pinpointed behaviours identified for this group are not accurate. It is possible that this group is a composite of
individual’s roles with different span of control (i.e., a supervisor in one plant might be called a manager in another). If so, this calls for a revision of the demographic questions related to ‘role.’ The researcher calls for targeted research to better understand this working group e.g., area and span of responsibility, degree of standardization in job expectations and tasks.

In summary, the work contributes significantly to the research domain as it revealed a model of cultural attributes with predictive validity that can be used by practitioners and academicians alike to further food safety effectiveness. When subjected to statistical analysis a scale for evaluating food safety culture maturity was confirmed and it is suggested that a combination of attributes discovered independently by four researchers are integrated and explored through multi-geographical empirical research.
5. Triangulation and the importance of establishing valid methods for food safety culture evaluation

Publication status: under review with Food Research International.

5.1. Abstract

This research evaluates maturity of food safety culture in five multi-national food companies using several methods to explore validity of culture dimensions via method triangulation. As such, the same food safety culture was evaluated using a self-assessment scale, performance document coding, and semi-structured interviews. Weaknesses associated with each individual method are well documented but there are few studies in food safety where a method triangulation approach is used for both data collection and data analysis and these are described here with a clear and transparent audit trail. 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 evaluation methods, it is possible to better direct investments and strengthen food safety culture. This important finding can add to the positive trajectory of the food safety culture paradigm to go beyond a single evaluation of results from generic culture surveys.

5.2. Keywords

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

- Established the importance of triangulation for valid food safety culture evaluation.
- Compared data from scale, performance documents, and semi-structured interviews.
- Confirmed the need for multiple methods for trustworthy evaluation of food safety culture.
- Applied culture coding framework to interview transcripts and performance documents.
- Inter-coder and construct validity, and discrimination in food safety culture profiles.
5.4. Introduction

The understanding of culture to enable organizational effectiveness has been studied at length since 1970 before (Hofstede, 1980, 2001, 2013) studied national culture through his research on cross-cultural organizations, starting with the international (IBM) survey in 1966, and showed predictive validity of his ‘Values Survey Module’ instrument to dimensions of national culture. Denison (1995, 1997, 2012) 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: behavioural data and financial data (Denison et al., 2012; Denison, 1997; 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 evaluation tools. (Guldenmund, 2000) discusses this dichotomy specific to people safety 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. This suggests that other approaches are needed to understand the heterogeneity of organizations. Organizations which are typically made up of sub-groups and macro-cultures (Schein & Schein, 2017).
5.4.1. Theoretical framework.

To link the food safety domain with existing models for organizational culture, safety climate/culture, and food safety climate/culture, a theoretical framework based on eight existing cultural evaluation models was developed (Jespersen et al., In press). The framework compared the details of the specific dimensions, named by each author differently i.e., dimensions, traits, capability areas, categories, elements, but aligned in this framework under the title ‘dimensions.’ Together the five dimensions (Figure 5) encompass all the individual dimensions in the eight culture evaluation systems, although none of the eight systems covers all five dimensions. The framework 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. Robach¹, 2016).

¹ Mike Robach, Global Food Safety Initiative chair of the board.
5.4.2. **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 (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 involves gathering 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 (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 (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 (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 therefore 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, multi-layered, and multi-dimensional 
perspectives of the phenomenon being studied (Kopinak, 1999) and increase credibility of 
scientific knowledge by improving both internal consistency and generalizability (Yeasmin & 
Rahman, 2012). Quoting (Allison & Rootman, 1996)“rigid adherence to one approach at the 
expense or to the exclusion of the other, is destructively parochial and results are often 
incomplete or even inaccurate explanations and by extension, wrongly focused research.” In the 
data analysis phase, triangulation offers several benefits such as verification of overlapping 
results; validation of quantitatively generated constructs through comparison; opportunity to 
probe to investigate potential causes of discrepancies due to the instrumentation or 
misrepresentation of data; construction of case studies to illustrate statistically derived models, 
and clarity of ambiguous and provocative replies or questions (Floyd, 1993). There are 
difficulties related to the application of method triangulation. There must be a 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 (Denzin & Lincoln, 2011). Also, triangulation is time consuming and will increase 
the time needed to complete a study, however the researcher would argue that this approach is 
essential to establish new evaluation methods. Lastly triangulation is carried out with complex 
research designs and there are limited guidelines available to researchers on 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 (Kopinak, 1999). Further literature discussion would be beneficial to overcome gaps in guidance; however, discussion of potential approaches with other researchers to reach consensus around triangulation 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.

5.5. Material and method

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. Three data sets were collected from 21 food manufacturing plants; food safety culture maturity self-assessment responses, food safety document coding, and semi-structured interviews with plant leaders (Figure 6).
Each method was selected to provide as much data as possible on the same phenomenon – food safety culture – to counter weaknesses in each method, and to make use of already existing data sources e.g., food safety documents. For validating the triangulation method, data from a sub-set of five plants were analysed (Table 19). One plant per company was selected randomly from the list of 21 plants.

Table 19: Sources by plant and data type for the five plants in the triangulation validation.

<table>
<thead>
<tr>
<th>Plant ID</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-assessment responses</td>
<td>63</td>
<td>14</td>
<td>10</td>
<td>15</td>
<td>71</td>
</tr>
<tr>
<td>Performance documents</td>
<td>5</td>
<td>1</td>
<td>6</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Semi-structured interviews</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

Figure 6: Methods and data triangulation applied to evaluate of food safety culture.
5.5.1. **Methods strengths and weaknesses.**

Three methods were selected for the study of triangulation (Figure 6). Strength and weaknesses of each of the three methods were identified to show how each method can mitigate weaknesses in others. Method 1- Scale: The strengths of scales or surveys 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 analysing 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, provide a moving trail of investigation based on the respondents answer. Interviews are especially suitable for collecting data on sensitive topics because of the interviewer’s 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 group's language and 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 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.

5.5.2. Response analysis of self-assessment scale.

All salaried staff in each manufacturing plant were invited to participate in an online survey between October 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 a total of 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, behavioural 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 et al., In press) amended with the findings from predictive analysis (Jespersen & Edwards, Under review).
5.5.3. **Content analysis of performance documents.**

The content analysis of food safety performance documents provided 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 October 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 the literature and analysis of food safety culture and organizational culture evaluation tools. The coding framework (Table 20) was defined using the theoretical framework (Figure 5) of food safety culture and translated into nodes in NVivo [Computer Software] QSR International, Doncaster, Australia. Each document was imported into NVivo and all documents were coded by two independent researchers.

5.5.4. **Content analysis of semi-structured interviews.**

Semi-structured interviews with a 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 blinded 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 20).
Table 20: Coding framework used in the content and textual analysis. Adapted from Jespersen, Griffiths, and Wallace (2017).

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

### 5.5.5. Content coding.

The content was coded using practices already applied in the food safety domain (Wallace, 2009). The process for coding content (Figure 7) was followed by the two independent coders to ensure validity of data. The process consisted 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 0.8 but that 0.9 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 0.9 (90% agreement). Detailed research
questions were defined (step 1), 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 connected the coded data to the theoretical framework and the research domain. Following this, coders were trained (step 4) and two documents coded by the same coders (step 5). The results were analysed by detailed review of textual 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 the test documents were recoded (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 textual data review (step 7). Midway through the process discussions between coders allowed comparison of experience, discussion of coding difficulties and issues. These results led to finalization of the 30 documents to be used in the study (step 8). Finally, the data were analysed to derive information to answer the research questions (RQs) (step 9).
Figure 7: Coding process applied to deriving data through content analysis.
5.5.6. Data triangulation.

