How Managers in Poland Use the Principles and Instruments of the Kaizen Philosophy in Their Personal Lives – The Personal Kaizen Approach

DOI: 10.12776/QIP.V24I2.1396

Iwona Burka

Received: 2020-02-11 Accepted: 2020-06-24 Published: 2020-07-31

ABSTRACT

Purpose: The aim of this paper is to show whether and how Polish managers working in the Kaizen/Lean environment, benefit from the knowledge gained during their professional work to develop competences in their personal lives and improve personal productivity. Do they transfer the knowledge gained about lean systems into their private lives and what are the benefits for them?

Methodology/Approach: To answer the research questions raised, the authoress conducted a literature review and a qualitative research – multiple-case studies. As part of the qualitative study, five in-depth interviews were made.

Findings: The literature review and the qualitative research carried out allowed to formulate conclusions, including, i.a., the following: managers transform the knowledge acquired during the training conducted as part of their professional work into their private lives; it is possible to distinguish a number of principles and instruments of the studied concepts that are applied to support the process of developing personal professional and non-professional competences.

Research Limitation/implication: Subjectivity of the analysis, owing to the application of a qualitative approach in the studies.

Originality/Value of paper: The literature review points to a small number of scientific publications connected with the implementation of the Kaizen approach and Lean Management principles and instruments in the process of developing personal competences and improving personal productivity. This publication is the first approach to creating a Personal Kaizen Toolbox.

Category: Research paper

Keywords: Kaizen; personal Kaizen; Lean Management; waste; personal improvement

1 INTRODUCTION

"Success in the knowledge economy comes to those who know themselves – their strengths, their values, and how they best perform." (Drucker, 2005, p.100)

We are living in an economy that is based on data, information, and knowledge; therefore, we must constantly learn to be able to develop, improve our work, and improve the quality of life. Garvin asked: How can you improve yourself without finding out something new? How to solve problems or improve processes without trying to look at the matter from a different, new perspective? (Garvin, 1993). That is why managers try to constantly develop their competences by designing their own personal development systems, or basing on solutions taken from the literature or training.

One of the approaches as regards Continuous Personal Improvement is Personal Kaizen. The term Personal Kaizen means constantly improving the quality of your personal life and well-being, by means of introducing a series of small, incremental changes (Suarez-Barraza, Ramis-Pujol and Dahlgaard-Park, 2013). We can also define Personal Kaizen as a set of principles and instruments for personal, continuous improvement in order to develop your own competences.

Kaizen is a philosophy of continuous development, improvement, and amelioration and the name derived from the Japanese words: kai ('change') and zen ('good'). Continuous, by definition, means constant, continuing without interruption, and constantly repeating. Thus, Kaizen is an endless process of correction, implemented in little steps (Imai, 2006; 2007; Emiliani, 1998a; Locher, 2012; Liker and Meier, 2011; Maurer, 2007; 2016; Womack and Jones, 2008). Everything can be improved, and improvement is the responsibility of each employee towards the organization (Stoller, 2015, p.65), and each person in relation to each other. Bicheno and Holweg (2009) describe Kaizen on the one hand as a philosophy, and on the other, as a tool set. Kaizen, according to Liker and Hoseus (2009, p.410), is one of the five values of a caring leader. Liker (2005) placed Kaizen among the fourteen management principles which he called the Toyota Way, and Graban (2011) emphasises that all of us (not only leaders) are obliged to constantly look for ways to improve our work. According to Womack and Jones (2008, p.541), the aim of Kaizen is to create more value with less waste.

The assumptions of the Kaizen philosophy may be applied in both the improvement of processes, structures, systems in organisations, but also in our personal life. They can serve the improvement of the quality of life (Suarez-Barraza, Ramis-Pujol and Dahlgaard-Park, 2013) and effectiveness of our actions. Although the improvement in Kaizen is implemented in small steps, it brings huge and lasting results in the long term (Imai, 2006). Maurer (2007, p.541) quotes the words of Wooden, an American basketball player and coach: "When you improve a little each day, eventually big things occur. When you improve conditioning a little each day, eventually you have a big improvement in conditioning. Not tomorrow, not the next day, but eventually a big gain is made.

Don't look for the big, quick improvement. Seek the small improvement one day at a time. That's the only way it happens – and when it happens, it lasts". Confucius, on the other hand, is credited with words that aptly reflect the meaning of Kaizen's philosophy: "It does not matter how slowly you go as long as you do not stop".

The Kaizen philosophy is based to a large extent on the development of the organisation, which is based on the hidden knowledge of employees, knowledge acquired in their non-professional lives. When planning a training and employee development system, and a motivating system to generate improvements, it is worth paying attention to the passions and interests of the employees. The greatest scientific discoveries of the recent decades have been the result of combining knowledge from different, often seemingly unrelated scientific fields. An appropriate approach to planning a training and employee development system influences the competitiveness of the organization and is responsible for its market success (Roberts and McDonald 1995, p.16), and the broader knowledge and interesting interests the employees have, the more efficiently they solve problems and create new solutions.

Kaizen is a philosophy of development, improvement, but it is also a set of tools (Bicheno and Holweg, 2009). The authoress in her considerations has assumed that since organisations that develop their systems, structures, and processes in accordance with Kaizen assumptions, use the so-called Lean Toolbox to identify and minimise waste, the set of principles and tools proposed in the literature may also be used on the path to develop one's own competences and improve the quality of life. Personal Kaizen is a type of individual Kaizen, oriented on the individual (Imai, 2007, p.136), the goal of which is a gradual but continuous improvement of one's productivity and quality of life, by minimising waste, simplifying, applying standards, taking care of one's own work environment, planning and supervising personal goals, learning from mistakes, but above all being familiar with one's own value system. It is a kind of an individual suggestion system that streamlines the process of developing the competencies of a given person, it is a way of life, as Imai (2006) used to say. A human that adheres to the principles of Personal Kaizen - 'Kaizen person', is someone who can change their way of acting at work and beyond it to develop in a stable and lasting way (Suarez-Barraza, Ramis-Pujol and Dahlgaard-Park, 2013).

In her work, the authoress has raised the following research questions:

RQ1. How and in what way do Polish managers use the knowledge gained within their professional work on Kaizen/Lean Management to develop competences in their personal lives and improve their personal productivity?

RQ2. What benefits do Polish managers achieve from implementing Personal Kaizen in their personal lives?

The authoress has carried out a literature analysis covering selected principles and instruments of the Kaizen approach and the Lean Management concept, in the context of their application as part of improving personal competences and personal development. This analysis was the basis for preparing a questionnaire for in-depth interviews. The authoress has conducted five in-depth interviews with managers who have been working for at least five years in Polish organisations which operate in accordance with the Kaizen/Lean principles.

The results of the literature analysis are included in point 2, and the results of the case study in point 4 of this elaboration.

2 PERSONAL KAIZEN PRINCIPLES AND INSTRUMENTS – LITERATURE REVIEW

The motto of the organisations which operate in the Kaizen spirit and implement the principles and instruments of Lean Management is to work smarter, not harder. Lighter, more effective, more efficient and more pleasant work should be the goal of both the organisation and each of its members – the employees. Many years of practice in organization management has shown that everything can be improved and there is always potential for improvement (Stoller, 2015), however, the most difficult thing is to improve our own activities, consistently develop and constantly strive to improve the quality of our lives. According to the words of Plato, the first and greatest victory is to conquer yourself. Overcoming weaknesses regarding, for example, one's eating habits, is just as difficult as creating a habit of regular reading or learning foreign languages. Paradoxically, the more we have diet books on the market, the more obese people there are in the world (about growing obesity see Keating, Backholer and Peeters 2014). Markovitz (2011) writes that overcoming his own natural tendencies to generate complexity, mess, and chaos is something extremely difficult and requires something more than just a simple, obvious 'recipe' for effectiveness presented in the guides. From the times we were children, our parents had taught us to put things away, clean up after ourselves, do not buy what we do not need, and yet they did not call it 5S. We buy toilet paper, when we only have one roll left, and yet we do not call this last roll a kanban. Personal Kaizen is not just a set of obvious solutions, but it is also a change of approach to personal development, in which consistency in action is of great importance. The largest number of gym passes is sold in January, and in February few people remember about the New Year's resolution on improving physical fitness. A sudden, big upsurge usually ends in failure. Small steps seem to be an interesting alternative here.

To answer the research questions raised, the authoress began with a scientific literature query. To this end, she analysed scientific journals collected in the Emerald Publishing Scientific Database, in terms of selected principles and instruments of the Kaizen approach and the Lean Management concept, in the context of their application as part of improving personal competences and improving personal productivity. The key words according to which the database

was searched are shown in Figure 1 and Figure 2. The literature analysis was carried out in January 2019. As a result of the analysis carried out, as well as numerous interviews conducted with managers working in a Lean environment, focused on continuous improvement, the authoress prepared a list of principles and instruments that can be an element of the Personal Kaizen Toolbox.

The literature analysis and the empirical research results enable listing the following principles and instruments of the Personal Kaizen approach (Personal Kaizen Toolbox):

- continuous improvement in accordance with the Kaizen philosophy;
- knowledge about your own value systems;
- eliminating waste, overloading, and irregularities;
- determining (not necessarily written down) own standards and norms;
- taking care of your own workplace, home, in accordance with the 5S principles;
- eliminating complicatedness and unnecessary complexity - simplifying the work system and personal processes;
- the art of asking questions, including in-depth question asking (5 Why);
- planning and setting goals;
- designing an appropriate task management system and education system adapted to the predispositions and personal needs;
- using the benefits of visual management (including personal kanban);
- cooperation and teamwork;
- introduction of a systemic approach to problem solving, including the A3 Report and the Ishikawa diagram;
- learning from mistakes, taking care of feedback, reflection, and self-awareness;
- the use of an appropriate system of work breaks and methods of relaxation;
- stimulating and training creativity - because to change something, we have to look at the problem in a different way;
- proper preparation for learning and development (personal Training Within Industry);
- attention management and concentration training to improve effectiveness of actions related to personal development (Suarez-Barraza, Ramis-Pujol and Dahlgaard-Park, 2013; Markowitz, 2011; Imai 2006; Emiliani, 1998a; Maurer, 2007; 2016; Baccarani, Mascherpa and Minozzo, 2013).

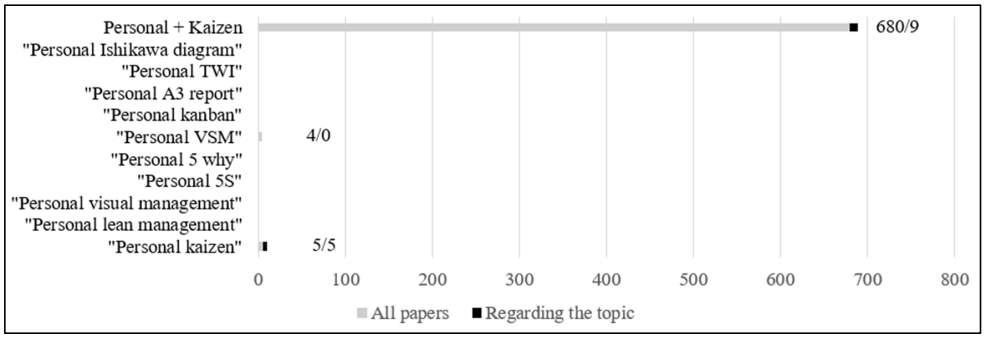

Figure 1 – Graphic Presentation of the Literature Analysis Results Including a List of Key Words

The literature query points to a shortage of scientific publications on the implementation of the Kaizen approach and Lean Management principles and instruments in the process of improving personal competences and improving personal productivity. The sparse publications focus on the process of personal improvement in personal life using the small steps of the Kaizen philosophy. There is even less information about research on the implementation of such instruments as: 5S, kanban, standardisation work, visual management, value stream analysis, in the process of human personal development. Those who studied the topics were, i.a. Suarez-Barraza, Ramis-Pujol and Dahlgaard-Park (2013) and Emiliani (1998a). The following publications are also recommended: Yoneyama (2007); Baccarani, Mascherpa and Minozzo (2013); Emiliani (1998b).

Taking the above into account, the authoress decided to conduct a qualitative case study among Polish managers, formulating the following research hypotheses:

H1. Polish managers working in the Kaizen/Lean environment, benefit from the knowledge on the principles and instruments of Kaizen/Lean Management gained during their professional work, to develop competences in their personal lives and improve personal productivity.

H2. Polish managers achieve specific benefits related to using the Personal Kaizen approach in their personal lives.

In point 4 of this publication, the authoress presents the results of a qualitative study conducted among Polish managers who work in the Kaizen/Lean environment.

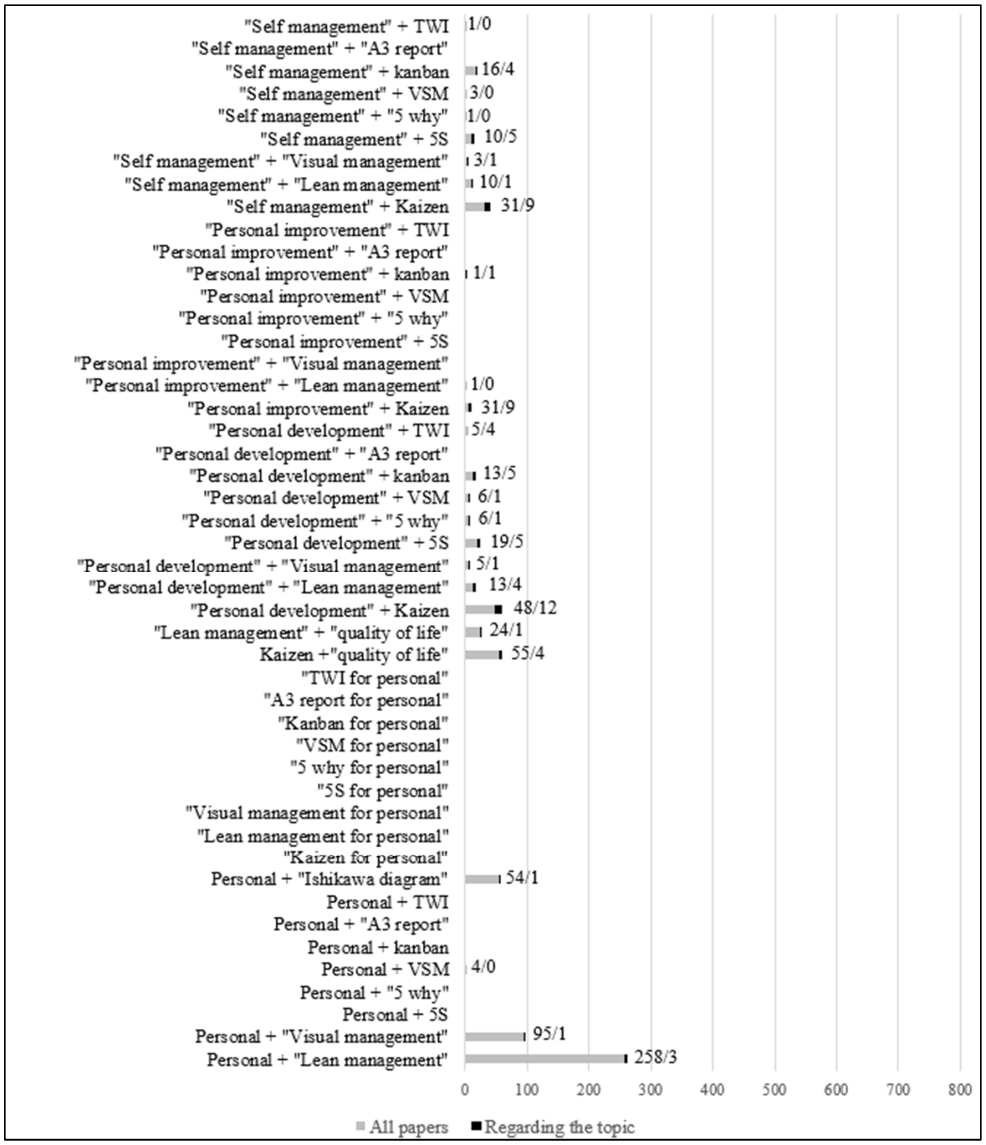

Figure 2 – Graphic Presentation of the Literature Analysis Results Including a List of Key Words

3 METHODOLOGY

The implementation of principles and instruments of Kaizen philosophy is more and more popular in Poland. A lot of training in this field is conducted in organisations. However, should the benefits of participating in training only translate into professional work?

In connection with the fact that the research questions start with the question: "How?" and "What?", the authoress decided on qualitative research (multiple-

case study) as the most appropriate research method. Thanks to the case study, it is possible to confront one's own reasoning with the behaviour of the actual participants of events and processes (Sławińska and Witczak, 2008, p.121). The case study's nature is empirical, it is reliable collection and processing of data that can be compared with each other (Yin, 1994, p.23). It is a research strategy that enables comparing cases between one another.

In order to answer the research questions, the authoress conducted literature research and empirical research – in-depth interviews. The idea is to get detailed opinions and information from the respondents in accordance with the logic of analytical, literal replication.

Information about the in-depth interview:

- the method of an in-depth, semi structured interview method was applied;
- the study consisted of two stages: Stage 1 – answering open-end questions contained in the interview questionnaire; Stage 2 – participants' own conclusions, including the indication of the benefits resulting from the implementation of selected Personal Kaizen principles and instruments in their lives and the assessment of the quality and effectiveness of the personal development process before getting acquainted with Kaizen/Lean and after completing vocational training and acquiring the necessary knowledge about Kaizen/Lean;
- the in-depth interview questionnaire was developed on the basis of a literature analysis and observations carried out by the authoress in a business environment; the questionnaire consisted of 30 open-ended questions divided into three groups;
- 5 participants (Polish managers) took part in the study – 3 women and 2 men;
- participants of the study have been working for at least five years in Polish organisations which operate in line with the principles of Kaizen/Lean and they have participated in many training sessions in the area of Kaizen/Lean;
- information about the possibility of participating in the survey was posted on LinkedIn in March 2019, and the survey was conducted at the turn of April and May 2019 (4 telephone interviews, 1 video call using an Internet connection);
- the interviews lasted 1.5 hours on average;
- 18 people expressed their willingness to participate in the study, from among whom, in line with the assumptions, the authoress chose five people who first reported, agreed to carry out a remote conversation with them, as part of a case study, and were available in the period scheduled for conducting the study.

It is possible that the willingness to participate in the study was expressed only by people who are aware of the importance of implementing the principles and instruments of Personal Kaizen in their personal lives and of the benefits resulting from it. People who do not apply the analysed solutions in their private lives may have deliberately not wanted to participate in the study. In the context of the research questions, however, what was important was to gain knowledge on how to transform Kaizen/Lean principles and instruments to improve the competences of the managers and improve their personal productivity as well as the benefits from implementing Personal Kaizen. The data and information obtained during the study were analysed and they constituted the basis for the formulation of the conclusions.

The purpose of this research is to create initial application assumptions and to pave the way for subsequent, wider research.

4 HOW MANAGERS USE PERSONAL KAIZEN TO DEVELOP THEIR OWN COMPETENCES AND IMPROVE THEIR PERSONAL PRODUCTIVITY – CASE STUDY

The interview questionnaires contained the following groups of issues:

Group 1. Introduction to personal development and Personal Kaizen

- the importance of personal development for the interviewee;
- contact with Kaizen/Lean principles and instruments during vocational training;
- the interviewee's approach to personal development (small steps or large, one-time upsurges) ;
- the private life sphere in which the interviewee applies the Personal Kaizen;
- diagnosis of one's strengths and weaknesses, competences, to take further steps connected with self-development.

Group 2. Personal Kaizen principles and instruments

- defining values in the life of the interviewee and applying the concept of flow and analysis of personal value streams;
- the importance of the philosophy of minimalism, limiting resources and simplifying in the interviewee's life;
- identifying sources of waste in non-professional life;
- setting measurable goals, supervising them, and their consistent implementation;
- management of personal tasks (methods and instruments);

- the importance and application of visual management and the role of standardised work in personal life;
- application of the 5S method (5 principles of workplace management) in non-professional life – teamwork, having a mentor with whom you could consult your own development, improvement;
- other Personal Kaizen principles and instruments not mentioned previously applied by the interviewee.

Group 3. Support for self-development

- developing habits;
- the importance of reflection (hansei), self-awareness, feedback, on the path to personal improvement;
- the importance of respect for people, active listening, suspending personal views;
- the importance of relaxation and attention management in a continuous path to excellence;
- approach to life, attitude (positive, negative) and its impact on the process of personal development.

The results of case studying have been presented in Table 1 to Table 5.

Table 1 – Case 1 (Source: Author)

Problem group	**Case study** *woman, 29 years old, manager, experience with Kaizen/Lean of approx. 9 years*
Introduction to personal development and Personal Kaizen	• personal development is very important and implemented in Kaizen small steps, but there are also radical changes (kaikaku), • analysis of return on investment in self-improvement for professional development, • areas of Personal Kaizen application: housework, creating habits, continuous improvement, maintaining order, taking care of the garden and developing its richness, taking care of own physical condition (regularity), getting rid of bad habits (e.g. using social networks), • analysis of own strengths and weaknesses (mainly psychological tests), • planning own personal development based on analysis results, • training connected with personal development dedicated to managers: international communication, interpersonal communication, and psychology.
Personal Kaizen principles and instruments	The participant of the study: • defines the values in her life (the most important are: family, love, self-fulfilment, and opportunity for development), • tries to carry out her work in accordance with the principle of flow and analyses repetitive value streams of values in her life (reflections, without putting them down), • applies a minimalistic approach to a limited extent, among others in relation to the clothes she has, mobile phone apps, desktop on the computer, data and

Problem group	**Case study** *woman, 29 years old, manager, experience with Kaizen/Lean of approx. 9 years*
	information, food, and shopping, • considers simplifying as very important in her life, it translates into time saving, • identifies sources of waste in private life; the main sources are: distractions (wasting time), food, clothing, • defines goals, but does not put them down (tries to remember them), which causes problems with supervision and consistent implementation, • uses kanban in the form of an electronic app, • applies visual management; examples: kanban for washing, kanban for fireplace wood, kanban associated with refuelling (sense of security), yellow and colourful post-it notes, • uses standardised work; examples: cooking (recipes written down in detail and according to a set rule), and instructions for the family, • uses 5S in: the dressing room, kitchen, garage, office, garden, computer, • as part of teamwork, she occasionally seeks advice of friendly business psychologists, conducting creative discussions with them, and she has sought advice of a career counsellor, • uses other principles and instruments: 5 Why, cause-and-effect diagram, and risk analysis.
Supporting self-development	The participant of the study: • develops habits (very important on the way to improvement), • tries to learn from mistakes; reflection and feedback, especially after public speeches, are very important for her, • is a patient person, for whom active listening is very important, without comments; she admits, however, that in professional life she has no problems with it, while privately she is constantly learning this, noticing the potential for improvement; believes that this approach has changed for better the relationships in her private life, • does not use relaxation techniques, but tries to avoid distractions, • is a person who has a positive attitude to the world around her and respects people.
Personal development before Personal Kaizen	• chaos and undefined personal development, • lack of possibility of supervision and the problem of consistent action, • worse quality of life and effectiveness of actions.
Personal development in the Personal Kaizen spirit	• effective personal development, efficient action and lack of discouragement, • good habits developed (e.g. regarding physical condition and health), • order, there is no need to search for things, • effective task management, • improvement in the sphere of self-discipline, assertiveness, self-control, • change of approach to personal development and lasting effects, • improvement of the quality of personal life, free time, improvement of health.