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 1 to Stage 5, (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 - 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.
5.6. Results

5.6.1. Self-assessment results.

Differences in overall, aggregated maturity ratings through the self-assessment scale for the five plants in the sub-set were not statistically significant for the overall maturity $F(4,182) = 0.273, p = 0.895$ (Table 21). However, in the larger study a difference was found between the plants and dimensions for all plants in the study $F(4,781) = 5.727, p = < 0.05$ (Jespersen and Edwards, 2017, under review).

Table 21: Plant results from maturity self-assessment scale, total and by dimension.

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

The dimensions of ‘Risk awareness’ and ‘Adaptability’ were not reported in the Jespersen et al (2016) tool and so do not feature in this aspect of the analysis (Jespersen, Griffiths and Wallace, 2017).
5.6.2. Coding comparisons.

A comparison of coders by dimension was completed. A total 4,522 references were coded in 10 interview transcripts and 20 performance documents. Coders are considered similar if within the set standard of minimum 90% agreement. Agreement between coders was calculated for each dimension and lowest level of pairwise agreement was calculated to be 90.4%. This result was obtained after coding and recoding. As such, content from two dimensions needed recoded; ‘Values and mission’ and ‘Risk awareness.’ The bar chart (Figure 8) shows that coders were within 90% agreement on scoring except for ‘Values and mission’ (69% agreement) and ‘Risk awareness’ (79% agreement).

![Figure 8: Codes by dimension with pairwise comparison and difference by coder.](image-url)
In looking at the sub-nodes for ‘Values and mission’ (Figure 9) most of this difference comes scoring in the 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 C\ 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.

Figure 9: ‘Values and mission’ dimension by sub-node and by coder.
For ‘Risk awareness’ (Figure 10), 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 records, 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 had more detailed word coding on hazards and risks and coder B coded references to specific bacteria and risks and hazards more generally.

![Figure 10: ‘Risk awareness’ dimension by coder and sub-nodes.](image)

5.6.3. **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 11). The Pearson’s coefficient shows values at or equal to 0.5 or above for similar items and values less than of 0.5 for items distinctly different. By use of the analysis, several items in the current dimensional framework were confirmed and new content deduced. 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/non-conformities, (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 food safety culture dimensions but also add more structure to the sub-nodes. This suggested dimensional framework (Figure 12) 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 analysed and thereby add to the overall analysis of an organizations food safety culture maturity.
Figure 11: Nodes clustered by word similarity.
Figure 12: Revised dimension framework and sub-nodes based on cluster analysis.
5.6.4. 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 13). 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.
5.6.5. 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. A mean score of maturity, from 1 to 5, was calculated for the self-assessment assessments, the performance document coding, and the interview coding results, for each of the five plants (Figure 14). 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 individual’s rating of food safety maturity, the documented performance, and what was said by leaders in conversation are telling similar stories. In a re-evaluation situation, it could be considered to only apply one of the three evaluation methods to save time and effort. P3 shows the greatest difference between methods. This means that individuals rate the plant’s 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 identify areas of change. P1 and P4 have comparatively low scores for the documented performance compared with their other measures and it would be interesting to look at the purpose of the submitted documents and if there is an opportunity to have better use 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 supervisors (Jespersen et al., 2016).

![Plant maturity - mean values as per method triangulation.](image)

**Figure 14:** Plant maturity - mean values as per method triangulation.

Ledger: Self-assessment scale result (dot), performance document coding result (diamond), and interview coding result (triangle).
5.7. Discussion

The objective of this research was to develop, apply, and analyse 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 analysed for inter-coder and construct validity, and ability to discriminate between dimensions, functions, and data sources. 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 evaluated 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 above 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 (Boeck et al., 2015; Wright) 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 5) 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 analyse 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 - Reactive’ to ‘Stage 3 - Know’ (Jespersen & Edwards, Under review; Jespersen et al., 2016) for all plants and documents. This could be due to the geographical dispersion of the plants, this subset all being in North America, and therefore under similar 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 bias 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.
6. A path to maturing culture for realizing financial gain

Under review with co-authors and targeted for publication in Food Protection Trends

6.1. Abstract

Research into the connection between organizational effectiveness and culture have been researched and documented since the early nineties. A clear connection between financial performance and organizational culture have been established directly linking strong cultural attributes to financial performance. This research proposes a similar link to food safety culture through the measure of maturity and cost of poor quality. Through data collected at five multi-national food companies the connection is explored and a dynamic model for food safety culture and climate as well as a food safety maturity model is suggested as the frameworks for evaluating food safety culture.

6.2. Keywords

Food safety culture, financial performance, evaluation models

6.3. Highlights

- Financial performance and food safety maturity
- Culture dynamics and actions
- Maturity model for food safety
6.4. Introduction

To solve the specific challenge of food safety performance it is now widely accepted that food safety culture plays an integral role. It is also accepted that to get to a stronger macro culture (e.g., safety culture, food safety culture, innovation culture), one must consider the broader organizational culture and its effectiveness (Schein & Schein, 2017). What comes first, effectiveness or culture, becomes a bit of a ‘chicken or the egg’ discussion but recognizing an organizational culture and its macro cultures as enablers of organizational effectiveness including food safety performance is the statement of importance. Quoting Harvard professor emeritus James L. Heskett “Organization culture is not a soft concept, its impact on profit can be measured and quantified. And in organizations with large numbers of customer-facing employees, the sum of the effects of employee turnover, referrals of potential employees by existing ones, productivity, customer loyalty, and referrals of new customers attributable to culture can add up to half of the difference in operating income between organizations in the same business.” It is this paradigm between perceived softness and hard measures that makes organizational cultures and macro cultures both intriguing and challenging for practitioners and scientists to understand.


(Kotter, 1992) studied culture in 207 U.S. firms through surveys and detailed interviews and found a direct connection between organizational culture and financial performance. While the authors stated that many confounding variables impact an organization’s financial performance, they also discovered a substantial difference in financial performance between performance-enhancing cultures and non-performance enhancing cultures for two groups of 12 companies (Table 22). In the group that invested in a performance-enhancing culture, the
increase across the financial indicators were all above 280%. Organizational values in the performance-enhancing cultures are described as managers deeply caring about customers, strongly value people and processes that create useful change. Conversely values in non-performance enhancing cultures are described as managers mostly caring about themselves and their immediate work group, valuing orderly and risk-reducing management processes.

Table 22: Financial performance differences between companies who invested in a performance-enhancing culture and those that did not (Kotter, 1992).

<table>
<thead>
<tr>
<th></th>
<th>Average increase for 12 firms with performance-enhancing cultures</th>
<th>Average increase for 12 firms without performance-enhancing cultures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Revenue growth</td>
<td>682%</td>
<td>166%</td>
</tr>
<tr>
<td>Employment growth</td>
<td>282%</td>
<td>36%</td>
</tr>
<tr>
<td>Stock price growth</td>
<td>901%</td>
<td>74%</td>
</tr>
<tr>
<td>Net income growth</td>
<td>756%</td>
<td>1%</td>
</tr>
</tbody>
</table>

Similar to Kotter and Heskett, (Graham, Harvey, Popadak, & Rajgopal, 2017)explored the connection between organizational culture and cultural effectiveness. An effective culture was defined by the authors as “one that promotes the behaviours needed to successfully execute the firm’s strategies and achieve its goals.” Data were gathered from 1,348 North American firms through online surveys and interviews with senior executives. The authors found that organizational effectiveness is the result of interaction between organizational values, norms, and formal systems. In this context, values are defined as the aspirations of the organization, norms the day-to-day practices that live out the values, and formal systems the written policies and procedures. Human behaviours are conditioned through the integration and adaptation of organizational norms, and norms are in turn an interpretation and adaptation to the organizations values and formal systems. The gathered evidence shows that norms enhance business outcomes but values do not. The research also suggests that the marketplace influences executives’
investment in culture as well as the organizational values they promote. This external adaption is also captured in definition of organizational culture as “… the accumulated shared learning of the group as it solves its problems of external adaptation and internal integration…” (Schein, 2004). Schein thereby integrates external and internal triggers of change as confirmed in the findings by (Graham et al., 2017).

The ‘Great Place to Work® Institute’ is a global organization dedicated to providing knowledge on how to build and sustain high performing work place culture (retrieved from greatplacetowork.com). ‘The Great Place to Work®’ database contains data from more than 5,500 companies operating in 45 countries collected through annual assessment surveys. Through analysis of the data, researchers found that proclaimed values appeared irrelevant to an organization’s effectiveness. This supports the findings from Graham et al. that values alone do not drive business outcomes but norms do (Guiso, Sapienza, & Zingales, 2014). The research also shows that if executives are perceived as trustworthy and ethical the company’s performance will be stronger. In analysing S&P 500 companies the researchers found that 80% of the companies mention ‘innovation’ followed ‘integrity and respect’ in their corporate values. A culture of integrity was found to add value and this type of culture was weaker among publicly traded companies than non-publicly traded companies. Integrity was also positively correlated with financial performance and attractiveness of job offerings and negatively correlated with degree of unionization.

6.4.2. Organizational culture and quality maturity.

(Crosby, 1972) is the first to define quality as ‘conformance to requirements’ and to make the claim that “management unintentionally cause an increased cost of quality for the organization by not understanding this simple definition.” Crosby also suggests, like Kotter and
Heskett (1992), that a culture revolution through a planned strategy is the key to reducing cost of quality in any organization. Through the ‘Quality management maturity grid’, Crosby defines six measurement categories by which an organization can evaluate its current stage of quality maturity. Using the grid, he demonstrates the connection between decreasing cost of quality and increasing quality culture maturity; thereby directly linking organizational culture to organizational financial performance. He shows how as much as 20% of sales can be lost as cost of poor quality in contrast to a high-level maturity culture at 2.5%. The American Society of Quality (ASQ) breaks cost of poor quality (COPQ) into four activities, prevention costs, appraisal costs, and internal and external failure costs (retrieved from asq.org) through these activities costs related to e.g., systems maintenance and training, conformance to specification, verification activities, waste and scrap, and complaints are track to quantify percentage of sales allocated to poor quality.

### 6.4.3. Organizational culture and food safety.

Research on connecting food safety culture to operational and food safety performance has thus far been focused on drivers of food safety behaviours and less about the integrative system from company values to norms as the means through which values are expressed as behaviours. Several authors have published material to educate food safety practitioners and scientists alike to understand food safety culture and how to construct a strong and positive food safety culture (Griffith, 2010; Griffith et al., 2010a; Griffith et al., 2010b; Jespersen & Huffman, 2014; Powell et al., 2011; Yiannas, 2009, 2015). More detailed research was conducted to explore the connection between organizational and safety culture and how to bring these learnings into the food safety domain (Boeck et al., 2015; Griffith, Jackson, & Lues, 2017; Griffith et al., 2010b; Jespersen & Huffman, 2014; Powell et al., 2011). Recently, scientists have
focused on demonstrating predictive validity of cultural attributes specific to the food safety domain (Ball et al., 2009; Betts, 2014; Clayton & Griffith, 2008; De Boeck et al., 2016; De Boeck et al., 2017; Hinsz & Nickell, 2015; Hinsz et al., 2007; Hinsz & S., 2004; Jespersen & Edwards, Under review; Nickell & Hinsz, 2011). To continue this positive trajectory, this research seeks to address two questions: 1) Is there financial value in focusing on food safety culture? and 2) What is the path to realizing this value?

6.5. Material and method

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. Total 816 individuals responded to the survey, 379 documents were coded and analysed, 42 on-site interviews were conducted and coded (Table 23).

Table 23: Data collected from the five participating companies.

<table>
<thead>
<tr>
<th>Category</th>
<th>Measure</th>
<th>Company 100</th>
<th>Company 200</th>
<th>Company 300</th>
<th>Company 400</th>
<th>Company 500</th>
<th>Mean (Total)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data</td>
<td>Responds rate</td>
<td>72</td>
<td>77</td>
<td>72.5</td>
<td>77</td>
<td>59</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>(Percentage)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Performance documents</td>
<td>268</td>
<td>3</td>
<td>33</td>
<td>50</td>
<td>25</td>
<td>(379)</td>
</tr>
<tr>
<td></td>
<td>(#)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Interviews (#)</td>
<td>22</td>
<td>8</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>(42)</td>
</tr>
</tbody>
</table>

Page 115 of 173
Three data sets were collected from 21 food manufacturing plants; food safety culture maturity self-assessment responses, food safety documents, and semi-structured interviews with plant leaders (Figure 15).

Figure 15: 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 method, and to make use of already existing data e.g., food safety documents.

6.5.1. 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, behavioural 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 et al., In press) amended with the findings from (Jespersen & Edwards, Under review)

6.5.2. 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 was defined using the theoretical framework of food safety culture (Jespersen et al., 2017) and translated into nodes in NVivo [Computer Software] QSR International, Doncaster,
Australia. Each document was imported into NVivo and all documents were coded by two researchers.

### 6.5.3. 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.

Since 11 of the 21 plants were in company 100 it was also to be expected that more data from responses, documents, and interviews were collected there in comparison to the other four companies. However, this provides a mean to assess the suitability of the evaluation system to large as well as smaller companies. Findings from demographics and cultural performance were summarized by company and a baseline calculated. Individual plant evaluations and aggregated company results were discussed with experts and company leaders for additional validation. Based on the learnings from the study a culture model was develop and the food safety maturity model revised. Generalizability was investigated by benchmarking demographics from participating companies to Compustat firms.

### 6.6. Results

Organizational culture as defined by Schein (2004) is about adapting and integrating change. As such, this research showed a valid connection between dimensions in the theoretical
framework (Jespersen et al., 2017), attributes demonstrated to predict food safety performance (Betts, 2014; De Boeck et al., 2017; Hinsz & Nickell, 2015; Jespersen & Edwards, Under review), and factors impacting organizational effectiveness (Denison, 1997; Kotter, 1992). Through this it was also seen that a valid culture evaluation system is one that evaluates the ‘actions’ between organizational culture effectiveness and organizational culture building blocks e.g., norms and values; working group macro-cultures and individual intent and behaviours. A valid culture evaluation system is, therefore, not only one that only checks the content of each building block (e.g., is there a learning system in place? are hygienic design guidelines applied?) but one that evaluates the actions (e.g., making decisions about training resources, managing engineering consequences for hygienic design) between building blocks. Actions that act as evidence for the organization’s ability to adapt and integrate change.

The evaluation system developed through this research made use of three research methods (i.e., survey, content analysis, and interviews) to capture the details of each action in the culture model (Figure 16). The model consists of cultural building blocks and actions. Each building block is connected to others through actions. There are four main building blocks; Individual intent and behaviours, working group learned and shared assumptions, and Organizational culture norms, and Organizational effectiveness.