Table 2 – Case 2 (Source: Author)

Problem group	**Case study** *woman, 44 years old, manager, experience with Kaizen/Lean of approx. 7 years*
Introduction to personal development and Personal Kaizen	• personal development is very important and implemented in small Kaizen steps, but mainly in the case of children's development, • areas of applying Personal Kaizen: housework, bringing up children; planned areas: sport, • periodic analysis of own strengths and weaknesses, • training related to personal development dedicated to managers: problem solving and post-graduate studies.
Personal Kaizen principles and instruments	The participant of the study: • has not defined the values in her life, she does not apply the principles of flow and has not analysed repetitive value streams in her own life, • applies a minimalistic approach to, among others, towards resources available, and simplification is very important in her life, • identifies sources of waste in her private life; the main source is wasting time; draws attention to environmental waste (decision to limit the purchase of products in plastic packaging), • defines and puts down measurable goals, oversees and consistently implements them, • uses an electronic calendar and a planner, common for all family members, to manage tasks, • applies visual management; examples: planner, calendar, visual instructions for the children (as part of standardised work), visual marking of cabinets in the children's room, • applies 5S in: the dressing room, kitchen, children's room; she sorts things mainly before Christmas, • occasionally seeks advice as part of teamwork.
Supporting self-development	The participant of the study: • develops habits, mainly food-related ones, but also related to learning languages (uses reminder apps), • tries to learn from mistakes; reflection and feedback are very important, • active listening is important to her; she believes that this approach has changed relationships in her private life for the better, • does not regularly use relaxation techniques, time spent traveling to work and home is time for making reflections, • is a person with a positive attitude towards the world and respects people, • works actively on developing creativity.
Personal development before Personal Kaizen	• chaos, • unspecified personal development, • greater garbage production, • worse quality of life and effectiveness of actions.
Personal development in the Personal Kaizen spirit	• self-consciousness and an organised life, 5S in resources, • systemic approach to personal development and goal and task management, • this period is for her "another universe", • permanent effects, • improving the quality of personal life and more free time.

Table 3 – Case 3 (Source: Author)

Problem group	**Case study** *woman, 33 years old, manager, experience with Kaizen/Lean of approx. 5 years*
Introduction to personal development and Personal Kaizen	• personal development is very important, but it is mainly based on large-scale upsurges, • areas of Personal Kaizen application: housework, bringing up children; eliminating unnecessary activities through modernisation of resources (automatic clothes dryer and dishwasher), goals in the spirit of *kata* (Toyota Kata), • no diagnosis of strengths and weaknesses and additional training.
Personal Kaizen principles and instruments	The participant of the study: • has defined values in her life, applies the principles of flow and has analysed repetitive value streams in her own life (without putting them down), • applies a minimalistic approach to: toys, equipment, crockery, food, friends with a negative attitude to life, • considers simplification as very important in her life, • identifies sources of waste in her private life; the main ones are: creating garbage, wasting time, doing the washing (after putting on the clothes once they are immediately put away to be washed), • defines and records measurable goals in the *kata* spirit; a slight problem with the consistent pursuit of their implementation; the main goal is to maximise the time spent with the family, • uses an electronic calendar to manage tasks, • applies visual management; examples: visual instructions for her child (as part of standardised work), • uses 5S in: the children's room, dressing room, documents, • as part of teamwork, she occasionally seeks advice of her brother, who at the same time has a motivating impact on her.
Supporting self-development	The participant of the study: • has a problem with developing habits, • tries to learn from her mistakes; reflection and feedback are very important for her (she herself first gives positive and then negative feedback), • active listening is important to her, but it is something with room for improvement; however, she believes that the change in approach has impacted improvement of relationships in her private life, • does not regularly apply relaxation techniques, • is a person with a rather positive attitude towards the world and she respects people, believes that it is of great importance for building positive and effective relationships, • actively works on developing her creativity.
Personal development before Personal Kaizen	• worse relationships, • chaos, • losses connected with generating waste, • worse quality of life and effectiveness of actions.
Personal development in the Personal Kaizen spirit	• positive attitude, • orderly life and resources, pursuing goals, • more lasting results, • less waste, • improvement of the quality of personal life and more free time.

Table 4 – Case 4 (Source: Author)

Problem group	**Case study:** *man, 42 years old, manager, experience with Kaizen/Lean of approx. 6 years*
Introduction to personal development and Personal Kaizen	• personal development is important, mainly in the Kaizen spirit, • areas of applying Personal Kaizen: organised learning, garage, dressing room, • no diagnosis of own strengths and weaknesses and no additional training.
Personal Kaizen principles and instruments	The participant of the study: • has defined values in his life, but he does not apply the flow principle and has not analysed repetitive streams of values, • applies a minimalistic approach to: shopping, • considers simplification to be very important in his life, • identifies sources of waste in his private life; the main sources are: food, clothes, children's toys, • defines, but does not put down his goals, problem with supervision, lack of consistency, • uses an electronic calendar to manage his tasks, • applies visual management; examples: yellow post-it notes, colour markers, • does not use standardized work, • applies 5S in: tools, clothes, documents, children's toys, • does not seek advice and works alone.
Supporting self-development	The participant of the study: • has a problem with developing habits, • tries to learn from his mistakes; • has not thought about reflection and feedback in his life, and active listening is a field for improvement, • does not regularly apply relaxation techniques, • is a person with a rather positive attitude towards the world and respects people, • he does not work on developing creativity; he has never thought about it.
Personal development before Personal Kaizen	• mess in documents, tools, always looking for something, • chaos, • losses associated with waste generation, • worse quality of life and effectiveness of activities.
Personal development in the Personal Kaizen spirit	• everything in its place ("I use instead of looking for it"), • more orderly life and striving for the implementation of life plans, • less waste, • improvement of the quality of personal life, more free time, • the interviewee, however, still has considerable potential for improvement.

Table 5 – Case 5 (Source: Author)

Problem group	**Case study** *man, 38 years old, manager, experience with Kaizen/Lean of approx. 12 years*
Introduction to personal development and Personal Kaizen	• personal development is very important and mainly consisting of small step improvement (Kaizen), • areas of Personal Kaizen application in personal life: everywhere, • the interviewee has diagnosed his strengths and weaknesses, • the interviewee has undergone additional training, mainly soft training.
Personal Kaizen principles and instruments	The participant of the study: • has defined values in his life, applies principles of flow in scientific development and implementation of tasks, he has also analysed the value streams associated with his development, • applies a minimalistic approach to: shopping, food, data, and information, considers simplification as something of great importance in his life, • identifies sources of waste in his private life; the main sources are: food, clothes, organised meetings, because it seems the thing to do, • defines and puts down the goals, supervises and consistently implements them, • uses a personal kanban and an electronic calendar to manage tasks, • applies visual management; examples: kanbans, pictorial markings, colour binders, shadow boards for tools, • uses standardized but informal work, • applies 5S in: information, tools, clothes, and documents, • seeks advice, but often works independently, • introduced the principles of Total Productive Maintenance (TPM) for household appliances and private cars as well as the Single Minute Exchange of Die (SMED) in connection with daily dressing up.
Supporting self-development	The participant of the study: • develops habits, the most difficult are regular, healthy meals, sports, • tries to learn from his mistakes; values reflection and feedback, • active listening is a field for improvement for him, • does not regularly use relaxation techniques (occasionally), but plans to change this, • is a person with a positive attitude towards the world and he respects people, • is not working on developing creativity, he has never given a thought to it, but he has been curious about it.
Personal development before Personal Kaizen	• disorder mainly in his mind, but also in resources, constantly looking for something, • losses related to the generation of waste, • problem with habits, • doing many things at the same time, • worse quality of life and effectiveness of activities.
Personal development in the Personal Kaizen spirit	• order, no need to look for things, • finally implementing his goals, • does not have to think about what else needs to be done by him, he has an effective system of managing tasks and noting down ideas, • less waste, • improving the quality of personal life and more free time.

Managers working in the Kaizen/Lean environment, participating in numerous training sessions covering both the principles and instruments of continuous improvement, simplification, and minimising resources, try to transfer the knowledge acquired during training to their personal lives. The participants of the study positively assessed the solutions taken from training and implemented in their personal lives. At the same time, they pointed out that this has significantly improved and organised their work on acquiring new knowledge and skills, and positively influenced the quality of life. Knowledge about reflection, respect for people, active listening made them realise how many mistakes they had been making in their relationship with their closest ones.

The results of the literature query carried out and the qualitative research conducted among Polish managers allow for formulating the following conclusions:

- managers transform the knowledge acquired during training undergone during their professional work into their private life;
- a number of principles and instruments of the Kaizen philosophy and Lean Management concepts can be distinguished, which find their application in supporting the process of developing professional and non-professional competences of a human being; we can define the set of such principles and instruments as the Personal Kaizen approach;
- managers most often transform the following principles and instruments into non-professional life: continuous improvement implemented in small steps, defining human-related values, identifying sources of waste, limiting the amount of resources held and processed, 5S, visual management (including kanban), work standardisation, and developing habits;
- the managers highlight that individual principles and instruments are often dependent on one another and the selection of one of them determines the necessity of applying the next one;
- the most frequently indicated sources of waste are: throwing away food, unnecessary purchases (clothes, equipment, children's toys); the most wasted resource, according to the managers, is time;
- the managers spend too little time on rest, do not use relaxation methods, attention management, improving concentration, and developing creativity;
- an interesting thing is the implementation of TPM principles to maintain good condition of devices and vehicles at home, as well as SMED principles for everyday changeovers, namely dressing up for work in the morning.

5 CONCLUSION

Living in a knowledge-based economy, we must constantly develop. Managers are aware of the fact that if they stop, they really begin to step back. In order for them to develop, they must find time for it, for example by improving their productivity. The goal of Lean Managers is to work smarter, not harder, and as a result, do scheduled work faster. Management sciences offer a whole spectrum of principles and instruments to improve the organisation, its structures, processes, and systems. By transforming them into our personal lives, we can improve our work, find time for development and rest, and thus improve the quality of our lives and develop our competences.

The analysis of literature and case studies carried out were aimed at obtaining answers to the research questions posed at the beginning and verifying the formulated hypotheses.

All managers participating in the study have declared transforming selected Kaizen/Lean principles and instruments into their private lives. It follows that there are Polish managers working in the Kaizen/Lean environment, benefiting from Lean Management systems and choosing an improvement path based on a series of small steps in order to develop their competences and improve the quality of their lives. The *H1* hypothesis has been confirmed.

In the case study, the managers pointed out, among others, the following examples of the benefits of implementing Personal Kaizen in their lives:

- order, no need to look for things,
- effective task management,
- less waste,
- improving the quality of personal life, more free time,
- self-consciousness and orderly life,
- systemic approach to personal development, visible effects,
- the use of small steps, resulting in lasting effects,
- effective action and lack of discouragement,
- good habits developed (e.g. regarding physical condition and health),
- improvement in the sphere of self-discipline, assertiveness, self-control, and relationships in the family.

As it follows from the above, managers, implementing the principles and instruments of Personal Kaizen achieve specific benefits, which confirms the *H2* hypothesis.

A compilation of the research questions, research hypotheses, and results of verification of the research hypotheses has been presented in Table 6.

The goal of the study conducted has been to identify the benefits of implementing Personal Kaizen and cannot be the basis for generalisation. The set of principles and tools proposed by the authoress constitutes a framework for personal improvement and can be a guidepost for all those for whom personal development, increasing competences, and improving the quality of life are of importance. It may also be an inspiration for personal development. Small steps allow for achieving lasting results, and a wide range of flexible instruments allows for choosing those that meet personalised needs. In the respondents' opinion, the fact of participating in the Lean and Kaizen projects contributed to a different perception of waste, realising the amount of time and resources wasted, the need for continuous improvement, changing the approach to problems and the way they are solved, cooperating with household members, willingness to develop thriftiness habits, maintaining order and better organisation of activities in personal life.

Table 6 – Verification of the Hypotheses (Source: Author)

Research question	**Hypothesis**	**Hypothesis verification**
RQ1. How and in what way do Polish managers use the knowledge gained within their professional work on Kaizen/Lean Management to develop competences in their personal lives and improve their personal productivity?	H1. Polish managers working in the Kaizen/Lean environment, benefit from the knowledge on the principles and instruments of Kaizen/Lean Management gained during their professional work, to develop competences in their personal lives and improve personal productivity.	5 people participated in the study and 18 people applied. Each of the interviewees declared applying selected Kaizen/Lean Management principles and instruments in the process of developing their own competences. The interviewees declared that the application of Kaizen/Lean Management principles and instruments in their personal lives has significantly improved their personal productivity. *The hypothesis has been confirmed.*
RQ2. What benefits do Polish managers achieve from implementing Personal Kaizen in their personal lives?	H2. Polish managers achieve specific benefits related to using the Personal Kaizen approach in their personal lives.	Polish managers declare achieving specific benefits connected with the use of Personal Kaizen principles and instruments in their personal lives. These benefits have been indicated in points 4 and 5 of this article. *The hypothesis has been confirmed.*

During the study, the participants repeatedly stated that they had just realised that they could improve something and discovered further areas with room for improvement in their personal lives yet. They suggest that it is worth promoting the assumptions and instruments of the Lean and Kaizen approach (the so-called Personal Kaizen Toolbox, mentioned in point 2 of the study) in order to reduce waste in all activities. The research results can constitute a benchmark for managers and they are illustrative, application and inspirational.

As part of further research, it is worth expanding its scope by next instruments and subsequent analysed cases. This will allow for to creating a set of Personal Kaizen Toolbox instruments, and maybe a diagnosis sheet for the current level of the manager's personal thriftiness.

REFERENCES

Baccarani, C., Mascherpa, V. and Minozzo, M., 2013. Zen and well-being at the workplace. *The TQM Journal,* [e-journal] 25(6), pp.606-624. DOI: 10.1108/TQM-07-2013-0077.

Bicheno, J. and Holweg, M., 2009. *The Lean Toolbox: The essential Guide to Lean Transformation.* 5th Ed. Buckingham: PICSIE Books.

Drucker, P.F., 2005. Managing Oneself. *Harvard Business Review.* [online] Available at: <https://hbr.org/2005/01/managing-oneself> [Accessed 14 July 2020].

Emiliani, M.L., 1998a. Continuous personal improvement. *Journal of Workplace Learning*, [e-journal] 10(1), pp.29-38. DOI: 10.1108/13665629810370021.

Emiliani, M.L., 1998b. Lean behaviors. *Management Decision*, [e-journal] 36(9), pp.615-631. DOI: 10.1108/00251749810239504.

Garvin, D.A., 1993. Building a Learning Organization. *Harvard Business Review*, 71(4), pp.78-91.

Graban, M., 2011. *Lean Hospitals: Improving Quality, Patient Safety, and Employee Engagement*. 2nd Ed. Wrocław: ProdPublishing.

Imai, M., 2006. *Gemba Kaizen: A Commonsense, Low-Cost Approach to Management.* Warsaw: MT Biznes Publishing.

Imai, M., 2007. *Kaizen: The Key to Japan's Competitive Success*. Warsaw: MT Biznes Publishing.

Keating, C., Backholer, K. and Peeters, A., 2014. Prevalence of overweight and obesity in children and adults. *The Lancet*, 384, pp.2107-2108.

Liker, J.K. and Hoseus, M., 2009. *Toyota Culture*. Warsaw: MT Biznes Publishing.

Liker, J.K. and Meier, D.P., 2011. *The Toyota Way Fieldbook*. Warsaw: MT Biznes Publishing.

Liker, K.L., 2005. *The Toyota Way: 14 Management Principles of The Toyota Way*. Warsaw: MT Biznes Publishing.

Locher, D., 2012. *Lean Office and Service Simplified: The Definitive How-To Guide*. Warsaw: MT Biznes Publishing.

Markovitz, D., 2011. *A factory of one*. Boca Raton: CRC Press Taylor & Francis Group.

Maurer, R., 2007. *One Small Step Can Change Your Life: The Kaizen Way*. Gliwice: HELION.

Maurer, R., 2016. *The Spirit of Kaizen: Creating Lasting Excellence One Small Step at a Time: Creating Lasting Excellence One Small Step at a Time*. Gliwice: HELION.

Roberts, CH., and McDonald, G., 1995. Training to fail. *Journal of Management Development*, 14(4), pp.16-31.

Sławińska, M. and Witczak, H., 2008. *Methodological basis of doctoral theses in economic sciences*. Warsaw: PWE.

Stoller, J., 2015. *The Lean CEO: Leading the Way to World-Class Excellence*. Warsaw: MT Biznes Publishing.

Suarez-Barraza, M.F., Ramis-Pujol, J. and Dahlgaard-Park, S., 2013. Changing quality of life through the Personal kaizen approach: a qualitative study. *International Journal of Quality and Service Science*, [e-journal] 5(2), pp.191-207. DOI: 10.1108/IJQSS-03-2013-0015.

Womack, J.P. and Jones, D.T., 2008. *Lean thinking: Banish Waste and Create Wealth in Your Corporation, Revised and Updated.* Wrocław: ProdPress.com.

Yin, R.K., 1994. *Case Study Research. Design and Methods.* Newbury Park: Sage Publication.

Yoneyama, E., 2007. Phenomenology of life, Zen and management. *Society and Business Review*, [e-journal] 2(2), pp.204-217. DOI: 10.1108/17465680710757411.

ABOUT THE AUTHOR

Iwona Burka – Katowice Business University (GWSH Górnośląska Wyższa Szkoła Handlowa), Poland, Faculty of Management, PhD., e-mail: iwonaburka@efektywnosc.net.pl, Author's ORCID: 0000-0002-7338-8830.

CONFLICTS OF INTEREST

The author declares no conflict of interest. The funder had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

New Approach to Portfolio Creation Using the Minimum Spanning Tree Theory and Its Robust Evaluation

DOI: 10.12776/QIP.V24I2.1450

Jakub Danko, Vincent Šoltés, Tomáš Bindzár

Received: 2020-04-20 Accepted: 2020-05-25 Published: 2020-07-31

ABSTRACT

Purpose: The aim of this paper is to describe another possibility of portfolio creation using the minimum spanning tree method. The research contributes to the existing body of knowledge with using and subsequently developing a new approach based on graph theory, which is suitable for an individual investor who wants to create an investment portfolio.

Methodology/Approach: The analyzed data is divided into two (disjoint) sets – a training and a testing set. Portfolio comparisons were carried out during the test period, which always followed immediately after the training period and had a length of one year. For the sake of objectivity of the comparison, all proposed portfolios always consist of ten shares of equal weight.

Findings: Based on the results from the analysis, we can see that our proposed method offers (on average) the best appreciation of the invested resources and also the least risky investment in terms of relative variability, what could be considered as very attractive from an individual investor's point of view.

Research Limitation/implication: In our paper, we did not consider any fees related to the purchase and holding of financial instruments in the portfolio. For periods with extreme market returns (sharp increase or decrease), the use of Pearson's correlation coefficient is not appropriate.

Originality/Value of paper: The main practical benefit of the research is that it presents and offers an interesting and practical investment strategy for an individual investor who wants to take an active approach to investment.

Category: Research paper

Keywords: portfolio creation; S&P 500; minimum spanning tree; graph theory; optimization

1 INTRODUCTION

In the last decade, investors in most developed countries have been facing an environment of persistently low-interest rates. As a result of the financial and economic crisis from the turn of the years 2008 and 2009, several central banks have sought to foster economic growth and stabilize their economies by using and applying an expansionary monetary policy. In the context of financial markets and investment, this aspect results in a low-interest-rate environment. Investors in financial markets are currently facing the challenge of how to optimally appreciate their investment in terms of investment risk. The range of investment instruments available for investors is wide, as is the range of viable investment strategies and approaches.

In this article, we present the possibility of creating an equity investment portfolio using the minimum spanning tree method. Using this approach, during the training period we focus on selecting specific stocks that are part of the S&P 500 stock index. When calculating the minimum spanning tree, we used adjusted correlation coefficients as a metric to determine the distance in these structures. As a result, equities that have a low correlation of returns are close to each other in these structures. This may be interesting in the context of portfolio risk diversification. To increase the robustness of the results, we use 10,000 simulations, and for each of the 10,000 training sets, we estimate the minimum spanning tree to build four portfolios in different ways. To ensure the objectivity of the comparison, all portfolios always consist of ten equities of equal weight. Subsequently, we compare the performance of these portfolios over a test period of one year. The goal is to prove that the portfolio we had proposed, marked as portfolio A, has the best performance compared to the other alternative portfolios marked as B, C, and D. Portfolios are compared mainly with regards to average annual return and their relative variability.

2 LITERATURE REVIEW

From the time since Markowitz introduced his Modern Portfolio Theory (1952;1959), there has been a lot of literature focused on new approaches to the portfolio creation process. In our paper, we use the graph theory approach for building the investment portfolio. We focus on the minimum spanning tree method (MST). One of the first authors to introduce this approach was Mantegna (1999). From this time, there were several authors dealing with this topic – see e.g. (Onnela et al., 2003a; 2003b; Bonanno et al., 2004).

Tola et al. (2008) show that clustering algorithms are suitable for improving portfolio reliability concerning the ratio of predicted and realized risk. Naylor, Rose and Moyle (2007) use two techniques – ultrametric hierarchical tree and minimum spanning tree for extraction of a topological influence map in the market of major currencies. Birch, Pantelous and Soramäki (2015) use three methods, including MST, and compare them for filtering information from the

DAX index. There have been many more authors dealing with this topic in recent times. Wang, Xie and Chen (2017) study the US stock market from a network perspective. They use the minimum spanning tree method and also planar maximally filtered graph method (PFG) and construct MST and PFG networks in the US market at different time scales. Danko and Šoltés (2018) use graph characteristics as a stock-picking tool and propose a portfolio while minimizing its standard deviation.