Each arrow is an action that causes the next building block to adapt and integrate change. In this study, each action was studied through one or more of the three methods. As such, data related to the adaption of organizational culture to its external environment (Figure 16 – arrow 1) and the reflection of values in the formal systems (Figure 16 – arrow 2) through on-site visits and the performance documents. Adaptation and integration of values in the organizational culture (Figure 16 - arrow 3), through on-site interviews. The adaptation of norms to the formal systems
through the performance document content analysis (Figure 16 - arrow 4). Adaptation of the working groups assumptions to the organizational culture (Figure 16 - arrow 5) through the survey and the on-site interviews. The individuals’ integration in the working group (Figure 16 - arrow 6) and the organizational culture (Figure 16 - arrow 7) through the self-assessment scale, including social desirability assessment. This clearly showed that reliance on a single method would not provide a comprehensive and trustworthy evaluation of culture. Not simply due to the inherent weaknesses in any research method, as discussed in Chapter 5, but also due to the inherent complexity of any organization. An example was drafted to illustrate this complexity.
Figure 16: Culture model: Organizational effectiveness, culture building blocks, and actions.

‘Actions’ i.e., adapt and integrate

Culture building blocks
ARROW 1: ‘Recall for *Salmonella* contaminated products sold by category competitor’ (marketplace) ⇨

ARROW 2: ‘Automatically orders are given to review procedures to ensure risk is managed at company x (norm) ⇨ ‘Management system includes written procedure for environmental monitoring’ (formal system).

ARROW 3: ‘Leaders automatically debate if the values at company x would help or hinder a similar event at company x’ (norm) ⇨ ‘At company x we care for customers, consumers, and the world we live in’ (value).

ARROW 4: ‘At company x we care for customers, consumers, and the world we live in’ (value) ⇨ ‘Management system includes written procedure for environmental monitoring’ (formal system).

ARROW 5: ‘A recall automatically increases food safety attention for a little while’ (norm) ⇨ ‘We know we will be talked to about food safety every time a competitor has a recall’ (working group assumption)

ARROW 6: ‘We know we will be talked to about food safety every time a competitor has a recall’ (working group assumption) ⇨ ‘I take it serious if my colleagues correct me if I make a food safety mistake but this mostly happens after a category recall’ (individual behaviour).

ARROW 7: ‘A recall automatically increases food safety attention for a little while’ (norm) ⇨ ‘I am confused about the importance of food safety but do what I think is expected of me’ (individual behaviour).
6.6.1. Company aggregated results and baseline.

Demographics and cultural performance results of each company were summarized to evaluate the content of the culture model and the maturity model. A baseline (mean and total) was established by aggregating all data from all companies (Table 24).

Table 24: Aggregated company demographics, culture performance results, and baseline (mean and total).

<table>
<thead>
<tr>
<th>Category</th>
<th>Measure</th>
<th>Company</th>
<th>Mean (Total)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>200</td>
</tr>
<tr>
<td>Demographics</td>
<td># plants</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Years in food industry</td>
<td>10-14</td>
<td>10-14</td>
</tr>
<tr>
<td></td>
<td>(mean)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Years in the company</td>
<td>10-14</td>
<td>5-9</td>
</tr>
<tr>
<td></td>
<td>(mean)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Years in current role</td>
<td>5-9</td>
<td>5-9</td>
</tr>
<tr>
<td></td>
<td>(mean)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Age (mean)</td>
<td>45-54</td>
<td>34-44</td>
</tr>
<tr>
<td>Functional distribution</td>
<td>(%MFG/%FSQ)</td>
<td>86/14</td>
<td>82/18</td>
</tr>
<tr>
<td>Role distribution</td>
<td>(%Leader/%Supervisor)</td>
<td>37/63</td>
<td>46/54</td>
</tr>
<tr>
<td>Cultural performance</td>
<td>Maturity [1-5]</td>
<td>3.36</td>
<td>3.31</td>
</tr>
<tr>
<td></td>
<td>COPQ (Bill. $)</td>
<td>2.4</td>
<td>0.039</td>
</tr>
</tbody>
</table>
6.6.1.1. **Demographics.**

Individual participants in each company (n=1,273) were asked a series of demographical questions to analyse potential impact on cultural performance. As such, mean age of all respondents (n=816) was 34-44 years, with 10-14 years of experience in the food industry and current company, and 5-9 years in current role. Comparing to baseline, respondents in company 100 were older – 45-54 years. Respondents in company 200 had less experience in both current company and role – 5-9 years. Respondents in company 300 also had less experience – 5-9 years in current company but 2-4 years in role and thereby the least experience in the study. Respondents in company 400 where older than benchmark – 45-54 years and had the longest tenure in the industry – 15-19 years and the company and role – 10-14 years. Respondents in company 500 also had low tenure in current role – 2-4 years, but unlike company 300, at benchmark for experience in both industry and company – 10-14 years. Mean industry tenure ($F(3, 925) = 6.88, p < .001$), company tenure ($F(3, 925) = 5.74, p < .001$), tenure in current role ($F(3, 925) = 5.89, p < .001$) and age ($F(4, 925) = 7.65, p < .001$) were all found to be significantly different between the companies.

A functional ratio was calculated as percentage manufacturing and percentage FSQ respondents of total respondents for each company (%MFG/%FSQ). As such, companies 100, 200, and 400 were similar – 86/13, 82/18, and 85/12. Respondents from company 300 were mostly manufacturing – 92/8 and company 500 had the lowest manufacturing participation – 78/22. Despite these differences, many respondents in all companies were, not surprisingly, from manufacturing. It should be noted that manufacturing in this context includes all functions except food safety and quality with direct reporting relationship to a senior manufacturing leader e.g., S. VP Manufacturing or plant manager e.g., sanitation, maintenance, and finance.
Span of control was analysed through respondents answer to what role the respondent was in, i.e., ‘supervisor’, ‘leader’, and ‘functional leader.’ A ‘supervisor’ was anyone with direct responsibility for plant floor operatives/associates or individuals in a technician or generalist role. ‘Leaders’ anyone with direct responsibility for supervisors, and ‘functional leaders’ those with direct responsibility for leaders. For this ratio (%Leader/%Supervisor), ‘leaders’ and ‘functional leaders’ were combined. Ratios for companies 100 and 500 were relatively similar – 37/63 and 35/65 – with most supervisors in the study. Respondents for company 200 were almost an even split at 46/54 and companies 300 and 400 had the most leaders of the five companies responding – 58/42 and 55/45.

6.6.1.2. Cultural and financial performance.

Aggregated maturity by company was calculated by calculating a mean score of all data collected through the three methods. As such, companies 100, 200, and 400 were in the ‘Know’ stage at 3.36, 3.31, and 3.05. Company 300 had the lowest maturity of 2.80 and in the ‘React’ stage. Finally, company 500 had the highest maturity of 4.01 and in the ‘Predict’ stage. Maturity was found to be significantly different ($F (4, 785) = 5.727, p < .001$) across the five companies. Cultural performance was also evaluated through the company social desirability score. Social desirability is a social science research measure that quantifies the tendency of study participants to answer questions in such a way as to be viewed favourably by others. It can take the form of over-reporting ‘good behaviour’ or under-reporting ‘undesirable behaviour’ (Chapter 3). In analysing social desirability, the companies were also found to be significantly different, ($F (4, 460) = 10.079, p < .001$). Companies 100 and 600 scored the lowest at mean 4.10 and 4.98 out of 18 possible scores. Company 300 had the highest score of all at 7.56 with companies 200 and 400 and lower at 6.67 and 7.16.
COPQ was calculated to evaluate the financial gain related with stage of maturity. By applying the guidelines by Crosby (1972) and ASQ (retrieved from asq.org) a percentage of sale was allocated to COPQ based on the stage of maturity found through the research. To show the path to financial gain or loss three estimates of COPQ were calculated; (1) based on actual evaluation, (2) one stage lower, and (3) one stage up from the maturity evaluation results (Figure 17). Using the evaluation results for total COPQ, company 100 spent most, due to it also having the highest annualized sales but this is followed by company 300 with the second highest COPQ due to its low maturity rating. Collectively the companies spent 1.14 billion sales dollars on COPQ annually in their current stages of maturity. If they all slide back one stage they spend 0.38 billion sales dollars more and if they all go up one stage they save an additional 0.43 billion sales dollars.
6.6.2. Path to food safety maturity.

The evaluations were conducted using the original food safety maturity model proposed by Jespersen et al (2016). Based on the findings in this research the model was further developed to incorporate learnings from the five companies and increase its applicability. A glossary was developed to bring clarity to the model (Table 25).

Table 25: Maturity model glossary.

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adaptability</td>
<td>Organizational culture dimension that describes organization’s approach to innovation and sustaining change.</td>
</tr>
</tbody>
</table>
### Term | Definition
---|---
Consistency | Organizational culture dimension that describes an organization’s adoption of formal systems, technology, and data to drive food safety activities and decisions.
Ethics | Moral principles that govern a person's behaviour or the conducting of an activity.
Integrity | The quality of being honest and having strong moral principles.
Mission | Communicates an organization's reason for existence.
Norms | In sociology, norms are social expectations that guide behaviour. Norms explain why people do what they do in given situations. From a sociological perspective, social norms are informal understandings that govern the behaviour of members of a society. Social psychology recognizes smaller group units, such as a team or an office, may also endorse norms separately or in addition to cultural or societal expectations. Norms are the means through which values are expressed in behaviour.
People system | Cultural dimension describing an organization’s decision around behavioural recognition, building employee competencies, and how these are used in food safety communication.
Responsibility | The state or fact of having a duty to deal with something or of having control over someone.
Trust | Firm belief in the reliability, truth, or ability of someone or something.
Values | Organizational culture dimension related to an organizations ethics and integrity related to food safety responsibilities and expectations.

The dimensions of the maturity model were aligned to the theoretical framework (Jespersen et al., 2017) and the content of the model structured around specific values derived from the values of each participating company and the findings from (Guiso et al., 2014). As such, the values align to specific food safety culture dimensions but fit into commonly defined
and used corporate values. The content for each value and stage was defined as norms by finishing the sentence ‘Food safety <VALUE> at company x can be described as <STAGE> through …’ (Table 26). This was different from the content of the original model where content was derived by summarizing the behaviours behind each capability area and stage. The adapted model tied dimensions, values, and norms to food safety culture through each stage of maturity, resulting in a model that was easier for organizations to apply in the context of their own organizational values and norms. This also provided a path to improve food safety culture directly tied to operational values and to organizational effectiveness as shown by (Graham et al., 2017; Guiso et al., 2014; Kotter, 1992).
<table>
<thead>
<tr>
<th>Dimension</th>
<th>Values</th>
<th>Stage 1 (Doubt)</th>
<th>Stage 2 (React)</th>
<th>Stage 3 (Know)</th>
<th>Stage 4 (Predict)</th>
<th>Stage 5 (Internalize)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Values and Mission</td>
<td>Integrity and trust</td>
<td>Employees have little trust that management will act on food safety without external pressure</td>
<td>Employees trust that management will act and do the right thing for food safety after an issue have occurred</td>
<td>Everyone trust that food safety issues are solved because we know it protects our business.</td>
<td>Everybody are trusted to invest in food safety information to make future performance stronger</td>
<td>Frontline employees are trusted to act to correct and celebrate food safety performance on their line/in their area</td>
</tr>
<tr>
<td>Being responsible</td>
<td></td>
<td>Everybody readily carries responsibility but it is unclear what that means.</td>
<td>Detailed food safety responsibility is written into job descriptions for everybody</td>
<td>Detailed food safety responsibility is written into job descriptions for everybody</td>
<td>Detailed food safety responsibility is written into job descriptions for everybody</td>
<td>Detailed food safety responsibility is written into job descriptions for everybody</td>
</tr>
<tr>
<td>Ethics</td>
<td></td>
<td>Moral principle... don’t look.</td>
<td>Moral principle... invest if we must</td>
<td>Moral principle... improve system</td>
<td>Moral principle... reduce cost by taking out variation</td>
<td>Moral principle... grow business</td>
</tr>
<tr>
<td>People System</td>
<td>Reward and recognize</td>
<td>Individuals complete food safety tasks out of fear for negative consequences.</td>
<td>Individuals are recognized sporadically after having solved a food safety problem</td>
<td>Leaders recognize teams and individuals according to a documented system of positive and negative consequences</td>
<td>Leaders reward teams for collectively improving food safety processes/procedures</td>
<td>Cross functional/level teams nominate other teams for being proactive and thinking strategic around food safety</td>
</tr>
<tr>
<td>Dimension</td>
<td>Values</td>
<td>Stage 1 Doubt</td>
<td>Stage 2 React</td>
<td>Stage 3 Know</td>
<td>Stage 4 Predict</td>
<td>Stage 5 Internalize</td>
</tr>
<tr>
<td>-------------------------</td>
<td>------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Competently communicating</td>
<td>Top-down ‘tell’ with little ‘why’ content and understanding of the importance of the task</td>
<td>Food safety information is communicated by FSQ as problems occur using, if available, facts discovered as the problem was solved</td>
<td>There is a deep understanding of the food safety system and performance is communicated by some functional on a regular basis</td>
<td>Frontline leaders are having regular communications on food safety system and performance using data and tracking the teams’ improvement actions.</td>
<td>Food safety communication cadence is an organizational habit that involves everybody in specific team discussions.</td>
<td></td>
</tr>
<tr>
<td>Together we make the difference Innovate</td>
<td>silos… problem communication… fragmented delivery of information…</td>
<td>Change is analyzed and incorporated into written food safety system including changes to competencies/job descriptions</td>
<td>Innovation is driven by data internally to reduce food safety costs</td>
<td>Innovation is suggested by frontline teams and bubbling up to impact companywide system. Quick to adapt as they have technology interface in their hands.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adaptability Innovate</td>
<td>Scrambling to meet changed requirements</td>
<td>We know change is coming and will deal with it last minute…</td>
<td>We know the change and have analyzed the impact on individuals and teams according to a pre-defined change curve…</td>
<td>We look for cost reduction opportunities and plan these in our continuous improvement program…</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Embrace and drive change</td>
<td>Nothing is stable so it does not matter if we must change…again</td>
<td>We know the change is coming and will deal with it last minute…</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consistency Data and reporting</td>
<td>Data are not used to solve problems and mostly sitting in a</td>
<td>It is left to the individual to identify needed</td>
<td>Leading indicators are used to find root cause of food safety</td>
<td>Leading indicators are continuously updated through</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Page 131 of 173
| Dimension | Values | Stage 1  
Doubt | Stage 2  
React | Stage 3  
Know | Stage 4  
Predict | Stage 5  
Internalize |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Technology enabled success</td>
<td>filing cabinet or in unused reports</td>
<td>data and ways to derive information from these</td>
<td>problems and solutions are built into the food safety management system</td>
<td>precisely and accurately collected data</td>
<td>to improve food safety systems</td>
<td></td>
</tr>
<tr>
<td>Unstructured problem solving to remove the immediate pain.</td>
<td>Technology is bought in reaction to a specific need e.g., faster pathogen testing results</td>
<td>Technology is seen in the context of the business system to integrate functions, procedures, and capabilities (e.g., ERP specification system)</td>
<td>Automation is used frequently and seen as an integral part of reducing food safety cost.</td>
<td>ERP is used in an integrated way with automate workflows that make the enterprise quick to adapt.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>'plan, do, check, act’ with emphasis on control and expectation of 100% perfect solutions from the get go.</td>
<td>'plan, do, study, act’ with emphasis on study and an iterative approach to improvement</td>
<td>Identifying risks through horizon scanning and continuous improvement followed by mitigation plans built into the food safety system.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
6.7. Discussion

To further this positive trajectory of the food safety culture domain, this research searched for answers to two questions; is there operational value in focusing on food safety culture? and is there a path to realizing this value?