Over time, two basic approaches to investment have been formed - active and passive. There is a persisting debate as to which is preferable and there are several key studies in this field.

Sorensen, Miller and Samak (1998) compare active and passive forms of investment by studying stock-picking skills and comparing them to index investing in various market conditions. They conclude that both strategies have justification: an active approach to the portfolio could be more effective when there are tough (bearish) periods on the market, but in bullish periods a passive investment strategy or indexing could outperform the active approach.

In his work, Blitz (2014) points out the problems of passive investment. He claims that passive investing is efficient only if there are a lot of active investors because they are inevitable for efficient capital markets. As an alternative to passive investing, he offers a factor investing approach.

Fahling, Steurer and Sauer (2019) examined the sample of 194 actively managed funds in German equity markets and compared the results with passive investing. The arithmetic average annual return of the sample is better than the benchmark. The results show that active funds perform slightly better in terms of risk-adjusted performance. In our paper, we focus on a method that we can classify as an active form of investment.

3 METHODOLOGY

In our research, we employ the closing prices of shares forming the S&P 500 index in the period from 2009-01-02 to 2019-10-18. In this period, we have complete data of 450 financial instruments for the period of 2,718 trading days (over 10 years).

From the daily closing prices of the stocks forming the analyzed index, we calculated the daily returns in the standard way:

$$r_i = \frac{P_i - P_{i-1}}{P_{i-1}} [\times 100\%], \quad (1)$$

where p_i is the closing price at the time i, $p_{(i-1)}$ is the closing price on the previous trading day (at the time $i - 1$) and r_i represents the daily rate of return at the time i. Thus, every single stock from the analyzed index is represented by a daily return vector.

The idea of the analysis is that we take any time interval of at least two years (assuming that one business year has approximately 253 business days) and divide it into two parts: a training and a testing set. The testing set is still 253 trading days (one year) long and the training set can be any length, at least 253 trading days and a maximum of 2,465 days (2,465 = 2,718 – 253), with the training set, immediately preceding the testing set. Simply put, we create a portfolio based on data from different time periods (at least for the previous 253 days – maximum for the last 9 years or 2,465 days) and we still test portfolio performance for the year after we acquired this data. Simplification is shown in Figure 1.

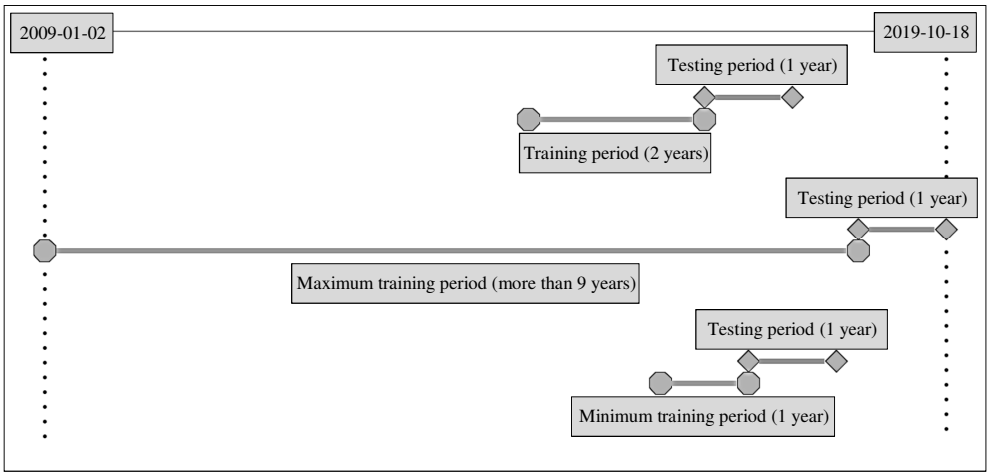

Figure 1 – Timeline

Because we want the results to be unaffected by the length of the training set, we perform 10,000 simulations and in each simulation, we select a training set of any length according to the abovementioned conditions. Thereafter we pair each training set with a testing set which immediately follows. On each training set, we calculate the correlation matrix of daily returns according to the following formula (2).

$$\rho_{i,j} = \frac{cov_{i,j}}{\sigma_i . \sigma_j}, \tag{2}$$

where $\rho_{i,j}$ is the correlation coefficient between returns of stocks i and j, $cov_{i,j}$ is the covariance between these returns and σ_i is the standard deviation of the stock's i return.

Correlations of daily returns are mostly positive, with negative correlations of daily returns in these periods of at least one year being very rare. If there are any, they are statistically insignificant and negatively linearly dependent (independent) with values close to zero. Nevertheless, these negative values must be taken into account, and the way we do this is outlined below. We decided not to transform these correlation coefficients into a distance matrix, as is common in

similar analyzes (Mantegna, 1999), but to calculate the minimum spanning tree based on correlation coefficients. Based on distance matrices and according to Mantegna, due to the character of the minimum spanning tree, they form a structure in which the most similar objects are closest to each other. Using the correlation coefficients, the least similar objects will be close to each other in these structures (those stocks that have a very low mutual correlation of returns).

Due to the fact that we need to know the distances between objects when estimating the minimum spanning tree, if we consider the correlation coefficients to be this metric, it is necessary to slightly adjust them to avoid negative values. The idea is to have the most dissimilar shares as close as possible. Taking into account that the correlation coefficient takes values from -1 (absolutely negative linear dependence) to 1 (absolutely positive linear dependence), we logically want lower correlation coefficient values to represent smaller distances and higher correlation coefficients greater distances, taking into account non-negativity. The simplest solution seems to be to adjust the correlation coefficients by adding the unit to them. Thus we get a kind of "pseudo-correlation" coefficient (distance measure), which takes values from 0 (absolutely negative linear dependence) to 2 (absolutely positive linear dependence). This fulfills the condition of non-negativity of the distance of objects, which is required for the calculation of the minimum spanning tree.

$$d(i,j) = \rho_{i,j} + 1 \tag{3}$$

In this way, we estimate the minimum spanning tree on each of the 10,000 training sets and build the portfolios in 4 different ways, denoted as A, B, C, and D in the work. To ensure a fair comparison of the results, we create the portfolios still out of ten shares with the same weights in the portfolio. With such a large number of analyzed stocks (450), the resulting structures that emerge still profile one large-scale stock (a kind of center from which the edges associate with other stocks). An example is shown in Figure 2 – *Stock Structure*, where we divided the stocks into three groups with the following color differentiation:

- The black color represents the center of the graph, i.e. (that is) stock with a maximum degree.
- The grey color shows all stocks that represent the neighbors of the graph center, i.e. we can get to the center of the graph through one edge.
- The white color represents all other stocks, i.e. we can get to the center of the graph through at least two edges.

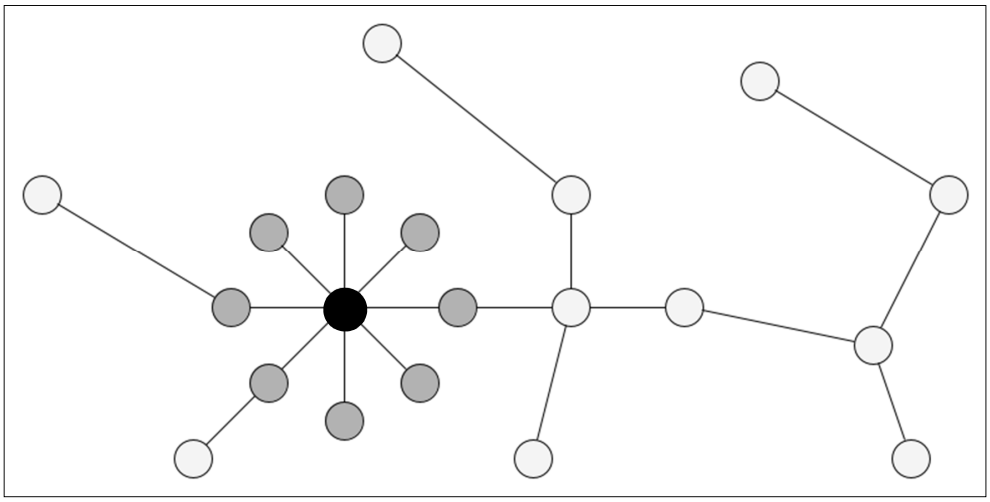

Figure 2 – Stock Structures

Type A Portfolios

We build the type A portfolios in the following way:

- We identify the center of the graph (the stock with maximum degree) that represents the first stock in the portfolio.
- Of all the shares that are incidental (adjacent) to this stock, we select nine closest to the center. This means that these ten stocks are a combination of shares that are the least similar based on the correlation of their daily returns.
- In this way, we will get ten stocks that create the type A portfolios, and each of those stocks has the same weight.

Type B Portfolios

We build the type B portfolios in the following way:

- We identify the center of the graph (the stock with maximum degree) that represents the first stock in the portfolio.
- Of all the shares that are incidental (adjacent) to this stock, we randomly select nine stocks.
- In this way, we will get ten stocks that create the type B portfolios. Each stock has the same weight.

Type C Portfolios

We build the type C portfolios in the following way:

- We identify the center of the graph (the stock with maximum degree) – this stock does not constitute an object of our interest.

- We identify all stocks that are incidental (adjacent) to this stock – these stocks do not constitute objects of our interest.
- Of all the other stocks, we randomly select ten stocks.
- In this way, we will get ten stocks that create the type C portfolios. Each stock has the same weight.

Type D Portfolios

We build the type D portfolios in the following way:

- Of all shares, we randomly choose ten shares.
- In this way, we will get ten stocks that create the type D portfolios. Again, each stock has the same weight.

From the set theory point of view, type C portfolios and type B portfolios have an empty intersection and their unification creates type D portfolios. Type A portfolios are a subset of type B portfolios and they are a subset of all possible portfolios (type D portfolios).

Assuming that using minimum spanning tree method we have identified the stocks that are, because of their greatest dissimilarity in the context of yield correlation, the most appropriate for portfolio creation, type A portfolios should perform best in terms of both real return and relative risk measured by the coefficient of variation. These should be followed by type B portfolios because they have the same principle of creation, but not under fully effective conditions (we do not select the nearest shares but randomly select the shares that are adjacent to the graph center). If we assume that the type C portfolios contain the stocks we consider to be the least appropriate, we expect these portfolios to provide us with the worst performance. We can consider type D portfolios as a benchmark, due to the fact that we build them by random stock selection. We expect that the type A portfolios we propose will perform better than portfolios B and that portfolios C will perform worse than portfolios A and B.

For each of the 10,000 simulations, we build a training set on which we estimate the minimum spanning tree, from which we create typologically 4 different portfolios. In the second part of the analysis, we consider a testing set with a length of one year (253 trading days) immediately following by the training set. In this testing set, we compare the portfolio's value at the beginning and at the end of the period to obtain the real annual return on a particular portfolio for each simulation. In this way, we obtain 10,000 real annual return values for all types of portfolios.

The risk quantification of these portfolios is a bit more complicated. We start from the covariance matrix of daily returns from the beginning to the end of the test period. The standard deviation of daily returns for this period is calculated simply by the following formula (4).

$$\sigma = \sqrt{w.\sum.w^T}, \tag{4}$$

where w represents row weight vector, $\sum$ is the covariance matrix of profitability and w^T transposed row weight vector (column weight vector).

In our case, the vector w for all portfolio types is a vector with ten equal weights (1/10). In this way, we calculate the standard deviation of daily returns, which we need to transform into a standard deviation of annual returns in order to compare it with real annual returns. We use the approximate calculation given by the formula (5).

$$\sigma_{year} = \sigma_{day} \cdot \sqrt{n_{days\ per\ year}}, \tag{5}$$

Where σ_{year} represents the standard deviation of annual returns, σ_{day} represents the standard deviation of daily returns, and $n_{days\ per\ year}$ represents the average number of days that the business year contains (we used 253 business days throughout the whole analysis).

In this way, we get the result for each simulation in the form of the composition of individual portfolios, their real annual profitability, and the approximate standard deviation of the annual profitability.

By dividing the standard deviation and the annual rate of return of a particular portfolio, we get the ratio indicator – the coefficient of variation of the annual rate of return of the portfolio, which represents a measure of relative variability according to the relation (6), which can be interpreted as inverted Sharpe ratio.

$$\frac{\sigma_P}{E(r_p)} \tag{6}$$

For portfolios of each type, we have 10,000 real annual return values and 10,000 approximate standard deviations (absolute risk) values. By averaging these values, we get the results that are presented in the next section.

4 RESULTS

In this section, we present the results we have obtained from comparing our proposed portfolio with alternative portfolios. The summary results are shown in Table 1.

Table 1 – Comparison of Portfolios

Portfolio type	A	B	C	D
Average annual rate of return	16.62%	13.95%	12.74%	13.75%
Average standard deviation	0.00192	0.00194	0.00181	0.00186
Average coefficient of variation	0.01156	0.01392	0.01422	0.01350

Type A portfolio

The type A portfolio proposed by us achieved an average annual return of 16.62% and an average standard deviation of 0.00192, which represents an average coefficient of variation of 11.56%.

As an interesting addition to the results of the analysis from the investor's point of view, we also include a more detailed description of the type A portfolios. These portfolios represent the core of our proposed method of portfolio creation. We describe these portfolios in terms of most represented stocks in them. As a central vertex in the 10,000 portfolios, the most frequently featured stock became Newmont Corporation (NEM) – 4,475 times, followed by Weyerhaeuser (WY) – 2,744 times, Expedia (EXPE) – 635 times, Dollar Tree (DLTR) – 599 times, and Netflix (NFLX) – 566 times. Other stocks were at the center of the graph less than three hundred times.

More important than the center of the graph is the success of the shares in terms of participation in the portfolio (proximity to the center), where out of 10,000 portfolios in 5,569 of them is stock NEM (Newmont Corporation), in 4,424 portfolios is stock DLTR (Dollar Tree), in 3,814 portfolios is stock EXPE (Expedia), in 3,352 portfolios is stock PEP (PepsiCo) and in 3,137 portfolios is stock ABMD (Abiomed). The frequency of other stocks is illustrated in Figure 3 by the size of the stock-specific rectangle.

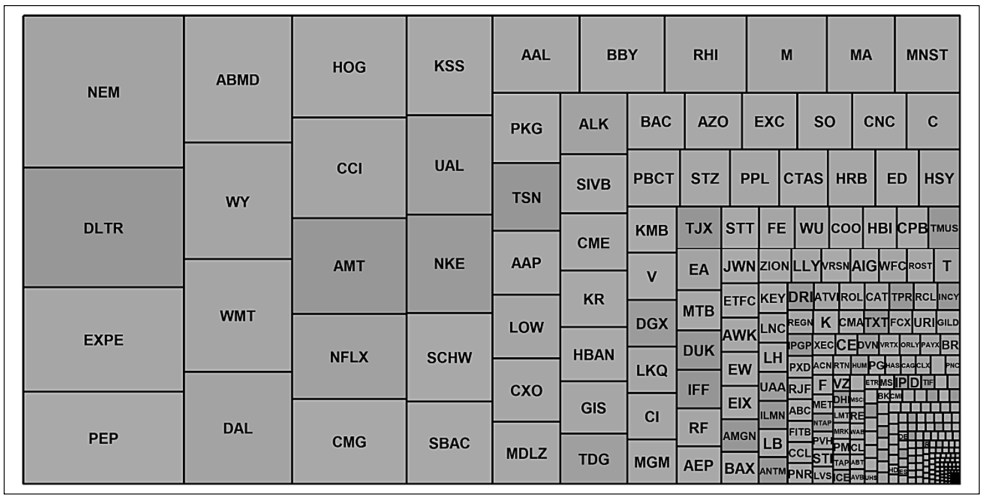

Figure 3 – Frequency in Type A Portfolio

Type B portfolio

The type B portfolio achieved an average annual return of 13.95%. The average standard deviation is 0.00194, and the average coefficient of variation is 13.92%. Compared to the portfolio proposed by us, portfolio B performed worse in all indicators.

Type C portfolio

The type C portfolio has an average annual return of 12.74%, an average standard deviation of 0.00181, and an average coefficient of variation of 14.22%. Our portfolio outperformed portfolio C in both average annual returns and the average coefficient of variation.

Type D portfolio

The type D portfolio achieved an average annual rate of return of 13.75% and an average standard deviation of 0.00186. The average coefficient of variation is 13.50%. Again, our portfolio outperformed the type D portfolio both in average annual profitability and in the average coefficient of variation.

The comparison clearly shows that in line with our assumptions, our portfolio A was able to outperform all alternative portfolios both in terms of average annual profitability and also in terms of average relative variability. From the investor's point of view, the strategy we propose brings the best appreciation of invested resources, but also the least risky investment in terms of relative variability.

5 CONCLUSION

In this work, using the minimum spanning tree method, we focused on a portfolio creation process appropriate for an individual investor. Portfolios were compiled based on data analysis from various long-term windows, while the portfolio performance evaluation was carried out based on the development for the following period of one year. The advantage of the analysis is that the results do not depend on the length of time that was chosen to build the portfolio, as we chose different time periods to increase the robustness of the results. Based on the results from the analysis, we can see that our proposed method offers (on average) the best appreciation of the invested resources and also the least risky investment in terms of relative variability – very attractive combination from an individual investor's point of view. The results we achieved were consistent with the assumptions made during the analysis. In this way, the article presents and offers an interesting and practical investment strategy for an individual investor who wants to take an active approach to investment.

REFERENCES

Birch, J., Pantelous, A. and Soramäki, K., 2015. Analysis of Correlation Based Networks Representing DAX 30 Stock Price Returns. *Computational Economics*, [e-journal] 47(4), pp.501-525. DOI: 10.1007/s10614-015-9481-z.

Blitz, D., 2014. The dark side of passive investing. *Journal of Portfolio Management*, [e-journal] 41(1), pp.1-4. DOI: 10.3905/jpm.2014.41.1.001.

Bonanno, G., Caldarelli, G., Lillo, F., Micciche, S., Vandewalle, N. and Mantegna, R.N., 2004. Networks of equities in financial markets. *The European Physical Journal B*, [e-journal] 38(2), pp.363-371. DOI: 10.1140/epjb/e2004-00129-6.

Danko, J. and Šoltés, V., 2018. Portfolio creation using graph characteristics. *Investment Management & Financial Innovations*, [e-journal] 15(1), p.180. DOI: 10.21511/imfi.15(1).2018.16.

Fahling, E.J., Steurer, E. and Sauer, S., 2019. Active vs. Passive Funds—An Empirical Analysis of the German Equity Market. *Journal of Financial Risk Management*, [e-journal] 8(2), p.73. DOI: 10.4236/jfrm.2019.82006.

Mantegna, R.N., 1999. Hierarchical structure in financial markets. The European Physical *Journal B-Condensed Matter and Complex Systems*, [e-journal] 11(1), pp.193-197. DOI: 10.1007/s100510050929.

Markowitz, H., 1952. Portfolio Selection. *The Journal of Finance*, [e-journal] 7(1), pp.77-91. DOI: 10.2307/2975974.

Markowitz, H., 1959. *Portfolio selection: Efficient diversification of investments*. New York: John Wiley & Sons.

Naylor, M.J., Rose, L.C. and Moyle, B.J., 2007. Topology of foreign exchange markets using hierarchical structure methods. *Physica A: Statistical Mechanics and its Applications*, [e-journal] 382(1), pp.199-208. DOI: 10.1016/j.physa.2007.02.019.

Onnela, J.P., Chakraborti, A., Kaski, K., Kertesz, J. and Kanto, A., 2003a. Asset trees and asset graphs in financial markets. *Physica Scripta*, [e-journal] 2003(T106), p.48. DOI: 10.1238/Physica.Topical.106a00048.

Onnela, J.P., Chakraborti, A., Kaski, K., Kertesz, J. and Kanto, A., 2003b. Dynamics of market correlations: Taxonomy and portfolio analysis. *Physical Review E*, [e-journal] 68(5), p.056110. DOI: 10.1103/PhysRevE.68.056110.

Sorensen, E.H., Miller, K.L. and Samak, V., 1998. Allocating between active and passive management. *Financial Analysts Journal*, [e-journal] 54(5), pp.18-31. DOI: 10.2469/faj.v54.n5.2209.

Tola, V., Lillo, F., Gallegati, M. and Mantegna, R.N., 2008. Cluster analysis for portfolio optimization. *Journal of Economic Dynamics and Control*, [e-journal] 32(1), pp.235-258. DOI: 10.1016/j.jedc.2007.01.034.

Wang, G.J., Xie, C. and Chen, S., 2017. Multiscale correlation networks analysis of the US stock market: a wavelet analysis. *Journal of Economic Interaction and Coordination*, [e-journal] 12(3), pp.561-594. DOI: 10.1007/s11403-016-0176-x.

ABOUT AUTHORS

Jakub Danko – (J.D.) University of Economics, Prague, Czech Republic, Faculty of Informatics and Statistics, Assist. Prof., e-mail: jakub.danko@vse.cz, Author's ORCID: 0000-0002-1790-3324.

Vincent Šoltés – (V.S.) Technical University of Košice, Košice, Slovak Republic, Faculty of Economics, Department of Finance, Prof., e-mail: vincent.soltes@tuke.sk, Author's ORCID: 0000-0002-2656-5582.

Tomáš Bindzár – (T.B.) Technical University of Košice, Košice, Slovak Republic, PhD. Student, Faculty of Economics, Department of Finance, e-mail: tomas.bindzar@student.tuke.sk, Author's ORCID: 0000-0003-1652-9737.

AUTHOR CONTRIBUTIONS

J.D. – methodology, software, visualization; V.S. – conceptualisation, supervision, validation, project administration; T.B. – writing—review and editing, formal analysis, data curation.

CONFLICTS OF INTEREST

The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Why Most University-Industry Partnerships Fail to Endure and How to Create Value and Gain Competitive Advantage through Collaboration – A Systematic Review

DOI: 10.12776/QIP.V24I2.1389

Arlindo Marinho, Rui G. Silva, Gilberto Santos

Received: 2020-01-20 Accepted: 2020-04-25 Published: 2020-07-31

ABSTRACT

Purpose: The collaborative work between universities and enterprises is increasingly important given the growing competitive environment, transformed by intense global competition, rapid technological change and shorter product life cycles. In this article several case studies are analysed, as well as relevant academic literature, to get an insight concerning the current relationship status between enterprises and universities.