Through the data collected and findings in literature a culture model was developed. The model consisted of cultural building blocks and actions. The content of the model was used to update the food safety maturity model to better include all levels of culture; organizational, working group, and individual. To illustrate how the path to internalized food safety could lead to financial gain a cost of poor quality value was calculated. This value was calculated by applying the guidelines from Crosby (1972) as a percentage of sales. Each stage in the maturity model relates to a percentage e.g., stage 1 ‘doubt’ COPQ equates to 20% of sales and stage 5 ‘internalized’ to 2.5% of sales. For a larger company, like 100, with 20 billion dollars in annual sales and in maturity stage 3, COPQ equates to 2.4 billion dollars annually. For a smaller company, like 200, with 325 million dollars in annual sales and in stage 3, COPQ equates to 39 million dollars annually. So proportionally COPQ is of significant magnitude regardless of company size but the effort of realizing a reduction in COPQ could differ significantly. As such, reducing e.g., COPQ through reduced verification activities could require greater effort in a larger company with more layers of employees and greater centralization of policies and procedures. It should also be noted that gathering specific data to calculate actual COPQ requires significant resources that might be easier for companies in higher maturity stages e.g., 500 in ‘Stage 4 - Predict’ to justify as this and ‘Stage 5 - Internalized’ are signified by better process control and a focus on removing waste.

Content of the maturity model was revised to connect values and norms to food safety and
quality norms. The model provides a path to reduced COPQ with focus on five dimensions, Values and mission, People systems,

Through the company study both demographic and cultural performance was analysed resulting in an aggregated maturity score, a social desirability score, and a COPQ for each company. Homogeneity of cultures are often assumed and collecting demographic data is one way of exploring this assumption. The five companies were significantly different in employee composition i.e., age, tenure, function, and role. Ranging from mature, experienced, and consistent to younger, new in role, and adaptable. Experience, as measured through tenure in industry, company, and role, appear to impact both maturity score and social desirability. Company 500 was the most mature culture and the respondents had same length of experience in industry and company but lowest experience in role. Does this suggest that rotation helps mature culture? But only if with the company for longer time? Comparing to the lowest maturity – company 300 – this would support this hypothesis and the researchers call for future confirmatory studies to further explore this potential connection. All food safety and quality participants evaluated maturity significantly higher than their colleagues in manufacturing and other functions. There was no difference social desirability between the functions leading to the conclusion that food safety professionals simply believe maturity to be stronger. This is to be expected that the function with the main responsibility and expertise see more process control and understanding of food safety. Is this an advantage or blind-spot? Does it reflect the food safety functions isolation or manufacturing functions scepticism? Coding the interview statements there was a slight tendency for both. Participants in company 400 were the most mature and tenured in the study. The maturity evaluation was significantly depending on the respondents span of control, the greater the control the higher the maturity rating. Supervisors in company 400 see the culture borderline in stage 2 ‘react’ whereas those with larger span of control were firmly
in stage 3 ‘know.’ This was directly opposite to companies 100 and 300 with supervisors rating maturity significantly higher. No significant difference was picked up in company 600. Does this indicate that culture homogeneity increase with increased maturity? And must a culture evaluation system differentiate between working groups e.g., supervisors to get a result that can be actioned and lead to financial gain? Does this suggest that a uniform evaluation system that makes use of the same questions for all employees will provide a measure of culture that is difficult to action causing follow up actions to fail on changes promised?

There are some limitations in the estimation of COPQ. The use of Crosby’s guideline for calculating cost of poor quality is aligned with that of other methods for assessing impact of culture on financial performance but must be validated through empirical research specific to the food industry; cost structure, market dynamics, and global differences. It is suggested that more research is conducted to compare food safety culture maturity across countries and sectors of the food supply chain e.g., growers and retailers. Future work will include a modified Delphi research method with practitioners and scientists to finalize the validation of the findings and to further incorporate learnings into each company’s effort to strengthen its food safety culture.

The research adds significant value to the current discussion in the food industry regarding culture. The findings enable any food company to quantify food safety maturity and assigns a financial value to its current stage but more importantly the potential financial gain by improving status quo. This is a novel connection that with further confirmatory studies can become that external motivator for leadership teams to invest in maturing culture and gaining financially from their efforts.
7. Discussion of findings

Data were collected on existing evaluation models and from five multi-national food companies to investigate the application of a theoretical framework, the maturity model, and the proposed triangulated evaluation system. Overall it was established that the proposed evaluation system was valid and reliable and that content deduced through evaluation of existing methods and the method triangulation can be added to the system to strengthen it further. Details of these findings will be discussed in the following sections.

7.1. Comparative analysis of existing food safety culture evaluating systems

Several intriguing results were discovered related to the research hypotheses. The food safety culture maturity model and evaluation system does add additional information in comparison to other evaluation systems as it enables food manufacturers to profile and predict strength of food safety culture over time. A comparative analysis of five scientifically-based and three commercial culture evaluation models, i.e., Ball, CEB, Denison, De Boeck, Jespersen, NSF, TSI, and Wright, was completed. 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 on research methodology (Shavelson & 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 few of the research quality principles. 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 a 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 this most comprehensively, for example through discussion of focus group discoveries and missing evidence around impact on culture assessment through organizational levels. Validation and reliability measures are important for the validity of any type of research (Cohen et al., 2007; Robson, 2011). Each culture evaluation system was analysed for the validation strategy applied. Most applied external validation through population, ecology, and peer review. Only the Ball and Denison models document predictive validity. Ball constructed a predictive model based on the Reason Action model and Denison showed a predictive relationship between strength of culture assessment and existing financial and organizational performance data. Few models documented reliability measures and this is a considerable gap. Again, Ball and Denison did incorporate reliability measures into their models and documented the method chosen to do so in detail. Content of each evaluation system was compared through content analysis. It was found that almost all culture evaluation systems contained some content related to an organization’s values and mission. The Wright model dedicates six of eight elements to this dimension. All systems covered content related to the ‘People system’ dimension. Five of the eight cultural evaluation systems covered ‘Consistency’ and Jespersen covered this in three of five dimensions. ‘Adaptability’ was covered by four systems and the Ball system dedicated two of six constructs to this dimension. ‘Risk awareness’ was only covered by the De Boeck and Wright models (Table 27).
Table 27: Culture evaluation systems comparative analysis.

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

Ledger: red = values and mission, green = people systems, blue = consistency, yellow = adaptability, and purple = risk and hazards.

The content findings suggest that the proposed five dimensions cover all of what each system independently covers and identify some areas that could be strengthened in some systems e.g., adaptability, consistency, and risk awareness (Table 27). By reviewing the detailed global, organizing, and basic themes and by looking at all content from all systems, a picture of not simply organizational culture but food safety culture emerges. It was concluded that the five dimensions could be used to unify research in the food safety culture domain and provide each system owner with input into the continuous improvement of all systems (Figure 18).
Figure 18: Food safety culture theoretical framework.