Methodology/Approach: The methodology is grounded on the literature review on University-Industry Partnerships, selected from indexed sources, and targeted to case-based experiences where practical feedback is given.

Findings: It is shown that university-industry knowledge transfer contributes significantly to an increase of quality, productivity and economic value of businesses. It is found that the lack of controlling mechanisms and lack of efficient communication paths have a strong negative impact in collaboration. Key collaboration factors that support stronger relationships were compiled and discussed in support of better mitigation strategies.

Research Limitation/implication: The limited availability of case studies reporting on operative improvements introduced by policy changes hinders the effectiveness of the findings.

Originality/Value of paper: The paper analyses the collaboration of the university with industry based on case studies with a focus on value creation and how to gain a competitive advantage through collaboration.

Category: Literature review

Keywords: collaboration; knowledge transfer; competitiveness

1 INTRODUCTION

Collaboration between universities and industry in Europe, presents a heterogeneous landscape. Cooperation enterprise – university is a recent phenomenon and incipient (Unger et al., 2018; Jarábková, Chreneková and Roháčiková, 2019). Companies and universities are different in their organization and cultures and hence technology management perspectives are different. There is still significant room for better and more intense cooperation between universities and businesses, but difficulties remain in terms of trust and understanding of the operation on both sides (Paľová, Czaja and Vejačka, 2018; Roud and Vlasova, 2018). Nevertheless, models or theoretical frameworks to understand how customers and other stakeholders are involved with companies in collaborative activities of co-creation of value (Frow et al., 2015) are still poor. Much of the attention in research has recently been placed on case studies of large multinational companies and well-known universities (Edmondson, 2012). Few studies have considered the collaboration between universities, SMEs and non-governmental organizations.

Higher education institutions and companies benefit from working together and collaboration encourages the transfer and sharing of knowledge hence aiding in the creation of long-term partnerships (Guan and Zhao, 2013). In the process of innovation, it is also increasingly highlighted the importance of collaboration between science, information and technology (Rebelo, Santos and Silva, 2015; Wu and Chiu, 2015; Unger et al., 2018) and that, in the advanced industrial economy, there is a strong integration of the activities of science and technology systems (Bravi, Murmura and Santos, 2018). Collaboration is becoming a powerful source of competitive advantage. However, the process of knowledge transfer has some weaknesses particularly with regards to communication and collaboration between universities and companies with influence on their capacity for innovation and hence compromise successful relationships. There are nevertheless key factors present in the determinants of interaction that, identified and properly addressed, will overcome some of the barriers and assist the management of collaboration university – enterprise making it more profitable and thus a source of value for both parties and society in general.

The aim of the paper is hence to explore, identify and discuss the importance of key issues pertaining to the relationship between universities and industry and how to mitigate the effects of neglected factors. The article first addresses the common strategies used in the university-industry relationship highlighting the benefits that can be brought through those partnerships. The next section highlights the role of innovation in getting sustained competitive advantage for industry and why universities are the perfect allies to achieve those goals. The following two sections report on the success of partnerships and highpoint major constraints that undermine the success and future of collaboration between industry and universities.

2 METHODOLOGICAL APPROACH

The methodology is grounded on the literature review on University-Industry Partnerships, selected from indexed sources, and targeted to literature reviews in the field as well as case-based experiences where practical feedback is given. This paper presents a critical review on the key aspects pertaining to the relationship between universities and industry. The approach used in elaborating the article consisted of two steps: (i) identify the central theme of the research, through a general literature search in ScienceDirect, Google Scholar, Springer, Taylor&Francis and major indexed journals in SCOPUS using the terms "Collaboration", "Knowledge Transfer", "Competitiveness", and "Value Creation", and (ii) select major case studies that dwell on the central topic and sent address the relationship problems. In order to ensure that relevant case studies and reviews on collaboration mechanisms pertaining to the University-Industry relationship were not neglected, relevance was taken into account as well as temporal limitation to the last 10 years, especially between 2014 and 2019, with some exceptions pertaining to historical reports, agreements or governing acts. The most accessed articles in the period were also reviewed to verify and categorize the themes.

3 COLLABORATION STRATEGIES AND COMPETITIVE ADVANTAGE

The word collaboration is associated with positive ideas of interaction and convergence of efforts in favour of a particular accomplishment. People who come to collaborate do it in the expectation of achieving, in interaction with others, something that previously did not have. This approach can result in professional development and contributions to improvements in the areas in which they operate. Collaboration is therefore an instrument that promotes the development of people and activities and allegedly also the organizations to which they belong (Reficco et al., 2018). When there is a collaborative relationship responsibility must be balanced and assumed by all stakeholders, regardless of their different functions (Lin, 2017). This process should take into account four aspects: Conceptual convergence, which highlights the importance of having a common understanding of what it means and what it implies to collaborate and agree on the interest to work together for a particular purpose, building a shared vision; Agreement on the definition of objectives, which touches the importance of being defined as a whole leading to a course of actions in which all see themselves and engage; Shared management, highlighting the assumption of responsibility for managing the process; and, finally, the anticipation of individual incomes that should result in benefits for each and for all participants, to the organizations to which they belong and the society, in its relationship with the developed activity and the people involved in it. In collaborative work the will to accomplish together should always be present, thus it implies having confidence in others, valuing their knowledge and experience

and believe that it is possible to go further together than alone (Un and Rodríguez, 2018). The generation of ideas to solve problems is the currency of the future and should be addressed by well-integrated teams (Edmondson, 2012). Thus, adaptable and talented companies able to manage change and with skills to make the business environment faster and disruptive will be better positioned to address current and future challenges.

Competitive advantage has been viewed according to two dominant theories, one considering a view based on the market, which argues that the structure and competitive dynamics of the industry in which it operates are the main determinants of the company's performance and the vision based on resources that emphasizes the importance of resources that companies use as the main source of competitive advantage (Ghadimi, Ghassemi Toosi and Heavey, 2018). The notion of core competencies is closely related to the vision based on resources, deriving from it also a vision based on knowledge and vision based on capabilities. A recent formulation, the relational view, has received considerable attention and it argues that in the current globalization environment, most companies cannot deal with the rapidly changing competition confined to their own resources and capabilities. The success of contemporary companies is achieved through a sustainable competitive advantage, achieved by their ability to develop a set of core competencies that ensure a better approach to meet the needs of customers compared with its competitors (Eftimoski and Milenkovski, 2012).

4 COMPETITIVENESS VS. INNOVATION

Business leaders are faced with increasing competition, with globalization, with technological changes and a new strategic thinking. These factors, in a competitive and dynamic changing environment, should therefore be managed in order to build and sustain a high-performance organization. It is therefore necessary to combine scarce resources to achieve critical mass to boost the introduction of innovation in the market and improve collaboration especially concerning research and innovation activities that determine productivity growth and industrial competitiveness. Innovation has long been considered, given the growing competition, as a form of survival (Smirnova, Rebiazina and Khomich, 2018) and the request for innovation can come from explicit demands of the market or as a result of research and development activities within companies. Innovation is not just a process of creating ideas, but also the result of a complex social interaction, communication and exchange of knowledge.

The generation of innovation within the company involves two elements: resources and the ability to build networks – although innovation is not a sole characteristic of companies that are involved in networks but does happen also in individual companies (Konsti-Laakso, Pihkala and Kraus, 2012), as there is also a positive relationship between growth, innovation and the use of external relations of various types. In fact, the development of innovation requires support

for research and development and cooperation with other organizations (Jørgensen and Ulhøi, 2010). Hence, the networks represent a possible solution that enables companies, particularly SMEs, to overcome their constraints and promote innovation. The literature also points to the active role of consumers as a value co-creation key component in innovation activities (Liao, Hu and Ding, 2017). This outlook incorporates new forms of customer engagement (Santos, Murmura and Bravi, 2019), their interaction and integration with the innovation activities of the company (Cui and Wu, 2016; Ferreira Rebelo, Silva and Santos, 2017; Araújo et al., 2019). The key element of this approach relates to the recently-enabled client and connected that leverages new technological developments and attempts to achieve greater involvement and control of the innovations within companies (Orviská et al., 2019; Barbosa, de Oliveira and Santos, 2018).

Knowledge is the ingredient that underpins the competitiveness of regions, countries, sectors and companies. In the global economy the creation of high added value is based on innovation grounded on the ability to create and transform new ideas into new products and processes commercially valued. The health and wealth of societies increasingly depend on its ability to innovate and science and technology appear as a key source of competitive advantage. Linking basic and applied research to the market through technology transfer and marketing mechanisms, including partnerships government – university – industry, and capital investment, will constitute the key driver of sustainable competitive advantage and prosperity (Carayannis and Campbell, 2012). The public – media, culture and civil society – uses and applies knowledge and by doing so, is also part of the innovation system, in fact, in a society of knowledge and advanced economy, knowledge flows in all spheres of society. Diverse and heterogeneous culture configurations should help promote creativity, necessary and essential for the creation and production of new knowledge and innovations. This encourages the development of "creative knowledge spaces", i.e., environments, contexts and surrounding areas that have a positive influence on human beings involved in creative work (Resetarits and Resetarits-Tincul, 2012).

An advanced knowledge economy embraces innovation and creativity at the same time and the more mature and advanced it becomes then the more creative it is required to be since more knowledge, innovation and creativity can be absorbed (Dubina, Carayannis and Campbell, 2012; Chi et al., 2018). The creative economy inter-relates creatively technological innovations with social innovations. Also, the competitiveness and superiority of a system of knowledge or the advanced level of development of a knowledge system is highly determined by their ability to adapt to combine and integrate different modes of knowledge and innovation (Carayannis and Campbell, 2012). One major fact is that knowledge is socially constructed and is better supported through collaborations in which participants share knowledge and discuss projects that incorporate characteristics, real world content and use of various sources of information, in teams. This is the most widely supported present approach,

especially when regarding the use of information technology, and hence collaboration and knowledge transfer are gaining popularity.

5 COLLABORATION WITH UNIVERSITIES

Multinational companies have set the example in developing international networks known as "Centres of Excellence" with universities and research institutions and assumes a strategic importance for public administration at two different levels: from a national point of view it contributes to enhance national research and knowledge capital; from a local point of view it promotes the involvement of major economic and academic actors and represent a unique resource to promote a sustainable and durable development in the region. Within this framework, the strategic partnership is a form of organization designed to integrate different tools of intellectual capital, with the focus on specific projects, where each partner brings operational issues and different business cultures. It is believed that European universities can significantly increase their attractiveness since partnerships are a clear priority and also there is a growing number of academics who have worked in companies (Edmondson et al., 2012). Reinforcing this idea, the European Union has recognized the important role higher education institutions play – through education, research and innovation - in the transfer of knowledge to society and its vital contribution to the competitiveness of the European economy (Plewa et al., 2013). And yet, when universities form partnerships with industry, often the potential for synergy is frustrated due to communication failures. Although new information and communication technologies makes it easier to access research results, there is still a large gap between the knowledge produced by researchers and the one which is used in practice. The "European paradox" depicts a strong research capacity but lacks the strength to translate this into innovative products (Ranga et al., 2013). The reasons for this gap is attributed mainly to researchers who often have much more interest and devote more time and effort to the production of new knowledge than to their disclosure.

Several authors argue that the key to successful knowledge transfer is an on-going dialogue process, an accumulation of social networks, hence a strong emphasis should be placed in building strong personal relationships (as opposed to institutional relations), and building trust, a key element in business (Johnston and Huggins, 2016). The collaboration university-company is based on a mental attitude, it is driven by intrinsic and psychological elements (trust, mutual commitment, common objectives) and not by rules or quantifiable elements (Edmondson et al., 2012). The literature suggests as influences facilitators, in creating and maintaining Industry – University relationships, the confidence (Schilke and Cook, 2013), and the existence of collaboration champions, acknowledging the impact they have on technology and learning outcomes. Although cooperation between organizations can be achieved through various mechanisms, it should be highlight the importance of trust as a means to establish

and maintain the relationship (Huang and Wilkinson, 2014). Trust is a crucial element for effective negotiations, when negotiators trust each other they are more likely to develop a strategy for value creation and problem solving, involving information sharing, needs and help the other party achieve its objectives (de Klerk, 2012). In the context of inter-cooperative arrangements, trust is mutual confidence - no agent should exploit the vulnerabilities of others. It is therefore associated with the "problem" specific risk and is a remedy for its resolution. In the absence of an environment of uncertainty or risk, trust has no meaning. Trusting means travelling with the risk of betting on the uncertain future (Nguyen and Liem, 2013). Trust is built on inter-organizational relationships over time and come to play a greater role in maintaining the relationship (Schilke and Cook, 2013). In summary, trust is the result of behaviour "right", "fair" and "serious" where decisions and morally correct actions are based on ethical principles that recognize and protect the rights and interests of others (Hewitt-Dundas, Gkypali and Roper, 2019).

A successful collaboration does not just happen; it must be carefully planned and nurtured, it is then important to fully understand what makes this collaboration become a success (Rajalo and Vadi, 2017). When new knowledge is created as a result of a collaborative project, the perspectives of the parties are often changed (Bruneel, D'Este and Salter, 2010) – Scholars have an incentive to disseminate the results at the beginning, to improve their reputation in the scientific community and companies, in contrast, have an incentive to postpone disclosure of results, in order to ensure a competitive advantage. Other potential conflicts may arise due to different attitudes about intellectual property rights or differences in language and values that affect communication (Hewitt-Dundas, Gkypali and Roper, 2019). It is important that the motives and objectives of each partner are clearly identified and explained. Moreover, it is also important that all partners have the opportunity to influence decisions that affect the partnership (Anderson, Michael and Peirce, 2012). Removing barriers, however, is not enough: the drivers of the partnership and the presence of obvious benefits, which act as motivators, are also necessary to promote collaboration (Edmondson et al., 2012), building trust among all stakeholders (Piacentini, 2013; Rehák, Šipikal and Lešková, 2019).

6 MAJOR CONSTRAINS IN COLLABORATION STRATEGIES

From the given review there are some key aspects that should be addressed by universities and enterprises when engaging in a relationship in order to gain a competitive advantage through collaboration and enable future collaborations to flourish. Below are presented recommendations that summarize findings pertaining to the sustainable development of a University – Enterprise partnership. Table 1 presents a summary of major aspects neglected by most partnerships, corroborated by numerous authors that often lead to unsuccessful collaborations. These aspects are next discussed and analysed from a constructive

perspective highlighting the advantages and contributions to healthier and fruitful partnerships.

Table 1 – Neglected Collaboration Factors in Partnerships

Collaboration Factor	**Source**
Partner Selection	Giesecke (2012), Ranga and Etzkowitz (2013), Bunduchi (2013), Thune and Gulbrandsen (2014), Le Roy, Robert and Lasch (2016), Mindruta, Moeen and Agarwal (2016), Johnston and Huggins (2016), Criscuolo et al. (2017), Yoon et al. (2017), Ankrah and Al-Tabbaa (2017), Hartman and Dhanda (2018), Frølund, Murray and Riedel (2018), Reichert (2019)
Collaboration Management	Giesecke (2012), Ranga and Etzkowitz (2013), Ankrah and Omar (2015), Sum Chau, Mark and Serbanica (2016), Grimpe, Goel and Göktepe-Hultén (2017), Jonkers et al. (2018), Hartman and Dhanda (2018), Merritt and Kelley (2018), Frølund, Murray and Riedel (2018), Reichert (2019)
Interface Management	Giesecke (2012), Ranga and Etzkowitz (2013), Plewa et al. (2013), Thune and Gulbrandsen (2014), Perkmann and Schildt (2015), Yoon et al. (2017), Ankrah and Al-Tabbaa (2017), Champenois and Etzkowitz (2018), Al-Tabbaa and Ankrah (2018), Reichert (2019)
The Use of Champions	Ranga and Etzkowitz (2013), Plewa et al. (2013), Bstieler, Hemmert and Barczak (2015), Sum Chau, Gilman and Serbanica (2016), Yoon et al. (2017), Veles, Carter and Boon (2018), Merritt and Kelley (2018), De Silva, Howells and Meyer (2018), Reichert (2019)
Long-term partnership	Perkmann and Salter (2012), Giesecke (2012), Ranga and Etzkowitz (2013), Thune and Gulbrandsen (2014), Yoon et al. (2017), Frølund, Murray and Riedel (2018), Reichert (2019)
Shared Vision and Strategy	Giesecke (2012), Ranga and Etzkowitz (2013), Thune and Gulbrandsen (2014), Bstieler, Hemmert and Barczak (2015), Sum Chau, Gilman and Serbanica (2016), Yoon et al. (2017), Ankrah and Al-Tabbaa (2017), Hartman and Dhanda (2018), Frølund, Murray and Riedel (2018), de Wit-de Vries et al. (2018), Reichert (2019)

The literature highlights two distinct phases in the development of the collaborative process, namely "partner selection" and "partnership management" which are the key elements to its overall success. In the selection of a partner the necessary conditions for the success of a partnership is to ensure the correspondence between the motivations of companies and strengths of the university. The collaboration history, geographical proximity and technological capabilities are also important factors to consider. In the second phase, managing the partnership assumes relevance since outsourcing of innovation is vulnerable to opportunism on both sides, given the limits of uncertainty and monitoring. Both academics and companies seek partners with complementary objectives and skills, whose evaluation depends on the success of the project. Hence a key linkage must exist to ensure that there is frequent, and ongoing, personal involvement between university researchers and industry managers. Collaborations only work well when they are managed by people who cross

borders with ease and that have a deep knowledge of the two cultures, thereby the use of Champions.

The first step to a healthy partnership must be given with the identification of main university strengths and the core competencies of the company while aiming to identify promising opportunities for collaboration. It is required a thorough understanding of the three different types of possible partnerships; strategic, operational or transactional, and select the one that best fits both partners. Clearly defined objectives, realistic and mutually agreed are also very important factors in the management of collaborative projects. Without clearly defined goals, projects tend to become too large and beyond their initial limits. Moreover, the importance of the project leader and effective communication should be emphasized in order to continuously balance ambitions and expectations. There is no shortcut to cultivate personal ties that could lead to more creative and promising collaborations. Universities should create opportunities for scholars and researchers from companies and executives with common interests to meet and develop a dialogue. The most fruitful partnerships take time to produce results.

7 CONCLUSION

Several studies highlight the importance of universities as providers of significant opportunities for the development of new knowledge, in addition to teaching and research, changing the role of universities in marketing and contributing significantly to innovation. Research findings show, however, that companies typically do not often use university relationships to strengthen and build core competencies given that the cultural and philosophical differences still exist between industry and academia. Universities have a time orientation and goals that are different from industrial companies. In addition, it seems to emerge from literature that companies are apprehensive about the dependence on universities, especially in areas that are crucial to the creation of competitive advantage, even though it is demonstrated that it is a neglected area and can be very useful for companies. There are therefore difficulties that remain in terms of trust and understanding of operations on both sides. Reducing complexity, fostering relationships and introducing effective intermediation, particularly for small businesses, are recommendations that sustain strong, trustworthy relationships between people in business and academia setting the foundations for successful collaboration. Partnerships often fail to properly address the expected outcomes for both parties and hence cause the deterioration of the relationship. The compiled list of key neglected factors highlights the necessity for the prioritization of aspects such as: Partner Selection; Collaboration Management; Interface Management; The Use of Champions; Long-term partnership; and, Shared Vision and Strategy. People are key to making any collaboration work in a productive and lasting way wish is often undermined by poor leadership and negligence concerning the stakeholders' interests. Long-term

alliances build the human capital necessary for the operation of industry-university collaborations and over time a well-managed partnership produces an increasing number of teachers and graduate students who can think and act through cultural division, connecting with the company's major research interests and work harmoniously to define broad and common goals. It is thou required that both companies as well as academia delineate policies to mitigate nefarious effects of the neglected key relationship factors.

REFERENCES

Al-Tabbaa, O. and Ankrah, S., 2018. Engineered" University-Industry Collaboration: A Social Capital Perspective. *European Management Review*, [e-journal] 16(3), pp.543-565. DOI: 10.1111/emre.12174.

Anderson, T.S., Michael, E.K. and Peirce, J.J., 2012. Innovative Approaches for Managing Public-Private Academic Partnerships in Big Science and Engineering. *Public Organization Review*, [e-journal] 12(1), pp.1-22. DOI: 10.1007/s11115-010-0142-3.

Ankrah, S. and Al-Tabbaa, O., 2017. Universities — industry collaboration : A systematic review ScienceDirect Universities — industry collaboration : A systematic review. *Scandinavian Journal of Management*, [e-journal] 31(3), pp.387-408. DOI: 10.1016/j.scaman.2015.02.003.

Ankrah, S. and Omar, A., 2015. Universities–industry collaboration: A systematic review. *Scandinavian Journal of Management*, [e-journal] 31, pp.387-408. DOI: 10.1016/j.scaman.2015.02.003.

Araújo, R., Santos, G., da Costa, J.B., and Sá, J.C., 2019. The quality management system as a driver of organizational culture: An empirical study in the Portuguese textile industry. *Quality Innovation Prosperity,* [e-journal] 23(1), pp.1-24. DOI: 10.12776/QIP.V23I1.1132.

Barbosa, L.C.F.M., de Oliveira, O.J. and Santos, G., 2018. Proposition for the alignment of the integrated management system (quality, environmental and safety) with the business strategy. *International Journal for Quality Research,* [e-journal] 12(4), pp. 925-940. DOI: 10.18421/IJQR12.04-09.

Bravi, L., Murmura, F. and Santos, G. 2018. Manufacturing labs: Where new digital technologies help improve life quality International. *International Journal for Quality Research,* [e-journal] 12(4), pp.957-974. DOI: 10.18421/IJQR12.04-11.

Bruneel, J., D'Este, P. and Salter, A., 2010. Investigating the factors that diminish the barriers to university–industry collaboration. *Research policy*, [e-journal] 39(7), pp.858-868. DOI: 10.1016/j.respol.2010.03.006.

Bstieler, L., Hemmert, M. and Barczak, G., 2015. Trust formation in university-industry collaborations in the U.S. biotechnology industry: IP policies, shared governance, and champions. *Journal of Product Innovation Management*, [e-journal] 32(1), pp.111-121. DOI: 10.1111/jpim.12242.

Bunduchi, R., 2013. Trust, partner selection and innovation outcome in collaborative new product development. *Production Planning and Control*, [e-journal] 24(2-3), pp.145-157. DOI: 10.1080/09537287.2011.647868.

Carayannis, E.G. and Campbell, D.F.J., 2012. Mode 3 Knowledge Production in Quadruple Helix Innovation Systems. In: E.G. Carayannis and D.F.J. Campbell, eds. 2012. *Mode 3 Knowledge Production in Quadruple Helix Innovation Systems: 21st-Century Democracy, Innovation, and Entrepreneurship for Development*. New York: Springer. pp.1-63. DOI: 10.1007/978-1-4614-2062-0_1.

Criscuolo, P., Dahlander, L., Grohsjean, T. and Salter, A., 2017. Evaluating novelty: The role of panels in the selection of R&D PROJECTS. *Academy of Management Journal*, [e-journal] 60(2), pp.433-460. DOI: 10.5465/amj.2014.0861.

Cui, A. and Wu, F., 2016. Utilizing customer knowledge in innovation: antecedents and impact of customer involvement on new product performance', *Journal of the academy of marketing science*, [e-journal] 44, pp.516-538. DOI: 10.1007/s11747-015-0433-x.

de Klerk, S., 2012. Establishing and developing business networks: The importance of trust. *African Journal of Business Management*, [e-journal] 6(18), pp.5845-5856. DOI: 10.5897/AJBM11.2637.

De Silva, M., Howells, J. and Meyer, M., 2018. Innovation intermediaries and collaboration: Knowledge–based practices and internal value creation. *Research Policy*, [e-journal] 47(1), pp.70-87. DOI: 10.1016/J.RESPOL.2017.09.011.

de Wit-de Vries, E., Dolfsma, W.A., van der Windt, H.J. and Gerkema, M.P., 2018. Knowledge transfer in university–industry research partnerships: a review. *The Journal of Technology Transfer*, 44, pp.1236-1255. DOI: 10.1007/s10961-018-9660-x.

Dubina, I., Carayannis, E. and Campbell, D., 2012. Creativity economy and a crisis of the economy? Coevolution of knowledge, innovation, and creativity, and of the knowledge economy and knowledge society. *Journal of the Knowledge*, [e-journal] 3(1), pp.1-24. DOI: 10.1007/s13132-011-0042-y.

Edmondson, A.C., 2012. Teamwork on the fly. *Harvard Business Review*. [online] Available at: <https://hbr.org/2012/04/teamwork-on-the-fly-2> [Accessed 10 July 2020].

Edmondson, G., Valigra, L., Kenward, M., Hudson, R. and Belfield, H., 2012. *Making industry-university partnerships work: Lessons from successful collaborations*. Belgium: Business Innovation Board AISBL.

Eftimoski, D. and Milenkovski, B., 2012. The Knowledge Competitiveness of Macedonian Economy-Comparative Analysis. *Journal of Competitiveness,* [e-journal] 4(3), pp.122-135. DOI: 10.7441/joc.2012.03.09.

Ferreira Rebelo, M., Silva, R. and Santos, G., 2017. The integration of standardized management systems: managing business risk. *International Journal of Quality and Reliability Management*, [e-journal] 34(3), pp.395-405. DOI: 10.1108/IJQRM-11-2014-0170.

Frølund, L., Murray, F. and Riedel, M., 2018. Developing Successful Strategic Partnerships With Universities Developing Successful Strategic Partnerships With Universities. *MIT Sloan Management Review*. [online] Available at: <https://sloanreview.mit.edu/article/developing-successful-strategic-partnerships-with-universities/> [Accessed 10 July 2020].

Frow, P., Nenonen, S., Payne, A. and Storbacka, K., 2015. Managing Co-creation: A Strategic Approach to Innovation. *British Journal of Management*, [e-journal] 26(3), pp.463-483. DOI: 10.1111/1467-8551.12087.

Ghadimi, P., Ghassemi Toosi, F. and Heavey, C., 2018. A multi-agent systems approach for sustainable supplier selection and order allocation in a partnership supply chain. *European Journal of Operational Research*, [e-journal] 269(1), pp.286-301. DOI: 10.1016/j.ejor.2017.07.014.

Giesecke, J., 2012. The Value of Partnerships: Building New Partnerships for Success. *Journal of Library Administration*, [e-journal] 52(1), pp.36-52. DOI: 10.1080/01930826.2012.629964.

Grimpe, C., Goel, R. and Göktepe-Hultén, D., 2017. Who instigates university – industry collaborations ? University scientists versus firm employees. *Small Business Economics*, [e-journal] 48(3), pp.503-524. DOI: 10.1007/s11187-016-9795-9.

Guan, J. and Zhao, Q., 2013. The impact of university-industry collaboration networks on innovation in nanobiopharmaceuticals. *Technological Forecasting and Social Change*, [e-journal] 80(7), pp.1271-1286. DOI: 10.1016/j.techfore.2012.11.013.

Hartman, L.P. and Dhanda, K.K., 2018. Cross-Sector Partnerships: An Examination of Success Factors. *Business and Society Review*, [e-journal] 123(1), pp.181-214. DOI: 10.1111/basr.12139.

Hewitt-Dundas, N., Gkypali, A. and Roper, S., 2019. Does learning from prior collaboration help firms to overcome the "two-worlds" paradox in university-business collaboration ?. *Research Policy*, [e-journal] 48(5), pp.1310-1322. DOI: 10.1016/j.respol.2019.01.016.

Huang, Y. and Wilkinson, I., 2014. A case study of the development of trust in a business relation: Implications for a dynamic theory of trust. *Journal of Business Market Management,* 7(1), pp.254-279.

Champenois, C. and Etzkowitz, H., 2018. From boundary line to boundary space: The creation of hybrid organizations as a Triple Helix micro-foundation. *Technovation*, [e-journal] 76-77, pp.28-39. DOI: 10.1016/j.technovation.2017.11.002.

Chi, M., Wang, W., Lu, X. and George, J.F., 2018. Antecedents and outcomes of collaborative innovation capabilities on the platform collaboration environment. *International Journal of Information Management*, [e-journal] 43, pp.273-283. DOI: 10.1016/j.ijinfomgt.2018.08.007.

Jarábková, J., Chreneková, M. and Roháčiková, O., 2019. University and practice – cooperation in research and science: case study of the slovak university of agriculture in Nitra. *Quality Innovation Prosperity*, [e-journal] 23(1), pp.136-154. DOI: 10.12776/QIP.V23I1.1168.

Johnston, A. and Huggins, R., 2016. Drivers of University–Industry Links: The Case of Knowledge-Intensive Business Service Firms in Rural Locations. *Regional Studies*, [e-journal] 50(8), pp.1330-1345. DOI: 10.1080/00343404.2015.1009028.

Jonkers, K., Tijssen, R., Karvounaraki, A. and Goenaga Beldarrain, X., 2018. *A Regional Innovation Impact Assessment Framework for universities*. [pdf] Publications Office of the European Union. Available at: <https://ec.europa.eu/jrc/en/publication/eur-scientific-and-technical-research-reports/regional-innovation-impact-assessment-framework-universities> [Accessed 10 July 2020].

Jørgensen, F. and Ulhøi, J., 2010. Enhancing innovation capacity in SMEs through early network relationships. *Creativity and Innovation Management*, [e-journal] 19(4), pp.397-404. DOI: 10.1111/j.1467-8691.2010.00577.x.

Konsti-Laakso, S., Pihkala, T. and Kraus, S., 2012. Facilitating SME innovation capability through business networking. *Creativity and Innovation*, [e-journal] 21(1), pp.93-105. DOI: 10.1111/j.1467-8691.2011.00623.x.

Le Roy, F., Robert, M. and Lasch, F., 2016. Choosing the Best Partner for Product Innovation. *International Studies of Management and Organization*, [e-journal] 46(2-3), pp.136-158. DOI: 10.1080/00208825.2016.1112148.

Liao, S.-H., Hu, D.-Ch. and Ding, L.-W., 2017. Assessing the influence of supply chain collaboration value innovation, supply chain capability and competitive advantage in Taiwan's networking communication industry. *International Journal of Production Economics*, [e-journal] 191, pp.143-153. DOI: 10.1016/j.ijpe.2017.06.001.

Lin, J.-Y., 2017. Balancing industry collaboration and academic innovation : The contingent role of collaboration-specific attributes. *Technological Forecasting & Social Change*, [e-journal] 123, pp.216-228. DOI: 10.1016/j.techfore.2016.03.016.

Merritt, C.C. and Kelley, D.C., 2018. What individual and organizational competencies facilitate effective collaboration? Findings from a collaborative governance simulation. *Journal of Public Affairs Education*, [e-journal] 24(1), pp.97-121. DOI: 10.1080/15236803.2018.1429812.

Mindruta, D., Moeen, M. and Agarwal, R., 2016. A Two-sided Matching Approach for Partner Selection and Assessing Complementarities in Partners' Attributes in Inter-firm Alliances. *Strategic Management Journal,* [e-journal] 37(1), pp.206-231. DOI: 10.1002/smj.2448.

Nguyen, N. and Liem, N., 2013. Inter-firm trust production: Theoretical perspectives. *International Journal of Business and Management,* [e-journal] 8(7), pp.46-54. DOI: 10.5539/ijbm.v8n7p46.

Orviská, M., Huňady, J., Pisár, P. and Hudson, J., 2019. From academic publications and patents to the technological development of the economy: Short and long run causalities. *Quality Innovation Prosperity*, [e-journal] 23(1), pp.74-101. DOI: 10.12776/QIP.V23I1.1166.

Paľová, D., Czaja, A. and Vejačka, M., 2018. Innovative approach to education improvement via enterprise-education collaboration. *Quality Innovation Prosperity*, [e-journal] 22(3), pp.68-82. DOI: 10.12776/QIP.V22I3.1171.

Perkmann, M. and Salter, A., 2012. How to Create Productive Partnerships With Universities. *MIT Sloan Management Reviews*, [online] Available at: <https://sloanreview.mit.edu/article/how-to-create-productive-partnerships-with-universities/> [Accessed 10 July 2020].

Perkmann, M. and Schildt, H., 2015. Open data partnerships between firms and universities: The role of boundary organizations. *Research Policy*, [e-journal] 44(5), pp.1133-1143. DOI: 10.1016/j.respol.2014.12.006.

Piacentini, A., 2013. *The University – Industry Partnership. A new format to promote innovation.* Australia: Government of South Australia.

Plewa, C., Korff, N., Johnson, C., Macpherson, G., Baaken, T. and Rampersad, G., 2013. The evolution of university–industry linkages—A framework. *Journal of Engineering and Technology Management,* [e-journal] 30(1), pp.21-44. DOI: 10.1016/j.jengtecman.2012.11.005.

Rajalo, S. and Vadi, M., 2017. University-industry innovation collaboration: Reconceptualization. *Technovation*, [e-journal] 62–63, pp.42-54. DOI: 10.1016/J.TECHNOVATION.2017.04.003.

Ranga, M. and Etzkowitz, H., 2013. *Triple Helix systems: an analytical framework for innovation policy and practice in the Knowledge Society*, *Industry and Higher Education.* Industry and Higher Education, [e-journal] 27(4), pp.237-262. DOI: 10.5367/ihe.2013.0165.

Ranga, M., Hoareau, C., Durazzi, N., Etzkowitz, H., Marcucci, P. and Usher, A., 2013. *Study on university-business cooperation in the US.* London: LSE Enterprise.

Rebelo, M.F., Santos, G. and Silva, R., 2015. Integration of management systems: Towards a sustained success and development of organizations. *Journal of Cleaner Production*, [e-journal] 127, pp.99-111. DOI: 10.1016/j.jclepro.2016.04.011.

Reficco, E., Gutiérrez, R., Jaén, M.H. and Auletta, N., 2018. Collaboration mechanisms for sustainable innovation. *Journal of Cleaner Production*, [e-journal] 203, pp.1170-1186. DOI: 10.1016/j.jclepro.2018.08.043.

Rehák, Š., Šipikal, M. and Lešková, A., 2019. Evolution of university third mission activities in slovakia: What role for a public policy?. *Quality Innovation Prosperity*, [e-journal] 23(1), pp.119-135. DOI: 10.12776/QIP.V23I1.1165.

Reichert, S., 2019. *The Role of Universities in Regional Innovation Ecosystems.* Geneva: European University Association.

Resetarits, A. and Resetarits-Tincul, A.-M., 2012. Fuzzy Concepts—A New Approach in the Description of Boundaries as Creative Knowledge Environments in Educational Sciences. *Journal of the Knowledge Economy*, [e-journal] 3(1), pp.53-67. DOI: 10.1007/s13132-011-0050-y.

Roud, V. and Vlasova, V., 2018. Strategies of industry‐science cooperation in the Russian manufacturing sector. *The Journal of Technology Transfer*, [e-journal] 45(3), pp.870-907. DOI: 10.1007/s10961-018-9703-3.

Santos, G., Murmura, F. and Bravi, L., 2019. Developing a model of vendor rating to manage quality in the supply chain. *International Journal of Quality and Service Sciences,* [e-journal] 11(1), pp.34-52. DOI:10.1108/IJQSS-06-2017-0058.

Schilke, O. and Cook, K., 2013. A cross-level process theory of trust development in interorganizational relationships. *Strategic Organization*, [e-journal] 11(3), pp.281-303. DOI: 10.1177/1476127012472096.

Smirnova, M.M., Rebiazina, V.A. and Khomich, S.G., 2018. When does innovation collaboration pay off ? The role of relational learning and the timing of collaboration. *Industrial Marketing Management*, [e-journal] 74, pp.126-137. DOI: 10.1016/j.indmarman.2017.11.001.

Sum Chau, V., Gilman, M. and Serbanica, C., 2016. Aligning University-Industry Interactions: the Role of Boundary Spanning in Intellectual Capital Transfer.Technological Forecasting and Social Change. *Kent Academic Repository*, [e-journal] 123, pp.199-209. DOI: 10.1093/bjc/azq038.

Thune, T. and Gulbrandsen, M., 2014. Dynamics of collaboration in university–industry partnerships: do initial conditions explain development patterns?. *Journal of Technology Transfer*, [e-journal] 39(6), pp.977-993. DOI: 10.1007/s10961-014-9331-5.

Un, C.A. and Rodríguez, A., 2018. Local and Global Knowledge Complementarity : R & D Collaborations and Innovation of Foreign and Domestic Firms. *Journal of International Management*, [e-journal] 24(2), pp.137-152. DOI: 10.1016/j.intman.2017.09.001.

Unger, M., Marsan, G.A., Meissner, D., Polt, W. and Cervantes, M., 2018. New challenges for universities in the knowledge triangle. *The Journal of Technology Transfer*, [e-journal] 45, pp.806-819. DOI: 10.1007/s10961-018-9699-8.

Veles, N., Carter, M.-A. and Boon, H., 2018. Complex collaboration champions: university third space professionals working together across borders. *Perspectives: Policy and Practice in Higher Education*, [e-journal] 23(2-3), pp.1-11. DOI: 10.1080/13603108.2018.1428694.

Wu, I.L. and Chiu, M.L., 2015. Organizational applications of IT innovation and firm's competitive performance: A resource-based view and the innovation diffusion approach. *Journal of Engineering and Technology Management*, [e-journal] 35, pp.25-44. DOI: 10.1016/j.jengtecman.2014.09.002.

Yoon, C., Lee, K., Yoon, B. and Toulan, O., 2017. Typology and success factors of collaboration for sustainable growth in the IT service industry. *Sustainability,* [e-journal] 9(11), pp.1-20. DOI: 10.3390/su9112017.

ABOUT AUTHORS

Arlindo Marinho – (A.M.) Universidade Lusíada – Norte, Portugal, Ph.D. Student, e-mail: arlindo_marinho@hotmail.com, Author's ORCID: 0000-0003-3668-9914.

Rui G. Silva – (R.G.S.) Universidade Lusíada – Norte, Portugal, Engineering and Technology Faculty, Assoc. Prof., e-mail: rgasilva@gmail.com, Author's ORCID: 0000-0002-7929-0367.

Gilberto Santos – (G.S.) Polytechnic Institute Cávado Ave (IPCA), Barcelos, Portugal, Prof., e-mail: gsantos@ipca.pt, Author's ORCID: 0000-0001-9268-3272.

AUTHOR CONTRIBUTIONS

A.M. – conceptualization, investigation, writing - original draft; R.G.S., G.S. – supervision; A.M., R.G.S., G.S. – methodology, writing - review & editing.

CONFLICTS OF INTEREST

The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Enterprise Risk Management – Approaches Determining Its Application and Relation to Business Performance

DOI: 10.12776/QIP.V24I2.1467

Jozef Klučka, Rudolf Grünbichler

Received: 2020-06-04 Accepted: 2020-06-30 Published: 2020-07-31

ABSTRACT

Purpose: As management systems, enterprise risk management and enterprise performance management pursue similar objectives and influence each other positively. The paper aims to provide an insight into the relationship between Enterprise Risk Management and Business Performance Management.

Methodology/Approach: The paper compares the results of an American study with the results of a Slovakian study. First, the American results are cited and interpreted. Then the Slovak results are presented and discussed. Then the results are compared. In the last part an overall conclusion is drawn, the relationships between the results are shown and practical implications are explained.

Findings: The results show similarities, but also differences to Enterprise Risk Management and the relationships between Enterprise Risk Management and Business Performance Management. The paper shows that there are differences in both the management approach and the impact on business performance between American and Slovak companies.

Research Limitation/implication: A limitation in both studies is the limited number of participating companies. This is accompanied by a higher probability of error.

Originality/Value of paper: The paper provides new information to the gap related to subjects enterprise risk management and business performance management and their relations.

Category: Research paper

Keywords: enterprise risk management; business performance management; empirical survey

1 INTRODUCTION

Enterprise Risk Management (ERM) is the sum of business activities focused on identifying, influencing and actively applying risks to achieve business objectives. The risks can deviate both positively and negatively from the planned state. Risk management is performance-related, aims at the growth of the company and serves to prepare a better basis for decision-making.

There is a wealth of literature on ERM, much of which is available in foreign languages. The Slovak literature and practice is determined by the fact that the application is not legally established and is not mandatory. Therefore, in Slovakia the topic is treated more often in theory than in practice. In fact, risk management is constantly applied in practice. Its application is more intuitive and less based on methods. The level of risk management also depends strongly on the industry in which the company operates and the size of the company. For this reason, it is possible to identify significant differences in the application of ERM in Slovakia in the banking and insurance sector and in other sectors as well as in large or small and medium-sized enterprises. Although studies have shown a link between risk and performance management, it is still a little-noticed topic (e.g. Rabechini and de Carvalho, 2013; De Bakker, Boonstra and Wortmann, 2012).

In 2015, we conducted a questionnaire survey among Austrian and Slovak small and medium-sized enterprises with the aim of surveying the current status of ERM among SMEs in both countries (Klučka and Grünbichler, 2016; Klučka, Grünbichler and Havko, 2017; Klučka, 2018). In parallel, other surveys on the status quo of ERMs in other countries were also conducted. For the purpose of this article we compare the results of the statistical survey in Slovakia with the statistical results published in the Disaster Recovery Journal (Belaouras, 2019). This source has been used for reasons of time actuality of data. The differences and similarities between companies from Slovakia and the USA are highlighted.

In America, companies from different sectors in North America were contacted. They were sent a questionnaire and 55 companies participated in the study. Of these, 44% had sales of less than USD 500 million. 40% of the respondents had between one and less than a thousand employees. In Slovakia, a standardized questionnaire was also created and sent to Slovakian companies. The questionnaire contained a wide range of ERM-relevant topics, such as the implementation and application of risk management within the company. The total number of companies surveyed was 162; however, not all questions from the companies were answered, so the number of answers per question varied. About 95% of the companies surveyed fall into the category of small and medium-sized enterprises.

2 METHODOLOGY

In this paper a qualitative approach is chosen. The procedure is designed in such a way that first the results of the American survey are processed and interpreted

and then the Slovakian results are presented and commented on. In the next part, measures are derived from the commented results for practical application and presented. Subsequently, the relationships between the management systems are shown. Those factors are identified which are covered by the management activities with regard to both risk management and performance management.

3 RESULTS

3.1 Use of Risk Management in Companies and Organizational Anchoring

The first question asked of American companies was to determine how Enterprise Risk Management is structured in the company. For this purpose, the companies were given several statements to choose from. The most frequent answer was that a formally established ERM program with the function of a chief risk officer was in place (36%). The second most common response was that there is a single director or head of risk who is responsible for selected risk areas, but does not cover the entire spectrum of an ERM program (25%).

The Slovakian questionnaire asked a question about the use of risk management by asking about rules and regulations of risk management in the company. Several answers were given for selection (multiple answers possible). 42% of the respondents stated that they have no written documentation of risk management in the company. Of the remaining answers, 33%, 31% and 28% stated that they have risk management documentation in the company as part of the internal control, organisational system or quality management documentation of the company. In another question, it was asked who in the company is responsible for risk management. In contrast to the American results, the responsibility for risk management in Slovakian companies is considered to be a task of the management (65%).

This implies that the issue of ERM is formally unregulated in Slovak companies, which is also reflected by the fact that there is no risk management position in the organizational structure outside the bank institutions and insurance companies, as well as in the vast majority of small and medium-sized enterprises.

3.2 Reporting of Risks and the Need for Risk Management Systems Adapted for SMEs

Another question was asked to which manager or department the most senior risk officer reports. Of the possible answers, this is the Chief Risk Officer (CRO) with 29%. This is followed by the Chief Information Security Officer (11%), others (11%) and Chief Finance Officer (CFO) (9%).

The Slovak version of the survey included one question regarding responsibility for risk management in the company. The respondents' answers have revealed a

big preference of the management or owner (65%), who is responsible for the risk management or the head of the department (25%). 8% of respondents' reactions were assigned to the risk manager.

The answer to this question shows a clear difference in risk management. In Slovak small and medium-sized enterprises, management, which includes risk management, is the task of the owner or head of unit, not a specialist with assigned competence. Especially small and medium-sized enterprises are affected by risks to a high degree and usually do not have the financial and human resources to assess and manage risks. In practice, this means that risk management systems have to be adapted to the needs of small and medium-sized enterprises so that they too can carry out risk management systematically.

However, the lack of a risk management system is not only an issue in Slovakian SMEs. Due to the lack of human and capital resources in most cases, it is a general issue. In general, small and medium-sized enterprises do not use a formalised risk management system, but individual key figures (Taticchi, Tonelli and Cagnazzo, 2010). Here, the need for research to develop appropriate systems for SMEs is evident.

When asked in the American questionnaire to which level the highest ranking employee reports, the most frequent answer was to the C-level or equivalent head of functional area (e.g. COC) (68%). For Slovakian conditions, it is clear from the above that the absence of a written report predominates and at the same time it is assumed that the boss (CEO) or head of department is responsible for risk management. It follows that this question is irrelevant for Slovakian circumstances. In Slovakia, it was also asked whether companies employ staff who are trained or educated in the field of risk management. Up to 75% of respondents stated that they do not have such employees. Here, too, the potential for optimisation in connection with risk management is evident in practice.