The analysis of NRC compliance rate and validation strategy provided 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 can make reliable 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. This is of great practical significance. Organizations invest in culture evaluations, like investments made in financial audits or market analyses, to make strategic decisions for the future of the organization. As such, if these decisions are made on questionable information, the investments in decisions and change can be lost. This in turn can influence the industry perception of the value of culture evaluations and potentially add to the fragmented perception of culture, safety culture, and organizational effectiveness. Also, few systems made use of planned triangulation, a method commonly applied to qualitative research (Denzin, 2012), 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 documented definitions it could guide the development of content and truly deliver on both climate and culture evaluation systems. It was also unexpected to find 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. The ‘Risk awareness’ dimension requires more discussion as to whether it should be defined as a cultural dimension or if it should remain part of the description of well implemented, effective food safety management systems. As such, if ‘Risk awareness’ is a dimension of food safety culture, what are the assumptions that are strong and mature specific to risks? How do these differ from assumptions described in the other four dimensions? Is it that ‘Risk awareness’ is the product of outrage and not a rational, strategic analysis by senior leaders (retrieved from www.psandman.com)? If this is the case, should ‘Risk awareness’ then be integrated in the ‘Values and mission’ dimension as a separate ‘value’ described through the maturity progression as risk perception and awareness norms?

The definition of food safety culture and that of organizational culture suggest that culture is learned and shared among people (Griffith et al., 2010a; Schein, 2004). Culture is based on accepted assumptions, values, and beliefs, is dynamic and impacted by an array of attributes and actions. 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 to implement them; and subsequently if consumers are put in harm’s 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 (Griffith et al., 2010b) and both scientists and practitioners must hold each other to a high standard to
minimize this risk. It is a concern that evaluation systems are marketed, portraying validity through well-designed sales tools, without publically available details. Practitioners who invest in such tools must take responsibility and demand such information and thereby actively push the food safety domain to higher performance based on valid and reliable science.

7.2. Triangulation and the importance of establishing valid methods for food safety culture evaluations

Triangulation of methods, in contrast to a single method, is required to gather data to provide food manufacturers with a comprehensive measure of food safety culture strength and an insight into the variables that drive food safety behaviours across plants. Capturing social desirability in food safety and evaluating impact on food safety culture evaluations was suggested to add validity to food safety culture evaluation results, these hypotheses were tested through the collection of data from 21 food processing plants.

7.2.1. Scale validity and predictive attributes of food safety culture and climate.

A scale was developed by Jespersen et al (2016) to evaluate evaluating food safety culture. The scale was distributed in this study to 1,273 salaried employees within the participating organization and a factor analysis and hierarchical regression analysis conducted to examine scale validity (Chapter 4). The results showed that the proposed scale, defined by six sub-scales, was reliable and valid to evaluate maturity of food safety culture. The valid scale was designed specifically to differentiate between functions (i.e., manufacturing, FSQ, and other) and individuals span of control (i.e., leaders and supervisors). Statistical results confirmed the scale’s discriminant validity. The degree of factor variance explained by the sub-scales ranged from 53.05 – 72.10%. A hierarchical regression analysis suggests that overall maturity can be predicted by the ‘Values and
mission’ dimension in combination with the ‘People system’ dimension while controlling for the participants ‘tenure in role’. This shows the scales predictive validity and also the importance of the ‘Values and mission’ dimension similar to the findings suggested by Denison and Kotter in their work on organizational culture (Denison, 1997; Kotter, 1992). A slightly weaker model was found predicting overall maturity by functions and roles. This confirms that importance of macro-cultures and working groups (e.g., ‘manufacturing’ vs. ‘FSQ’ and ‘leaders’ vs. ‘supervisors’) as highlighted by both Ball and Schein (Schein & Schein, 2017). A model was proposed based on detailed findings derived through predictive studies completed by four researchers; Ball, De Boeck, Hinzs, and Jespersen. Each researcher has led work resulting in the identification of individual factors predicting food safety performance. Consolidating the results of these studies resulted in the development of a model for food safety climate and culture (Figure 19).
The model was derived by segmenting individual factors into three levels; organizational, working group, and individual. By focusing on this collective set of factors future researchers can continue the positive trajectory for the food safety culture domain to better understand what cultural attributes contribute to improved food safety performance.

### 7.2.2. Socially desirable responding.

Many of the existing evaluation systems make use of scales in surveys to evaluated food safety culture. Inherent in this data collection method is the potential for response bias. This can be evaluated and controlled for by use of a scale that captures social desirability. By
applying this validated scale, the research sought to capture the degree of social desirability and interpret these results in the context of food safety culture. This was done by investigating the applicability of an established ecological scale for measuring social desirability to the domain of food safety. The results in Chapter 3 showed that the validated scale for quantifying social desirability in the ecological domain, when modified to fit the food safety domain, can provide a valid measure of social desirability in food safety or, said differently, capture the degree to which employees ‘walk the food safety talk’. The FSDSR scale was found to be both reliable and valid using data collected from the five companies. Through factor extraction it was shown that 61.8% of the variation in the social desirability score across the participants could be explained through three social desirability components; self-deception – assertion of positives, image management, and self-deception – denial of negatives. The proposed FSDSR scale consists of 14 items and these three subscales. Respondents were asked to indicate level of agreement with each statement on a five-point answer scale. These three indicators of social desirability are in line with social desirability theory and previous empirical research on social desirability in other domains. The data and analysis showed that social desirability can be captured and evaluated in the food industry by the 14 statements of the FSDSR scale and thereby control for this potential response bias. The FSDSR scale is an integral part of the self-assessment scale and adds both scientific and practical significance. Controlling for social desirable responding can potentially reduce the overall maturity results as a highly desirable response (e.g., a score of 2.0) influences the triangulation results. As such, a self-assessed score of 4.5 is reduced to 3.25 when controlling for social desirability. This result moves an organizations culture evaluation from Stage 4 – Predict’ to ‘Stage 3 – Know.’
7.2.3. **Triangulating food safety culture evaluations.**

Assessing social desirability is one means to increase food safety culture evaluation validity. Another is the use of the method of triangulation. As described in Chapter 5, data from multiple sources were collected and evaluation results from each plotted on the food safety culture maturity model. Data were analysed for inter-coder consistency and construct validity, and capability of discrimination within a food safety culture maturity profiling system. Data from three methods were analysed; 816 self-assessment survey responses, codes from 126 performance documents, and codes from 42 interviews conducted on-site at each of the 21 plants (Figure 20).

![Method Triangulation Diagram](image)

**Figure 20**: Food safety culture evaluation – method triangulation.

The dispersion between the mean results per method per plant confirms the need to apply triangulation to get a comprehensive 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 – React’ is largely about creating a personal connection and building a strong foundation of “why food safety is important to you?” A program in ‘Stage 3 -Know’ 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 not seen as valuable to business results. This is an important finding as it has consequences for the trustworthiness of both evaluations and subsequent investments made in tactics to mature food safety culture. The results showed that mean maturity for all plants was generally higher when assessed through the self-assessment survey; ranging from 3.06 – 3.18. The results from the semi-structured interviews were closer to the self-assessment survey 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 suggested by responses of manufacturing leaders. The findings from the two functions were found to be significantly different both in maturity self-assessments and semi-structure interview codes. Mean maturity scores derived from performance document codes were the lowest of the three measures except for one plant. In general, more action content, e.g., tasks, follow up actions, was captured in the performance document codes and this was to be expected given the original purpose of the documents e.g., meeting minutes and inspection reports.