3.3 ERM Responsibilites and Perceived Risks

Another question included in the questionnaire was to what extent the ERM system or ERM efforts are responsible for selected risk areas. The participating companies were able to choose to rate the risk areas on an ordinal scale. Table 1 shows the distribution to the options fully and mostly responsible (Balaouras, 2019, p.47).

The Slovakian questionnaire asked which risks are relevant for the companies (multiple answers possible). In Slovakia, the most frequent risks were indicated: departure of key employees (28%) and risks emanating from the state or political decisions or risks (21%). Other risks mentioned are risks arising from the development of raw material prices or risks arising from liability for damages caused.

Table 1 – Results of Risk Assignment

	Fully responsible	Mostly responsible
Operational risk	27%	30%
Compliance risk	24%	16%
Reputational risk	22%	19%
Financial risk	27%	11%
Strategic risk	22%	16%

Given the different formulation of the question, it is not possible to establish a clear comparison. However, with some generalization, it can be stated that, like US companies, Slovak companies have identified operational risk and events that affect operations.

3.4 ERM and Business Continuity (BC)

In an american enterprise where the occurrence of risk is associated with high financial / non-financial implications, in addition to the ERM agenda covered by the institutional department and its own procedures, there is also a business continuity management (BCM) agenda. The question concerns the cooperation of ERM and BCM departments/people. In response, 31% of respondents said that the BCM team is working closely with the ERM team. Very close (29%) was the second response, stating that BCM team reports are directed directly to the ERM team (department).

In the survey, which was addressed to Slovak companies, the question was formulated whether there is also a business continuity plan (based on the emergency plan). Up to 67% of respondents answered that they had no problems with business continuity. It is important that a Slovak SME is legally obliged to process e.g. an evacuation plan, but the term BCM is not used terminologically. It is not part of Slovak legal norms to process business continuity plans. The professional public in small and medium-sized enterprises is usually not familiar with the contents and tasks related to the BCM agenda.

3.5 Risks in General and due to Critical Events

The question, whose answer may illustrate the importance of the ERM agenda, has focused on the frequency of critical risks over the past 3 years. Critical risks were understood e.g. shooting, natural disasters, IT dysfunction with impact on traffic, cyber-attacks. In response, 38% of respondents stated that no critical events had occurred in the past 3 years and 32% of respondents reported that these events were in the 1-5 range; 12% of respondents reported the occurrence of critical events in the interval 11 to 20.

The last question concerned the experience with risks in American companies and their occurrence in the last 3 years. The top 10 were (Balaouras, 2019, p.48):

- IT failure a business-critical system or application
- Extreme weather or natural disaster
- Theft of intellectual property
- Cyberattack (data breach, ransomware, DDoS attack, etc.)
- Critical infrastructure failure (power, water, transportation, etc.)
- Customer privacy abuse, data breach, or fraud
- Supply chain disruption/failure
- Geopolitical events/social unrest
- Customer backlash/adverse media exposure/social activism
- Workplace misconduct

The latter two questions were not part of the questionnaire aimed at small and medium-sized enterprises in Slovakia. In Slovakia, the question was asked which risks were considered most significant at the time of the survey. The answers showed that the most important risks are 67% competition and 42% problems with customers.

4 CONCLUSION

The relation between risk and performance is theoretically known (e. g. Aureli and Salvatori, 2012) and also applied in practice by managers of Slovak companies. However, it is rather an intuitive approach, which is the responsibility of the company's management and the relevant line manager. Nevertheless, it is not explicitly stated in the tasks and obligations of these employees. In fact, given the circumstances of the transforming post-socialist economy, Slovak managers had to implement the risk in their management decisions. As the questionnaire suggests, risk management was rather intuitive, without data support and appropriate methods, know-how, and trained employees, providing background information for management decisions. Most importantly, risk management is not explicitly established as a legal or business culture obligation. Thus, risk management (as well as BCM issues) in Slovakia is reduced to companies that have to process these, as their parent company abroad requires this.

ACKNOWLEDGEMENTS

This work was supported by the Slovak Research and Development Agency under the contract No. Sk-AT – 2017-0003.

REFERENCES

Aureli, S. and Salvatori, F., 2012. An investigation on possible links between risk management, performance measurement and reward schemes. *Accounting and Management Information Systems*, 11(3) pp.306-334.

Balaouras, S., 2019. State of Enterprise Risk Management 2019. *Disaster Recovery Journal,* Spring 2019, pp.46-50.

De Bakker, K., Boonstra, A. and Wortmann, H., 2012. Risk managements' communicative effects influencing IT project success. *International Journal of Project Management*, 30(4), pp.444-457.

Klučka, J., 2018. Business performance management in small and medium enterprises. In: L. Štofová and P. Szaryszová, eds., *New trends in process control and production management: Proceedings of the international conference on marketing management*, *trade, financial and social aspects of business (MTS 2017).* Košice, Slovakia, and Tarnobrzeg, Poland, 18-20 May 2017. Leiden: Balkema. pp.259-262.

Klučka, J., Grünbichler, R. and Havko, J., 2017. The risk management in small and medium enterprises: its meaning and expectations (comparison between Austria and Slovakia). In: K. Kadvanová et al., eds., Zborník príspevkov zo 17. medzinárodnej vedeckej konferencie: Medzinárodné vzťahy 2016: Aktuálne otázky svetovej ekonomiky a politiky. Smolenice, Slovakia, 1-2 December 2016. Bratislava: Ekonomická univerzita. pp.374-383.

Klučka, J. and Grünbichler, R., 2016. Risikomanagement: Verbreitung, Bedeutung und zukünftige Erwartungen: Ein Vergleich zwischen Österreich, Slowakei und Deutschland. *Controller Magazin: Arbeitsergebnisse aus der Controller-Praxis*, 41(5), pp.49-54.

Rabechini, R. and de Carvalho, M., 2013. Understanding the impact of project risk management on project performance: an empirical study. *Journal of Technology, Management & Innovation*, 8, pp.64-78.

Taticchi, P., Tonelli, F. and Cagnazzo, L., 2010. Performance measurement and management: a literature review and a research agenda. *Measuring Business Excellence*, 14(1), pp.4-18.

ABOUT AUTHORS

Jozef Klučka – (J.K.) University of Žilina, Slovakia, Faculty of Security Engineering, Department of Crisis Management, Assoc. Prof., e-mail: jozef.klucka@fbi.uniza.sk, Author's ORCID: 0000-0002-0994-6439.

Rudolf Grünbichler – (R.G.) University of Applied Sciences CAMPUS 02, Austria, Department of Financial and Management Accounting, Prof. (FH), e-mail: rudolf.gruenbichler@lv.campus02.at, Author's ORCID: 0000-0002-5872-5771.

AUTHOR CONTRIBUTIONS

J.K. – methodology, resources, data curation, supervision; R.G. – formal analysis, writing—original draft preparation, writing—review and editing, visualization, project administration; J.K., R.G. – conceptualization, funding acquisition.

CONFLICTS OF INTEREST

The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

A Scale for Measuring Sustainable Manufacturing Practices and Sustainability Performance: Validity and Reliability

DOI: 10.12776/QIP.V24I2.1385

Yousif M. Ibrahim, Norsiah Hami, Susan S. Abdulameer

Received: 2020-01-18 Accepted: 2020-03-24 Published: 2020-07-31

ABSTRACT

Purpose: In quantitative studies, providing a valid and reliable instrument is necessary to ensure accurate results when measuring sustainable manufacturing practices (SMPs) and sustainability performance (SP). Therefore, this study aims to evaluate the validity and reliability of the measurements of SMPs and SP.

Methodology/Approach: The population of this study is top managers who have experience in the oil and gas industry (O&GI) in Iraq. Two tests were carried out in the present study: the pre-test and the pilot test.

Findings: In the pre-test, the comments made by six academician experts and three practitioners were used to rephrase the measurements items and modify them according to the requirements of the O&GI by the Iraqi context. Moreover, in the pilot test was all the items were reliable and were sufficiently correlated with their constructs.

Research Limitation/implication: There are some limitations to the current study. First, due to the small size of the study population, the pilot test sample in the current study was only 12 respondents. Future researchers can increase the sample size for the pilot test when they have a large population. Second, the validity and reliability of the measurements were tested in the current study in the oil and gas industry only. Future studies can test these measurements in other industries or small and Small and medium-sized enterprises (SMEs).

Originality/Value of paper: Theoretically, there are four contributions to the body of knowledge: first, introduce measures for SMPs according to the product life cycle view, it is limited in the literature. Second, these measures can be used by researchers to study the extent of SMPs and the SP of companies using descriptive statistics. Third, these measures can be used to investigate the impact of SMPs on SP by regression testing or structural equation modelling. Fourth, measures can be modified into open-ended questions for use in qualitative or

mixed studies. Practically, there are two practical implications which explain in the conclusion.

Category: Research paper

Keywords: sustainable manufacturing practices; sustainability performance; pre-test; pilot test; oil and gas industry

1 INTRODUCTION

Sustainability performance (SP) become an essential issue and a significant concern in the oil and gas industry (O&GI) in Iraq (Ibrahim, Hami and Othman, 2019). This is because of the imbalance among the dimensions of SP that includes: economic, environmental and social sustainability. For example, the report of the ESCWA reported the proportion of Iraqi exports of oil equivalent to 99% of the total annual exports (UN-ESCWA, 2018). This establishes the importance of this industry in the economic development in Iraq. Nonetheless, this industry considers the main reason for environmental emissions and social damage (Elhuni and Ahmad, 2017).

Actually, due to their nature and size, the O&GI has main impacts of health, safety global and environment (Schneider et al., 2011). Also, particulate matter and volatile compounds of filters in oil and gas companies cause many diseases, both for workers and the community in the same area, such as cancer diseases and respiratory diseases (EPA, 2003). Moreover, the main areas for the Extract and production of oil in Iraq, 70% of them have pollution matters in the environment and involve areas such as Kirkuk, Maysan, Basrah, Salah al-Din, Baghdad and Mosul (Al-Haleem, Awadh and Saeed, 2013). Nevertheless, to obtain a balance among the pillars of SP, there should be sustainable practices and activities in the O&GI.

In this respect, empirical evidence in literature has confirmed that sustainable manufacturing practices (SMPs) according to the product life cycle view: sustainable product design (SPD), sustainable manufacturing process (SMP), sustainable supply chain management (SSCM) and sustainable end-of-life management (SEoLM) (Abdul-Rashid et al., 2017a), improve economic, environmental and social sustainability and thus balance it. To achieve this improvement and balance, there should be measurements valid and reliable regarding SMPs and SP.

According to Creswell and Creswell (2018), in quantitative studies, when there is any adapt on one or combine measurements, the prior validity and reliability may not apply for the new measurements. Hence, it has become significant to assessed new validity and reliability for the new measurements before conducting the main study. Therefore, this study aims to evaluate the validity and reliability of the measurements of SMPs and SP among the O&GI in Iraq. The results of this study can be beneficial in several aspects. The researchers will

have a valid and reliable instrument to measure SMPs and SP, particularly in the O&GI. Besides, top management and managers in the O&GI will obtain a deeper perception of how to measure SMPs and SP. Theoretically, there are four contributions to the body of knowledge: first, introduce measures for SMPs according to the product life cycle view, it is limited in the literature. Second, these measures can be used by researchers to study the extent of SMPs and the SP of companies using descriptive statistics. Third, these measures can be used to investigate the impact of SMPs on SP by regression testing or structural equation modelling. Fourth, measures can be modified into open-ended questions for use in qualitative or mixed studies. Practically, there are two practical implications. First, managers can measure their companies' implementation of SMPs using a 6-point Likert scale. Second, measure the level of SP achieved and compare with the performance in previous years.

2 METHODOLOGY

The O&GI is one of the most top sectors that largely contributed to the GDP of Iraq (OPEC, 2018). This contribution is very important, especially in employment opportunities and exports. Consequently, the population of this study is top managers and senior executives who have experience in the O&GI in Iraq. In parallel to the enormous contribution to the GDP of the country and its large contribution to the environmental and social impacts to the nation because its harmful operational activities. on the other side, the pre-test was conducted to verify the face validity (Hair et al., 2013). Then, the pilot test was carried out to establish the reliability of the measurements used in the current study (Saunders, Lewis and Thornhill, 2016).

Scaling design of the items will be used on a "6-point Likert scale": "1" = "Strongly Disagree" (SD); "2" = "Moderately Disagree" (MOD); "3" = "Slightly Disagree" (SLD); "4" = "Slightly Agree" (SLA); "5" = "Moderately Agree" (MOA); and "6" = "Strongly Agree" (SA). The reason for using the 6-point Likert scale was to ensure that participants did not easily check the "indifference" option or "midpoint", as usually occur with a 5-point scale. The midpoint means the neutral response when answering the questionnaire with of exist an odd number of categories used in a scale (Hair et al., 2017). He also emphasised that the researcher usually uses the scale without the midpoint when many respondents are expected to choose neutrals on a particular issue. This is because it provides an easy option that needs a few efforts and is easily justified (Krosnick and Fabrigar, 1997). Garland (1991) argued that the participants would answer based on the content of the questions when given an even number of a response scale. Additionally, participants from Asian countries tend to choose the middle category response than those from Western countries (Si and Cullen, 1998; Thrulogachantar and Zailani, 2011). It was also found that the validity and reliability of the findings tend to be higher for the even number response scale a six-point in particular (Chomeya, 2010) when compared to the odd number

response scale (Krosnick and Fabrigar, 1997; Andrews, 1984; Alwin and Krosnick, 1991; Birkett, 1986; Coelho and Esteves, 2007).

The designed questionnaire was divided into three sections, which were the first section: focuses on SMPs implemented by oil and gas companies in Iraq, second section: focuses on identifying SP that can be achieved through the implementation of SMPs and the third section: provide a profile of the company and personal. Appendix (Table A1) states the items in the first and second sections of the questionnaire and the references adapted from them.

3 RESULTS AND DISCUSSION

3.1 Pre-test (Validity)

The pre-test process involves face validity. The face validity is done through systematic assessment of the measurement based on subjective judgment by the experts (Hair et al., 2013) to verify the measurement's ability to measure what it is supposed to measure in the study (Hair et al., 2017). They also pointed out that this validation method is commonly used in management and business research. Therefore, the study measurements that adapted from previous studies for the SMPs and SP were sent to six experts who are familiar with the constructs of this study to attest the face validity of the measurements. Additionally, three oil and gas industry's practitioners were also contacted for the same purpose. Their feedback, recommendations and comments have been made. Results of face validity by the experts in the pre-test is shown in Table 1.

Table 1 – Results of Face Validity by Experts in the Pre-Test

Expert Type	Variable	Comment	Action
Academicians	SMPs	• Add the words "Our company practices" at the below of each dimension of sustainable manufacturing practices.	• *"Our company practices".*
		• Modify the item by changing the word "Eliminating" to "Eliminates" in the item of "Eliminating the use of hazardous materials during the design of the products".	• *"Eliminates the use of hazardous materials during the design of the products".*
		• Modify the item by changing the worded "Design the products which will prolong the life of materials".	• *"Design the products which will prolong its lifetime".*
		• Modify the item by changing the words "Savings of energy" to "Save energy" in the item of "Savings of energy during the manufacturing process".	• *"Save energy during the manufacturing process".*
		• Modify the item by changing the word "process" to "processes" in the item of "Utilise lean production process".	• *"Utilise lean production processes".*
		• Modify the item by changing the word	• *"Adopts of sustainable*

Expert Type	Variable	Comment	Action
		"Adoption" to "Adopts" in the item of "Adoption of sustainable suppliers".	*suppliers".*
		• Modify the item by changing the word "Using" to "Use" in the item of "Using a less, cleaner or reusable packaging".	• *"Use a less, cleaner or reusable packaging".*
		• Modify the item by changing the worded "Providing recycling support using components and material coding standards".	• *"Provide recycling support for materials and components used".*
Academicians	SP	• Summarising the existing statement below the sustainability performance then put it at the below of each dimension of performance.	• *"In the last three years, please describe your company's achievements for economic performance caused by the current practices (as you described in sections one and two)".*
		• Delete "s" from "emissions" and add "es" to "gas" in the item "Reduced emissions of greenhouse gas".	• *"Reduced emission of greenhouse gases".*
		• Delete "s" from "wastes" in the item "Reduced solid wastes".	• *"Reduced solid waste".*
		• Delete "s" from "wastes" in the item "Reduced liquid wastes".	• *"Reduced liquid waste".*
Practitioners	Profile of company and personal	• Change the options "Private/ local" and "Private/ foreign" in the question "What is ownership of your company?" in section three to one option only as "Private" and add one more option as "Foreign".	• Done.
		• Add the options as "OHSAS 18001", "ISO 29001" and "All" in the question "Does your company have the following certifications?" in section three.	• Done.
		• Add the option as "General Manager" and "Chief executive officer" in the question "What is your current position in your company?" in section three.	• Done.

3.2 Pilot Test (Reliability)

After the questionnaire is constructed (Sekaran and Bougie, 2016; Kumar, 2014), it is necessary to test it before using it to actual data collection (Dawson, 2009; Oppenheim, 2000; Adams, Khan and Raeside, 2014; Fink, 2017). This is because, without a trial test, we will not be able to tell if the questionnaire will succeed (Saunders, Lewis and Thornhill, 2016). Moreover, because this study adapted the measurements from different sources (Hair et al., 2014) regarding the constructs of SMPs and SP.

Data collection in the research process usually begins with a pilot test (Cooper and Schindler, 2014). Saunders, Lewis and Thornhill (2009) defined a pilot test as "small-scale study to test a questionnaire, interview checklist or observation schedule, to minimise the likelihood of respondents having problems in answering the questions and of data recording problems as well as to allow some assessment of the questions' validity and the reliability of the data that will be collected".

There are many essential purposes for conducting the pilot test include understanding or interpreting questions by respondents (Kumar, 2011; Sekaran and Bougie, 2010) and clarity of wording of the questions and estimate achievement times (De Vaus, 2002; Adams, Khan and Raeside, 2014). Also, the pilot test will help to clarify the extent the flow and sequences of questions (Bryman and Bell, 2015; Oppenheim, 2000), as well as it will enable to get some evaluation of the validity of the questions and the potential reliability of the data to be collected (Saunders, Lewis and Thornhill, 2016). In order to achieve the purposes of the pilot test, Bell and Waters (2014) suggested to give the respondents a short questionnaire attached to the original questionnaire of the study includes a set of questions to know the following: (1) How long did it take to complete the questionnaire?, (2) were the questionnaire instructions clear?, (3) are there any unclear or vague questions? If there is, please specify, and why?, (4) do you have any objection to answering any question?, (5) do you think any significant topic has been deleted?, (6) do you think the layout of the questionnaire is clear/ attractive? and (7) any other comments?

Indeed, to make sure the questionnaire has achieved the purposes mentioned above (Oppenheim, 2000), the pilot test should be conducted with respondents who are similar to those that will be used in the full study (Saunders, Lewis and Thornhill, 2016; Zikmund et al., 2013; Hair et al., 2014). Naturally, the closer the link between the pilot sample and the final sample, the better (De Vaus, 2002). Given this, the reliability test must be carried out.

Principally, reliable measurements mean that they achieve the same result on repeated occasions (De Vaus, 2002). Cronbach's alpha was used for this purpose based on the recommendations of several researchers (e.g., DeVellis, 2016; Saunders, Lewis and Thornhill, 2016; Colton and Covert, 2007). Cronbach's alpha ranges from 0 to 1, the lowest acceptance value 0.70 (Hair et al., 2013). Moreover, the item analysis method was used by corrected item-total correlation test to estimate the reliability of responses within an instrument (Colton and Covert, 2007; Field, 2009), as well as, explains the most correlated items with the construct, meanwhile the value of any item is less than .30, it is deleted (Nunnally and Bernstein, 1994; Field, 2009; Hair et al., 2017). Bradburn, Sudman and Wansink (2004) recommended that no more than 10-12 participants are sufficient to detect the difficulties and weaknesses in the pilot test questionnaire. Likewise, Van Belle (2008) recommended that the sample size for the pilot test should be not a minimum of 12 participants. Besides, Fink (2013) stated that the minimum number of participants for a pilot test is 10 (cited in

Saunders, Lewis and Thornhill, 2016). In short, the literature has illustrated that the sample of the pilot test is few and is not considered an issue either in quantitative or qualitative studies (Khattab and Wahid, 2015). Thus, 12 sets of questionnaires were distributed; all the questionnaires that were returned were usable. Using SPSS.V.25, the reliability of the measurements and the item analysis was analysed (Field, 2013; Pallant, 2011), as displayed in Appendix (Table A2).

Based on the pilot study feedback by the seven questions that attached to the original questionnaire of the study (Bell and Waters, 2014), certain words were reconstructed to provide a better understanding to respondents in the main survey. Also, the pilot test revealed that on average, respondents took about 15 to 20 minutes to complete the survey instrument. In addition, Table A2 (Appendix) shows that the result of reliability ranges from 0.794 to 0.906 suggested that all the Cronbach's alpha values were greater than 0.70, which indicate that the 56 measurements were reliable (Hair et al., 2014). Besides, based on the item analysis, all the items correlate higher than 0.30 for the corrected item-total correlation, which ranged from 0.344 to 0.946. These indicate that all items are correlated with their constructs. Therefore, all items have been retained without the need to delete any of them.

4 CONCLUSION

In conclusion, providing a valid and reliable instrument is necessary to ensure accurate results when measuring SMPs and SP. In this respect, two tests were carried out in the present study: the pre-test to establish the validity of the measurements, and the pilot test to check the reliability of the measurements. In the pre-test, the comments made by academician experts and practitioners were used to rephrase items and modify them according to the requirements of the O&GI by the Iraqi environment. Moreover, in the pilot test, some significant findings were identified: an average, respondents took about 15 to 20 minutes to complete the questionnaire, the response rate was 100% high, all the items were reliable and were sufficiently correlated with their constructs. Therefore, this study provides valid and reliable measurements that can give a better understanding to researchers, top management and managers in the O&GI on how to measure SMPs and SP. Theoretically, there are four contributions to the body of knowledge: first, introduce measures for SMPs according to the product life cycle view, it is limited in the literature. Second, these measures can be used by researchers to study the extent of SMPs and the SP of companies using descriptive statistics. Third, these measures can be used to investigate the impact of SMPs on SP by regression testing or structural equation modelling. Fourth, measures can be modified into open-ended questions for use in qualitative or mixed studies. Practically, there are two practical implications. First, managers can measure their companies' implementation of SMPs using a 6-point Likert

scale. Second, measure the level of SP achieved and compare with the performance in previous years.

REFERENCES

Abdul-Rashid, S.H., Sakundarini, N., Ghazilla, R.A.R. and Ramayah, T., 2017a. Drivers for the adoption of sustainable manufacturing practices: A Malaysia perspective. *Int. J. Precis. Eng. Manuf.-Green Tech.,* [e-journal] 18(11), pp.1619-1631. DOI: 10.1007/s12541-017-0191-4.

Abdul-Rashid, S.H., Sakundarini, N., Ghazilla, R.A.R. and Ramayah, T., 2017b. 'The impact of sustainable manufacturing practices on sustainability performance: Empirical evidence from Malaysia. *International Journal of Operations & Production Management,* [e-journal] 37(2), pp.182-204. DOI: 10.1108/IJOPM-04-2015-0223.

Adams, J., Khan, H.T.A. and Raeside, R., 2014. *Research Methods for Business and Social Science Students*. 2nd ed. London: SAGE Publications.

Al-Haleem, A.A., Awadh, S.M. and Saeed, E.A.-J., 2013. Environmental Impact from Drilling and Production of oil Activities: Sources and Recommended Solutions. Ir*aqi Journal of Science*, pp.46-53.

Alwin, D.F. and Krosnick, J.A., 1991. The Reliability of Survey Attitude Measurement: The Influence of Question and Respondent Attributes. *Sociological Methods & Research,* [e-journal] 20(1), pp.139-181. DOI: 10.1177/0049124191020001005.

Andrews, F.M., 1984. Construct Validity and Error Components of Survey Measures: A Structural Modeling Approach. *The Public Opinion Quarterly,* [e-journal] 48(2), pp.409-442. DOI: 10.1086/268840.

Bansal, P., 2005. Evolving sustainably: a longitudinal study of corporate sustainable development. *Strategic Management Journal,* [e-journal] 26(3), pp.197-218. DOI: 10.1002/smj.441.

Bell, J. and Waters, S., 2014. *Doing Your Research Project: A Guide for First-Time Researchers*. 6th ed. Berkshire, UK: McGraw-Hill Education.

Birkett, N.J., 1986. Selecting the number of response categories for a Likert-type scale. In: N. Rothwell*, XII. Cognitive Aspects Of Survey Methodology*. Washington, D.C. pp.488-492.

Bradburn, N.M., Sudman, S. and Wansink, B., 2004. *Asking Questions: The Definitive Guide to Questionnaire Design -- For Market Research, Political Polls, and Social and Health Questionnaires*. 2nd ed. Hoboken, NJ: John Wiley & Sons.

Bryman, A. and Bell, E., 2015. *Business Research Methods*. 4th ed. Oxford: Oxford University Press.

Coelho, P.S. and Esteves, S.P., 2007. The Choice between a Five-point and a Ten-point Scale in the Framework of Customer Satisfaction Measurement. *International Journal of Market Research,* [e-journal] 49(3), pp.313-339. DOI: 10.1177/147078530704900305.

Colton, D. and Covert, R.W., 2007. *Designing and constructing instruments for social research and evaluation.* San Francisco: John Wiley & Sons.

Cooper, D. and Schindler, P., 2014. *Business Research Methods.* 12th ed. USA: McGraw-Hill Higher Education.

Creswell, J.W. and Creswell, J.D., 2018. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches.* 5th ed. USA: SAGE Publications.

Dawson, C., 2009. *Introduction to Research Methods: A practical guide for anyone undertaking a research project.* 4th ed. UK: Little, Brown Book Group.

De Vaus, D.A., 2002. *Surveys in social research.* 5th ed. New South Wales: Allen & Unwin.

DeVellis, R.F., 2016. *Scale development: Theory and applications.* 4th ed. USA: SAGE Publications.

Elhuni, R.M. and Ahmad, M.M., 2017. Key Performance Indicators for Sustainable Production Evaluation in Oil and Gas Sector. *Procedia Manufacturing,* [e-journal] 11 (Supplement C), pp.718-724. DOI: 10.1016/j.promfg.2017.07.172.

EPA (Environmental Protection Agency), 2003. Environmental Impact of the Petroleum Industry. *Environmental Update #12*, [onlien] June 2003. Available at: <https://cfpub.epa.gov/ncer_abstracts/index.cfm/fuseaction/display.files/fileID/14522> [Accessed 1 July 2020].

Field, A., 2009. *Discovering Statistics Using IBM SPSS Statistics.* 3rd ed. USA: SAGE Publications.

Field, A., 2013. *Discovering Statistics Using IBM SPSS Statistics.* 4th ed. USA: SAGE Publications.

Fink, A., 2013. *How to Conduct Surveys: A Step-by-step Guide.* 5th ed. USA: SAGE Publications.

Fink, A., 2017. *How To Conduct Surveys: A Step-by-Step Guide.* 6th ed. USA: SAGE Publications.

Frank, M.D., Nwuche, A.C. and Anyanwu, S.A.C., 2016. Operations management activities and organizational sustainability in oil and gas companies in rivers state. *International Journal of Advanced Academic Research,* 2(11), pp.34-56.

Garland, R., 1991. The mid-point on a rating scale: Is it desirable?. *Marketing Bulletin,* 2, pp.66-70.

Hair, J.F., Black, W.C., Babin, B.J. and Anderson, R.E., 2014. *Multivariate Data Analysis.* 7th ed. USA: Pearson Education Limited.

Hair, J.F., Celsi, M., Money, A.H., Samouel, P. and Page, M.J., 2017. *Essentials of business research methods*. USA: Routledge.

Hair, J.F., Wolfinbarger, M., Bush, R.P. and Ortinau, D.J., 2013. *Essentials of Marketing Research.* 3rd ed. USA: McGraw-Hill Higher Education.

Chomeya, R., 2010. Quality of psychology test between Likert scale 5 and 6 points. *Journal of Social Sciences,* [e-journal] 6(3), pp.399-403. DOI: 10.3844/jssp.2010.399.403.

Ibrahim, Y.M., Hami, N. and Othman, S.N., 2019. Integrating sustainable maintenance into sustainable manufacturing practices and its relationship with sustainability performance: A conceptual framework. *International Journal of Energy Economics and Policy,* [e-journal] 9(4), pp.30-39. DOI: 10.32479/ijeep.7709.

Infante, C.E.D.., Mendonça, F.M.d., Purcidonio, P.M. and Valle, R., 2013. Triple bottom line analysis of oil and gas industry with multicriteria decision making. *Journal of Cleaner Production,* [e-journal] 52, pp.289-300. DOI: 10.1016/j.jclepro.2013.02.037.

Khattab, T.H. and Wahid, N.A., 2015. Penang User's Perception of Domestic Water Quality, Health Risk Concern and Willingness to Pay: A Pilot Study. *Jurnal Teknologi,* [e-journal] 74(11), pp.93-107. DOI: 10.11113/jt.v74.4878.

Krosnick, J.A. and Fabrigar, L.R., 1997. Designing Rating Scales for Effective Measurement in Surveys. *Survey Measurement and Process Quality*, pp.141-164.

Kumar, R., 2011. *Research Methodology: A Step-by-Step Guide for Beginners.* 3rd ed. USA: SAGE Publications.

Kumar, R., 2014. *Research Methodology: A Step-by-Step Guide for Beginners.* 4th ed. USA: SAGE Publications.

Nunnally, J.C. and Bernstein, I., 1994. *Psychometric Theory.* 3rd ed. New York: McGraw-Hill.

OPEC (Organization of the Petroleum Exporting Countries), 2018. *Annual Statistical Bulletin*. [pdf] Vienna: OPEC. Available at: <http://www.thegulfintelligence.com/mediafiles/downloadfile/4833753a-f159-46f2-8dc0-f2335344ebe6.pdf> [Accessed 01 July 2020].

Oppenheim, A.N., 2000. *Questionnaire Design, Interviewing and Attitude Measurement.* USA: Bloomsbury Academic.

Pallant, J., 2011. *Spss Survival Manual.* 4th ed. New York: McGraw-Hill.

Paulraj, A., 2011. Understanding the relationships between internal resources and capabilities, sustainable supply management and organizational sustainability. *Journal of Supply Chain Management,* [e-journal] 47(1), pp.19-37. DOI: 10.1111/j.1745-493X.2010.03212.x.

Saunders, M., Lewis, P. and Thornhill, A., 2009. *Research Methods for Business Students.* 5th ed. UK: Prentice-Hall.

Saunders, M., Lewis, P. and Thornhill, A., 2016. *Research Methods for Business Students.* 7th ed. UK: Pearson Education Limited.

Sekaran, U. and Bougie, R., 2010. *Research Methods For Business: A Skill Building Approach.* 5th ed. UK: Wiley & Sons.

Sekaran, U. and Bougie, R., 2016. *Research Methods For Business: A Skill Building Approach.* 7th ed. UK: Wiley & Sons.

Schneider, J., Vargo, C., Campbell, D. and Hall, R., 2011. An Analysis of Reported Sustainability-Related Efforts in the Petroleum Refining Industry. *The Journal of Corporate Citizenship,* 44, pp.69-84. DOI: 10.9774/GLEAF.4700.2011.wi.00006.

Si, S.X. and Cullen, J.B., 1998. Response Categories And Potential Cultural Bias: Effects Of An Explicit Middle Point In Cross-Cultural Surveys. *The International Journal of Organizational Analysis,* [e-journal] 6(3), pp.218-230. DOI: 10.1108/eb028885.

Thrulogachantar, P. and Zailani, S., 2011. The influence of purchasing strategies on manufacturing performance: An empirical study in Malaysia. *Journal of Manufacturing Technology Management,* [e-journal] 22(5), pp.641-663. DOI: 10.1108/17410381111134482.

UN-ESCWA (United Nations Economic and Social Commission for Western Asia), 2018. *External Trade Bulletin Of The Arab Region.* [pdf] New York: United Nations. Available at: <https://www.unescwa.org/recurring-publication-identifier/external-trade-bulletin-arab-region> [Accessed 01 July 2020].

Van Belle, G., 2008. *Statistical Rules of Thumb.* 2nd ed. UK: Wiley & Sons.

Zhu, Q. and Sarkis, J., 2004. Relationships between operational practices and performance among early adopters of green supply chain management practices in Chinese manufacturing enterprises. *Journal of Operations Management,* [e-journal] 22(3), pp.265-289. DOI: 10.1016/j.jom.2004.01.005.

Zikmund, W.G., Babin, B.J., Carr, J.C. and Griffin, M., 2013. *Business Research Methods.* 9th ed. USA: Cengage Learning.

ABOUT AUTHORS

Yousif Munadhil Ibrahim – (Y.M.I.) Universiti Utara Malaysia, Malaysia, College of Business, School Of Technology Management & Logistic, Dr., e-mail: yousifmonadhil58@gmail.com, Author's ORCID: 0000-0002-7314-2638.

Norsiah Hami – (N.H.) Universiti Utara Malaysia, Malaysia, College of Business, School Of Technology Management & Logistic, Senior Lecturer, e-mail: norsiahami@uum.edu.my, Author's ORCID: 0000-0002-9673-8012.

Susan Sabah Abdulameer – (S.S.A.) Universiti Utara Malaysia, Malaysia, College of Business, School Of Technology Management & Logistic, Dr., e-mail: susansabah50@gmail.com, Author's ORCID: 0000-0003-4441-9930.

AUTHOR CONTRIBUTIONS

Y.M.I. – writing—original draft preparation; N.H. – supervision; N.H., S.S.A. – review and editing.

CONFLICTS OF INTEREST

The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

APPENDIX

Table A1 – Measurement Items of SMPs and SP and Their References

Code	Item	Reference
SPD		
SPD.1	Eliminating the use of hazardous materials during the design of the products.	Abdul-Rashid et al. (2017a), Abdul-Rashid et al. (2017b)
SPD.2	Design the products which will facilitate disassembly of retired products, separation of parts according to materials, as well as reprocessing of materials.	
SPD.3	Design the products which will facilitate repair, rework and refurbishment.	
SPD.4	Design the products which will reduce material use.	
SPD.5	Design the products which will reduce energy consumption.	
SPD.6	Use environmental-friendly materials (e.g. recyclable materials).	
SPD.7	Design the products which support maintenance.	
SPD.8	Design the products which will prolong the life of materials.	
SMP		
SMP.1	Savings of energy during the manufacturing process.	Abdul-Rashid et al. (2017a), Abdul-Rashid et al. (2017b)
SMP.2	Emissions reduction during the manufacturing process.	
SMP.3	Improve manufacturing and machines efficiency.	
SMP.4	Utilise lean production process.	
SMP.5	Commitments to sustainable programmes, standards or regulations.	
SMP.6	Setting sustainable targets and objectives.	
SMP.7	Measure and inspection of material flows or wastes.	
SSCM		
SSCM.1	Adoption of sustainable suppliers.	Abdul-Rashid et al. (2017a), Abdul-Rashid et al. (2017b)
SSCM.2	Influence suppliers to practice sustainable initiatives.	
SSCM.3	Sustainable collaboration with suppliers.	
SSCM.4	Impact customers to accept sustainable practices, services or products.	
SSCM.5	Using a less, cleaner or reusable packaging.	
SSCM.6	Use energy-efficient transportation.	
SSCM.7	Use energy-efficient logistics (e.g. warehouse location and routes).	
SEoLM		
SEoLM.1	Prolong the service life of products or materials by providing support services to customers.	Abdul-Rashid et al. (2017a), Abdul-Rashid et al. (2017b)
SEoLM.2	Providing hazardous waste treatment in the company for products after recovery from the market.	
SEoLM.3	Providing and managing product warranty returns.	

Code	Item	Reference
SEoLM.4	Providing and managing recalls (e.g. reconditioning, reselling).	
SEoLM.5	Providing recycling support using components and material coding standards.	
Economic Sustainability		
EcS.1	Increased net profits.	Bansal (2005), Paulraj (2011), Elhuni and Ahmad (2017), Zhu and Sarkis (2004)
EcS.2	Increased revenue growth.	
EcS.3	Increased revenue through the sale of waste products.	
EcS.4	Increased return on assets.	
EcS.5	Increased return on investment.	
EcS.6	Decreased costs.	
EcS.7	Commitment to production plan %.	
EcS.8	Improving delivery performance.	
Environmental Sustainability		
EnS.1	Reduced emissions of greenhouse gas.	Frank, Nwuche and Anyanwu (2016), Paulraj (2011), Zhu and Sarkis (2004), Elhuni and Ahmad (2017)
EnS.2	Reduced flaring gas.	
EnS.3	Reduced solid wastes.	
EnS.4	Reduced liquid wastes.	
EnS.5	Reduced water usage.	
EnS.6	Reduced oil spills.	
EnS.7	Reduced energy consumption.	
EnS.8	Reduced consumption of hazardous/harmful/toxic materials.	
EnS.9	Reduced environmental accidents.	
Social Sustainability		
SoS.1	Increased local procurement and supplier development.	Elhuni and Ahmad (2017), Frank, Nwuche and Anyanwu (2016), Bansal (2005), Infante et al. (2013)
SoS.2	Increased preventing corruption.	
SoS.3	Increased workforce diversity.	
SoS.4	Increased workforce engagement.	
SoS.5	Increased workforce training and development.	
SoS.6	Decreased rates of work-related injury frequency.	
SoS.7	Decreased rates of work-related occupational illnesses.	
SoS.8	Decreased rates of work-related deaths.	
SoS.9	Participation in community affairs.	
SoS.10	Provide societal health facilities.	
SoS.11	Improved health and safety community.	
SoS.12	Increased social investment.	

Table A2 – Results of the Reliability Test in the Pilot Test

Dimension	Item	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted	Cronbach's Alpha
SMPs				
SPD	SPD.1	0.796	0.884	0.906
	SPD.2	0.596	0.903	
	SPD.3	0.818	0.882	
	SPD.4	0.532	0.907	
	SPD.5	0.727	0.893	
	SPD.6	0.818	0.882	
	SPD.7	0.652	0.898	
	SPD.8	0.670	0.896	
SMP	SMP.1	0.613	0.877	0.886
	SMP.2	0.406	0.903	
	SMP.3	0.661	0.872	
	SMP.4	0.744	0.861	
	SMP.5	0.848	0.848	
	SMP.6	0.744	0.861	
	SMP.7	0.753	0.860	
SSCM	SSCM.1	0.344	0.800	0.794
	SSCM.2	0.726	0.728	
	SSCM.3	0.591	0.755	
	SSCM.4	0.434	0.784	
	SSCM.5	0.592	0.754	
	SSCM.6	0.592	0.754	
	SSCM.7	0.403	0.790	
SEoLM	SEoLM.1	0.836	0.853	0.892
	SEoLM.2	0.927	0.823	
	SEoLM.3	0.775	0.861	
	SEoLM.4	0.685	0.882	
	SEoLM.5	0.531	0.908	
SP				
Economic Sustainability	EcS.1	0.819	0.880	0.904
	EcS.2	0.665	0.895	
	EcS.3	0.730	0.888	
	EcS.4	0.741	0.889	

Dimension	Item	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted	Cronbach's Alpha
	EcS.5	0.946	0.868	
	EcS.6	0.574	0.902	
	EcS.7	0.639	0.897	
	EcS.8	0.488	0.909	
Environmental Sustainability	EnS.1	0.638	0.846	0.864
	EnS.2	0.409	0.868	
	EnS.3	0.840	0.825	
	EnS.4	0.364	0.869	
	EnS.5	0.476	0.861	
	EnS.6	0.545	0.855	
	EnS.7	0.763	0.833	
	EnS.8	0.592	0.850	
	EnS.9	0.753	0.836	
Social Sustainability	SoS.1	0.847	0.879	0.899
	SoS.2	0.517	0.896	
	SoS.3	0.579	0.893	
	SoS.4	0.528	0.895	
	SoS.5	0.596	0.892	
	SoS.6	0.800	0.881	
	SoS.7	0.402	0.902	
	SoS.8	0.860	0.878	
	SoS.9	0.600	0.892	
	SoS.10	0.501	0.897	
	SoS.11	0.477	0.898	
	SoS.12	0.774	0.885	

Anomaly Detection for Noisy Data with the Mahalanobis–Taguchi System

DOI: 10.12776/QIP.V24I2.1441

Masato Ohkubo, Yasushi Nagata

Received: 2020-04-02 Accepted: 2020-04-28 Published: 2020-07-31

ABSTRACT

Purpose: Condition-based maintenance requires an accurate detection of unknown yet-to-have-occurred anomalies and the establishment of anomaly detection procedure for sensor data is urgently needed. Sensor data are noisy, and a conventional analysis cannot always be conducted appropriately. An anomaly detection procedure for noisy data was therefore developed.

Methodology/Approach: In a conventional Mahalanobis–Taguchi method, appropriate anomaly detection is difficult with noisy data. Herein, the following is applied: 1) estimation of a statistical model considering noise, 2) its application to anomaly detection, and 3) development of a corresponding analysis framework.

Findings: Engineers can conduct anomaly detection through the measurement and accumulation, analysis, and feedback of data. Especially, the two-step estimation of the statistical model in the analysis stage helps because it bridges technical knowledge and advanced anomaly detection.

Research Limitation/implication: A novel data-utilisation design regarding the acquired quality is provided. Sensor-collected big data are generally noisy. By contrast, data targeted through conventional statistical quality control are small but the noise is controlled. Thus various findings for quality acquisition can be obtained. A framework for data analysis using big and small data is provided.

Originality/Value of paper: The proposed statistical anomaly detection procedure for noisy data will improve of the feasibility of new services such as condition-based maintenance of equipment using sensor data.

Category: Research paper

Keywords: engineered system; Gaussian graphical model; liner Gaussian model; statistical modelling; Taguchi method

1 INTRODUCTION

In recent years, the condition-based maintenance of equipment using sensor data has been put into practical use in the Japanese manufacturing industry based on the progress of technologies related to the Internet of Things (IoT). Condition-based maintenance requires an accurate detection of unknown anomalies that have yet to occur, and thus the establishment of an anomaly detection procedure for sensor data is needed. However, because sensor data are noisy with features observed by adding noise to the true measurement value, a conventional analysis procedure cannot always be used to conduct an appropriate analysis. The purpose of this study is to develop a novel anomaly detection procedure for noisy data.

Figure 1 shows a schematic diagram illustrating the process of generating noisy data. The symbols $\boldsymbol{x}$, $\boldsymbol{y}$, and $\boldsymbol{e}$ in the figure are P-dimensional random variables indicating the true measurement value, observed value, and noise, respectively. For example, when assuming the formation process of a porous film, the true value of measurement $\boldsymbol{x}$ takes a value measured using precision-measuring equipment as the physical properties of the porous film (e.g., film thickness, film width, and pore diameter). By contrast, the observed value $\boldsymbol{y}$ takes a value measured using sensors for such physical properties. Here, noise is caused by measurement errors derived from the sensors or by various other factors. Thus, such noise is additively superimposed over the true value of the measurement, and $\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{e}$ is established. As described in Section 3, this model is also related to the engineered system used in the Taguchi method.

When dealing with such noisy data, problems can arise in which it becomes difficult for engineers to consider adopting anomaly detection algorithms from the viewpoint of intrinsic technology. Taking the example of the formation process of a porous film, it can be stated that it is easy for skilled engineers to consider the relation between the physical properties of the porous film from the true measurement value. By contrast, in the case of noisy data, correlations that should be found among the physical properties of a porous film cannot be inferred from the data. Considering the influence of noise, anomaly detection based on a model in which the essential correlation structure in the true measurement value is ignored can be useful for improving the anomaly detection. Therefore, there is a risk that the adopted anomaly detection algorithms will lack validity from the viewpoint of intrinsic technology.