A coding framework was deduced by analysing results from the literature and the theoretical framework and applied to derive data using content analysis. The framework was consistently applied by two researchers with 90% agreement except for two dimensions ‘Values and mission’ and ‘Risk awareness’. 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 ‘Risk awareness’ 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 awareness’
is, in fact, a stand-alone dimension. Is content related to ‘hazards’ and ‘CCPs’ relevant for
evaluating culture? Or, as mentioned earlier, is it part of a related food safety management
system that interacts with food safety culture? Because of this issue and the fact that only two
systems (Boeck et al., 2015; Wright) have separated out ‘Risk awareness’ as a stand-alone
dimension, it is worth discussing if this dimension should remain in the food safety culture
theoretical framework or best considered in the evaluation as part of the other dimensions.

7.3. A path to realizing value by maturity food safety culture

The method triangulation rendered a valid and reliable method for evaluating food
safety culture. To further refine the model and create a stronger connection to organizational
values the maturity model was updated to reflect the behavioural content in the methods in
the form of norms statements. The details of the validated triangulation method provide the
detailed evaluation method and content and the revised maturity model identifies the way
forward. As such, the revised maturity model is found valid and reliable as an evaluation
model when combined with method triangulation. These details, as outlined in Chapters 3, 4,
and 5 support the concept that the food safety culture maturity model and evaluation system
is valid and reliable and adds additional information in comparison to other applied methods
as described in Chapter 2.
7.4. Evaluating and predicting effectiveness and maturity of food safety culture in food manufacturing

Organizational culture as defined by Schein (2004) is about adapting and integrating change. As such, this research has showed a valid connection between dimensions in a theoretical framework to elements already proven to predict food safety performance to elements already shown to drive organizational effectiveness. We can therefore say that a valid culture evaluation system is one that evaluates the actions between organizational culture effectiveness and organizational culture building blocks e.g., norms and values; working group macro-culture and individual intent and behaviours (Figure 21). It is not a system that checks the content of each building block (e.g., is there a learning program?) but the actions between building blocks that act as evidence for the organization’s ability to adapt and integrate change.
Figure 21: Organizational effectiveness with cultural building blocks and actions.

‘Actions’ i.e., adapt and integrate

Culture building blocks
The evaluation system developed through this research (i.e., self-assessment scale including social desirability, triangulation method, and maturity model) offers organizations and academia alike an improved starting point for future food safety culture research and improvements. As such, data related to the adaption of organizational culture to its external environment (Figure 21 – arrow 1) and the reflection of values in the formal systems (Figure 21 – arrow 2) through on-site visits and the performance documents. Adaptation and integration of values in the organizational culture (Figure 21 - arrow 3), through on-site interviews and the adaptation of norms to the formal systems through the performance document content analysis (Figure 21 - arrow 4). Adaptation of the working groups assumptions to the organizational culture (Figure 21 - arrow 5) through the scale and the on-site interviews. The individuals’ integration in the working group (Figure 21 - arrow 6) and the organizational culture (Figure 21 - arrow 7) through the self-assessment scale, including social desirability assessment. This shows that reliance on a single method will not provide a comprehensive evaluation of culture; for reason of method weaknesses as discussed in Chapter five, but also due to the inherent complexity of building blocks and actions taking place in any organization.

This study has several implications for professional practice in both academia and industry. Notably, the study’s developments offer insights about future education and training of food science students and researchers. Academicians in the food science field should continue their involvement and investment in established organizations e.g., GFSI, SSAFE, Salus: The Food Safety Culture Science Group, to continue the positive trajectory of establishing a food safety culture research tradition. Practitioners must learn to better acknowledge proven principles of science in the food safety culture domain and seek to understand food safety culture as critical to the survival of consumers and business alike and
not choose the easiest solution if it is not fit for the purpose. The researcher calls for practitioners to be more critical of the owners of evaluation systems to ensure their return on this investment and to push those that develop and profit from such tools to be held to higher performance standards. Quoting a global food company leader “Doing the right thing isn’t always easy. But it is always right.”

7.5. Limitations of this study and recommendations for future research.

As with all research, this study has limitations. The study was based on data collected from five multi-national companies and a sub-set of 21 North American plants. Although this was a mix of small and large plants across multiple product categories, the study’s validity could have been increased by including plants outside North America. It is suggested that more research be conducted to compare food safety culture maturity across countries and sectors of the food supply chain e.g., growers and retailers.

As this research might be biased or influenced by human factors such as employees’ food safety motivation and their willingness to participate and fill in self-assessment tools ((Boeck et al., 2015; Clayton et al., 2002; Guldenmund, 2007). It is proposed that employees’ tendency to respond in a social desirable way regarding food safety issues should be captured in future food safety culture surveys. The FSDSR allows quantification of this tendency by estimating the extent to which employees deceive themselves by asserting positive behaviours and denying negative behaviours on the one hand and managing their image on the other hand. An opportunity exists to apply the scale among other stakeholder and groups e.g., associates/operatives, and inspectors to compare social desirability between groups. The researchers also recommend further psychometrical research to explore the scale’s predictive validity.
It was unexpected that the sub-scales for food safety and quality supervisors was not valid and reliable. This is a significant result and suggests that the pinpointed behaviours identified for this group are not accurate. It is possible that this group is a composite of individual roles with different spans of control (e.g., a supervisor in one plant might be called a manager in another). If so, this calls for a revision of the scale’s demographic question related to ‘role.’ Targeted research is needed to better understand this working group e.g., area and span of responsibility, degree of standardization in job expectations, and tasks.

Further, it was unexpected that such similarity was found between the five plants in the triangulation analysis (Chapter 5). Maturity ranged from ‘Stage 2 – React’ to ‘Stage 3 – Know’ for all plants. This could be due to the geographical dispersion of the plants, this subset all being in North America, and therefore under similar legal systems and customer expectations. It could also be a case of selection bias as the participating companies were not randomly identified nor were they subject to quasi-random assignment, rather through senior leader interest and board willingness to participate in the research. It is reasonable to assume some sampling bias due to the voluntary nature of the participants.

8. Conclusion

In conclusion, “Organizational culture is not a soft concept, its impact on profit can be measured and quantified” and, as a macro-culture, food safety is an integral part of any food company’s organizational effectiveness and culture. As food companies recognize the strategic importance of their food safety culture, the validity of evaluation systems and methods will gain greater importance.

This research contributes significantly to the food safety culture research domain as it revealed a model that can be used by practitioners and academicians alike to further food safety performance. When subjected to statistical analysis, a scale for evaluating food safety
culture maturity was confirmed and it is suggested that a combination of attributes discovered independently by four researchers are integrated and explored through multi-geographical empirical research. 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 of 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.

Knowing the strength of your food safety culture is crucial to your company’s food safety performance, your organization’s effectiveness and, ultimately, its survival. As illustrated in the introduction, direct links between food safety failures and organizational culture have been made and we must move beyond the discussion of causation into a discussion of creating and sustaining strong cultures. The evaluation system presented in this research is a substantial leap forward towards this new discussion for scientists and practitioners alike.
9. References


doi:10.1177/1757913914532620

Jespersen, L., T., M., & Vlerick, P. (In press). Development and validation of a scale to capture social desirability in food safety *Food Control.*


Loewenthal, K. M. (2004). *An introduction to psychological tests and scales*


Wright, M. A tool to diagnose culture in food business operators, Food Standards Agency Research Report, 2013.