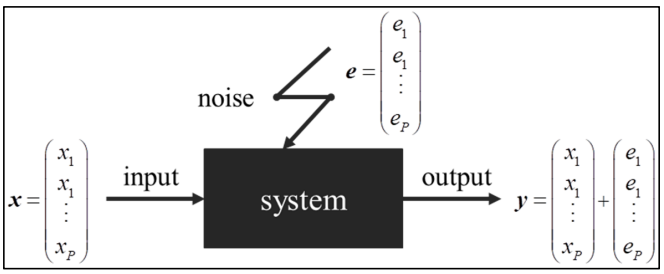

Figure 1 – Process of Generating Noisy Data

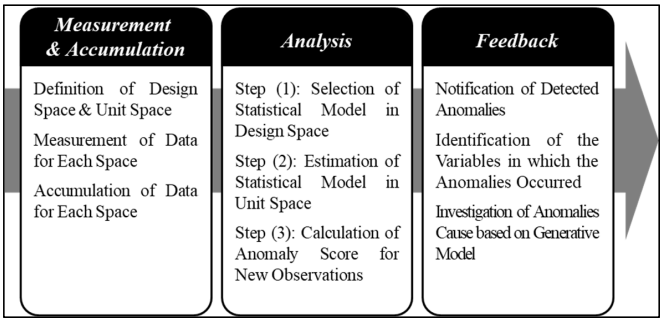

Figure 2 – Overview of Proposed Procedure

To solve this problem, a novel anomaly detection procedure is proposed that improves the Mahalanobis–Taguchi (MT) method developed by Taguchi and Jugulum (2002). Figure 2 shows an overview of the proposed procedure. From Figure 2, the proposed procedure consists of three stages, namely the measurement and accumulation, analysis, and feedback of data. Through these stages, an anomaly detection procedure that allows engineers to consider algorithms from the viewpoint of intrinsic technology while maintaining a predictive performance is expected to be realised.

In the first stage, the measurement and accumulation of data are conducted for two spaces, the design space and the unit space. The design space is a novel concept not found in a conventional MT method. The concept of a design space is proposed with the intention of incorporating the technical knowledge of the engineer into the analysis procedure. The details of this are described in Section 3.2.1.

In the following analysis stage, anomaly detection is conducted using three steps: (1) the selection of a statistical model in the design space, (2) an estimation of a statistical model in the unit space, and (3) the calculation of an anomaly score for new observations. Step (1) is a novel stage. It has been recognised that statistical models that are consistent with the engineer's technical knowledge often differ from models showing a high predictive performance. The proposed procedure is expected to fill in the gap between these two models. The details of this are provided in Section 3.2.2.

In the final feedback stage, anomalies detected in the analysis stage are notified to the engineers. This helps investigate the causes of anomalies by identifying the variable that is suspected of incurring an anomaly. It is also expected to make it possible to investigate the causes of anomalies more accurately by analysing the statistical model estimated during the analysis stage according to the proposal by Ohkubo and Nagata (2017). The details of this are described in Section 3.2.3.

The structure of this paper is as follows. In Section 2, the MT method as used in previous studies is applied and its anomaly detection procedure is described. In Section 3, a novel anomaly detection procedure for noisy data is proposed. In Section 4, the usefulness of the proposed procedure is confirmed through a

Monte Carlo simulation. Finally, in Section 5, the implications of the proposal are discussed.

2 MAHALANOBIS–TAGUCHI SYSTEM

The MT system is a generic term for multivariate analysis methods proposed using the Taguchi method. There are many applications in various fields centring on the manufacturing industry (e.g., Jin and Chow, 2013; Peng et al., 2017). Among the different MT systems, in recent years, the MT method, which is a multivariate analysis procedure for anomaly detection, has attracted the attention of practitioners. In fact, the MT method has been used as an algorithm forming the core of an anomaly prediction system of equipment using sensor data (Takahama and Mikami, 2012). In this section, an overview of the MT method is described in terms of the measurement and accumulation, analysis, and feedback to facilitate a comparison with the proposed procedure.

2.1 Measurement and Accumulation Stage

In the measurement and accumulation stage, after defining a dataset called a unit space, the measurement and accumulation of data are conducted. A unit space is a group forming a homogeneous population. In general, in the field of statistical anomaly detection, after sampling normal data, namely data with a normal label, the data are used as training data to estimate a statistical model. The unit space applied in the MT method can be considered training data consisting only of data with normal labels.

Here, when conducting anomaly detection, it is necessary to prepare test data to evaluate whether the performance of the algorithm is sufficient for practical use. Test data consist of data with normal labels and data with anomaly labels. In the proposed procedure, although test data are also required, the description is omitted because it becomes redundant. In addition, it is necessary to prepare a dataset called “signal data” separately from the unit space in the MT method. However, it has been pointed out that the definition and use of signal data remain controversial (e.g., Woodall et al., 2003; Inoh et al., 2012). In this paper, signal data are interpreted as test data, and signal data or the analysis procedure related to it are not discussed.

2.2 Analysis Stage

In the analysis stage, anomaly detection is conducted using two steps: (1) an estimation of the Mahalanobis distance in the unit space and (2) the calculation of an anomaly score for new observations. In general, in the field of anomaly detection, the deviation from the normal state is quantified by a certain scale, and an anomaly is considered if the scale exceeds a predetermined threshold, and a normal state is determined if it does not. The MT method can be said to be an

anomaly detection procedure that uses the Mahalanobis distance (MD) as its scale.

First, the MD is estimated from an observed value in the unit space. Let $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ be the population mean vector and population covariance matrix of the *P*-dimensional variable $\boldsymbol{y}$ observed for a population. Then, the MD on the population is defined as follows:

$$\boldsymbol{\Delta}(\boldsymbol{y} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \left(\boldsymbol{y} - \boldsymbol{\mu}\right)^T \boldsymbol{\Sigma}^{-1} \left(\boldsymbol{y} - \boldsymbol{\mu}\right), \tag{1}$$

where *T* is the transpose of a vector or matrix. Note that the vectors in this paper are vertical vectors, and their transposes are horizontal vectors. Then, with the MT method, the MD on the population is estimated using the sample mean vector and sample covariance matrix obtained from *N* individuals belonging to the unit space as the estimators of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$.

In (2), the anomaly score for the new observations is calculated. Here, a function for quantifying the degree of deviation from the normal state is called an anomaly score function. In addition, the value of this function taken in practice is called an anomaly score. The anomaly score function in the MT method is the following function in which $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ in equation (1) are replaced by their respective estimator $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$:

$$\boldsymbol{\Delta}(\boldsymbol{y} \mid \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}) = \left(\boldsymbol{y} - \hat{\boldsymbol{\mu}}\right)^T \hat{\boldsymbol{\Sigma}}^{-1} \left(\boldsymbol{y} - \hat{\boldsymbol{\mu}}\right). \tag{2}$$

At this time, after substituting the new observed value $\boldsymbol{y}_{\text{new}}$ into equation (2) and calculating the anomaly score, an anomaly is determined if the value exceeds the predetermined threshold; otherwise, it is judged as normal.

2.3 Feedback Stage

In the feedback stage, anomalies detected during the analysis stage are notified to the engineers. In addition, the variable suspected of having the anomaly is identified through a causal diagnosis using an orthogonal array. Here, a causal diagnosis using an orthogonal array is conducted as follows: (1) assign each variable on a two-level orthogonal array as a factor, taking 1 when using the *j*-th variable of observed variable $\boldsymbol{y}$, and 0 when not using it, (2) calculate the anomaly score through the combination specified in each row, (3) calculate the factorial effects of each variable, and (4) create a graph of the factorial effects. See Taguchi and Jugulum (2002) for more details on this.

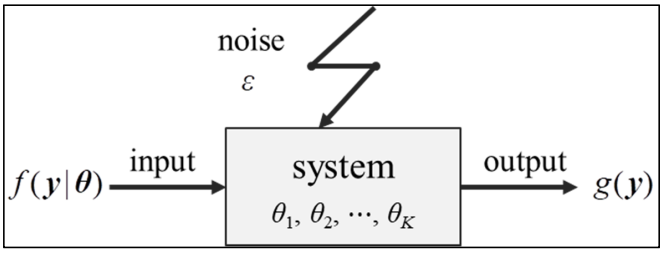

Figure 3 – Modelling Philosophy of the Proposed Procedure

3 PROPOSED PROCEDURE

In this section, the proposed anomaly detection procedure is described, which is an improvement over the MT method under the premise of its application to noisy data.

Figure 3 shows a schematic diagram illustrating the modelling philosophy, following the parameter design of the Taguchi method. The symbols $f(\boldsymbol{y}|\boldsymbol{\theta})$, $g(\boldsymbol{y})$, and ε in the figure indicate parametric statistical models, the true distribution of $\boldsymbol{y}$, and noise, respectively. The symbols θ_1, θ_2, ···, θ_K in the figure are the elements of $\boldsymbol{\theta}$ in $f(\boldsymbol{y}|\boldsymbol{\theta})$, respectively. Figure 3, which shows a system chart of the Taguchi method, illustrates the estimation of $f(\boldsymbol{y}|\boldsymbol{\theta})$ that minimizes the effect of noise ε by appropriately applying the set of parameters θ_1, θ_2, ···, θ_K. In addition, the entire system shown in the system chart is called an engineered system.

The novel anomaly detection procedure is formulated according to this modelling philosophy. The aim is to prove that the proposed procedure will allow engineers to consider the algorithm from the viewpoint of intrinsic technology while maintaining a predictive performance by introducing some concepts from the Taguchi method. The details of the proposed procedure are described for each stage, namely the measurement and accumulation, analysis, and feedback of data, in the following sections. Note that Taguchi, Chowdhury and Wu (2005) is used to indicate the Taguchi method.

3.1 Measurement and Accumulation Stage

In the measurement and accumulation stage, after defining two datasets, called the "design space" and "unit space", a measurement and accumulation of the data are conducted for each space. In this section, the definition of each space and the established relation between the true measurement value $\boldsymbol{x}$, observed value $\boldsymbol{y}$, and noise $\boldsymbol{e}$ in the design and unit spaces, respectively, are described.

First, the design space is defined as a set of observed values in an ideal state. Such data are actively measured and accumulated by engineers mainly using

precision instruments with the purpose of performing statistical modelling from the viewpoint of intrinsic technology. At this time, it is assumed that $\boldsymbol{y} = \boldsymbol{x}$ is established between the true value of measurement $\boldsymbol{x}$ and the observed value $\boldsymbol{y}$.

Next, the unit space is defined as a set of observed values in a usual state. Such data are measured and accumulated passively from sensors with the purpose of applying statistical modelling from the viewpoint of a predictive performance. At this time, it is assumed that the observed value $\boldsymbol{y}$ is generated along with the noise $\boldsymbol{e}$ according to the generation process shown in Figure 1. That is, $\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{e}$ is established.

Here, the concept of a design space is additionally described. Although both data types in the design and unit spaces are P-dimensional observation variables with a normal label, the data collection method and generative model differ, as described above. Therefore, merging these data and treating them as single learning data is not recommended. The intention of using the design space is considered the same as using a test piece with the Taguchi method. In the Taguchi method, it is considered useful to create a model, called a test piece, to capture the essence of a technology. The design space is also measured and accumulated to grasp the essential structure of the data.

Note that the notations of the probability density function in each space are as follows. First, the true distribution of $\boldsymbol{y}$, i.e. $g(\boldsymbol{y})$, is described as $g_0(\boldsymbol{y})$ when it emphasises the distribution of the design space. Similarly, when it emphasises the distribution of the unit space, it is described as $g_1(\boldsymbol{y})$. Next, the true distribution of $\boldsymbol{x}$ is described as $s(\boldsymbol{x})$ in both cases, assuming it does not change in the design or unit space. Finally, the parametric statistical model $f(\boldsymbol{y}|\boldsymbol{\theta})$ is described as $f_0(\boldsymbol{y}|\boldsymbol{\theta}_0)$ and $f_1(\boldsymbol{y}|\boldsymbol{\theta}_1)$ in the design and unit spaces, respectively, as in the case of $g(\boldsymbol{y})$.

3.2 Analysis Stage

In the analysis stage, anomaly detection is applied using three steps: (1) selection of the statistical model in the design space, (2) estimation of the statistical model in the unit space, and (3) the calculation of an anomaly score for new observations.

In (1), statistical modelling is conducted on the set of observed values in the design space under the assumption of a Gaussian graphical model. At this time, the best model is selected from a family of models while actively using the technical knowledge. In (2), statistical modelling is conducted on the set of observed values in the unit space under the assumption of a linear Gaussian model. That is, an estimation of a statistical model considering the existence of noise is conducted using the information of the model selected in (1). In (3), anomaly detection is conducted based on the statistical model estimated in (2) according to the general theory of anomaly detection. The details of this procedure are described below in order of (1) to (3).

3.2.1 Selection of Statistical Model in Design Space

Now, when a random variable $\boldsymbol{z}$ follows a P-dimensional normal distribution with the P-dimensional vector $\boldsymbol{\mu}$ and the $P \times P$ symmetric positive-definite matrix $\boldsymbol{\Sigma}$ as parameters, its probability density function is described as $\mathcal{N}(\boldsymbol{z}|\boldsymbol{\mu},\boldsymbol{\Sigma})$. At this time, assume $g_0(\boldsymbol{y}) = \mathcal{N}(\boldsymbol{y}|\boldsymbol{\mu},\boldsymbol{\Sigma})$ and $s(\boldsymbol{x}) = \mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu},\boldsymbol{\Sigma})$. In this step, Gaussian graphical modelling (GGM) described by Lauritzen (1996) is conducted on the set of observations valued in the design space under this assumption.

GGM is one of the basic methods used in a multivariate analysis, which is conducted by applying the following steps: (1) estimating the parameters in a multivariate normal distribution and (2) evaluating and selecting statistical models, assuming that most of the non-diagonal elements of $\boldsymbol{\Omega}$ take a value of zero where the inverse matrix of the covariance matrix $\boldsymbol{\Sigma}$ is $\boldsymbol{\Omega}$. There are many applications of GGM; applied procedures have also been proposed in the fields of anomaly detection (e.g., Ide et al., 2009). Ohkubo and Nagata (2017) proposed introducing the MT method, and indicated that the method is useful for both improving the anomaly detection performance and pursuing the cause of the anomaly.

In this study, GGM is conducted to understand the generative model, as applied by Ohkubo and Nagata (2017), for the usefulness of the feedback step described later. Therefore, when using the information criteria, it is recommended to use criteria that have the same concept as the Bayesian information criteria (BIC) developed by Schwarz (1978). However, a statistical model should be selected from the viewpoint of intrinsic technology while referring to such model of the evaluation criteria. As mentioned later, the model selection in this step does not significantly affect the anomaly detection performance, and it can therefore be useful to obtain a model to understand the essential correlational structure of the data.

3.2.2 Estimation of Statistical Model in Unit Space

In this step, estimating a parametric statistical model $f_1(\boldsymbol{y}|\boldsymbol{\theta}_1)$ against a set of observed values in the unit space is considered. Here, a linear Gaussian model is assumed for the true distribution $g_1(\boldsymbol{y})$ in the unit space. Now, assume that the probability density function of noise $\boldsymbol{e}$ is $\mathcal{N}(\boldsymbol{e}|\boldsymbol{0}_P,\boldsymbol{\Lambda}^{-1})$, where $\boldsymbol{0}_P$ is a P-dimensional zero vector and $\boldsymbol{\Lambda}$ is a $P \times P$ symmetric positive-definite matrix. The observed value $\boldsymbol{y}$ then follows a multivariate normal distribution from the assumption in the signal space, $\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{e}$. In addition, from the reproducibility of the normal distribution, the marginal distribution of $\boldsymbol{y}$ is described as follows:

$$g_1(\boldsymbol{y}) = \mathcal{N}(\boldsymbol{y}|\boldsymbol{\mu},\boldsymbol{\Omega}^{-1} + \boldsymbol{\Lambda}^{-1}), \tag{3}$$

where $s(\boldsymbol{x}) = \mathcal{N}(\boldsymbol{x} \mid \boldsymbol{\mu}, \boldsymbol{\Omega}^{-1})$ from the assumption of the design space, $\boldsymbol{y} = \boldsymbol{x}$. At this time, an estimation of the parameters can be conducted under the model $f_1(\boldsymbol{y} \mid \boldsymbol{\theta}_1)$, which is assumed to be a linear Gaussian model where $\boldsymbol{\theta}_1 = \{\boldsymbol{\mu}, \boldsymbol{\Omega}, \boldsymbol{\Lambda}\}$. Note that the expectation–maximisation (EM) algorithm can be easily implemented for this model (e.g., Bishop, 2006).

3.2.3 Calculation of Anomaly Score for New Observations

In this step, an anomaly score function is defined based on the general theory of anomaly detection (e.g., Yamanishi and Takeuchi, 2002). Specifically, $f_1(\boldsymbol{y} \mid \hat{\boldsymbol{\theta}}_1)$ is defined, which is a statistical model in which the parameters of $f_1(\boldsymbol{y} \mid \boldsymbol{\theta}_1)$ have been replaced by the estimators $\hat{\boldsymbol{\theta}}_1$, as the anomaly score function:

$$Score\left(\boldsymbol{y} \middle| \hat{f}_1\right) = -\ln f_1(\boldsymbol{y} \mid \hat{\boldsymbol{\theta}}_1). \tag{4}$$

Here, the plug-in estimator is applied such that the estimator of the parameter is plugged into the original probability density function as an estimator of the parametric statistical model, and $\hat{f}_1(\boldsymbol{y} \mid \boldsymbol{\theta}_1) = f_1(\boldsymbol{y} \mid \hat{\boldsymbol{\theta}}_1)$ is established. At this time, after substituting the new observed value $\boldsymbol{y}_{\text{new}}$, which is the target of judgement, into equation (4) and calculating the anomaly score as in the case of the MT method, an anomaly is determined if the value exceeds the predetermined threshold; otherwise, the value is judged as normal.

3.2.4 Relationship with Modelling Philosophy

This sub-subsection additionally describes the relation between the analysis stage in the proposed procedure and the modelling philosophy.

According to the modelling philosophy shown in Figure 3, the purpose of the proposed procedure is to estimate a $f_1(\boldsymbol{y} \mid \boldsymbol{\theta}_1)$ that minimizes the effect of noise ε by setting $\boldsymbol{\theta}_1$ appropriately. At this time, it is assumed that the following relation holds among $f_1(\boldsymbol{y} \mid \boldsymbol{\theta}_1)$, $g_1(\boldsymbol{y})$, and ε:

$$Score\left(\boldsymbol{y} \middle| g_1\right) = Score\left(\boldsymbol{y} \middle| f_1\right) + \varepsilon, \tag{5}$$

where the *Score* function has the same definition as in equation (4). Equation (5) can be interpreted as the state in which the *Score* functions of $g_1(\boldsymbol{y})$ and $f_1(\boldsymbol{y} \mid \boldsymbol{\theta}_1)$ match, which can be interpreted as the ideal state. In addition, it can be stated that minimising noise ε while aiming at an ideal state is a means to realise the above-mentioned modelling philosophy. Note that, in the case of $\varepsilon = 0$, equation (5) is called an ideal function in the Taguchi method.

When aiming at such an ideal function, it is possible to consider the expected noise value ε against $g_1(\boldsymbol{y})$ as in the following evaluation index:

$$E[\varepsilon] = \int \left\{ Score\left(\boldsymbol{y} \middle| g_1 \right) - Score\left(\boldsymbol{y} \middle| f_1 \right) \right\} g_1(\boldsymbol{y}) d\boldsymbol{y}$$
$$= -\int \left\{ \ln \frac{g_1(\boldsymbol{y})}{f_1(\boldsymbol{y} \mid \boldsymbol{\theta}_1)} \right\} g_1(\boldsymbol{y}) d\boldsymbol{y} \tag{6}$$

Here, the right-hand side of the second equal sign is the negative Kullback–Leibler (KL) divergence from the equation form. Because a KL divergence is non-negative (e.g., Bishop, 2006), the minimisation problem in equation (6) can be said to be the KL divergence minimisation problem. Note that this evaluation index can be stated to be the same concept as the signal-to-noise (SN) ratio in the Taguchi method.

From the above discussion, the rational modelling procedure in (2) is a modelling procedure that aims to estimate $f_1(\boldsymbol{y} \mid \boldsymbol{\theta}_1)$, which accurately approximates $g_1(\boldsymbol{y})$ in equation (3). That is, the estimated covariance structure is an estimated value of $\boldsymbol{\Omega}^{-1} + \boldsymbol{\Lambda}^{-1}$, where the effect of noise is added to the essential correlation structure. Therefore, it is difficult to grasp the essential correlation structure only through (2). In the same way, the same problem occurs in the conventional MT method. By contrast, with the proposed procedure, it is possible to realise the modelling philosophy by executing (2) after (1) while grasping the essential correlation structure.

3.3 Feedback Stage

In the feedback stage, anomalies detected during the analysis stage are notified to the engineers. In addition, as with a conventional MT method, identifying the variable that is suspected of having an anomaly through a causal diagnosis using an orthogonal array helps in investigating the cause of the anomaly. At the same time, information on the covariance structure among the parameters of the statistical model previously estimated during the analysis stage is simultaneously provided.

Ohkubo and Nagata (2017) proposed applying the GGM framework to investigate the cause of an anomaly when it occurs. It is possible to consider the cause of an anomaly through the conditional independence between variables by observing the precision matrix, which is the analysis result of the GGM. However, in the case of noisy data, because the covariance structure is affected by noise in the unit space, it is difficult to grasp the essential structure of the correlation even if the precision matrix is observed. By contrast, with the proposed procedure, even in the presence of noise, it is possible to obtain an estimate of the precision matrix $\boldsymbol{\Omega}$ with high accuracy, and thus it is possible to capture an essential correlation structure. Therefore, the proposed procedure is expected to be useful for investigating the cause of an anomaly more accurately even in the case of noisy data.

4 MONTE CARLO SIMULATION

In this Section, through Monte Carlo simulation, the usefulness of the proposed procedure is verified from two perspectives: predictive performance and consistency with the intrinsic technology.

4.1 Dataset Overview

In this experiment, 100 sets of design space, unit space, and test data are prepared. Here, the dimension of each data is 10 and the sample size is 50, 200, 50,000 in order.

The generative model is $\mathcal{N}(\boldsymbol{y} \mid \boldsymbol{\mu}_0, \boldsymbol{\Omega}_0^{-1})$ for design space and $\mathcal{N}(\boldsymbol{y} \mid \boldsymbol{\mu}_0, \boldsymbol{\Omega}_0^{-1} + \boldsymbol{\Lambda}_0^{-1})$ for unit space, using $\boldsymbol{\mu}_0$, $\boldsymbol{\Omega}_0$, and $\boldsymbol{\Lambda}_0$ described later. That is, according to the generation process shown in Figure 1, it is assumed that the observed value $\boldsymbol{y}$ is generated with noise. The test data is composed of data with normal labels and anomaly labels, and its generative model of normal state is the same as unit space. By contrast, the generative model of anomaly state is $\mathcal{N}(\boldsymbol{y} \mid \boldsymbol{\mu}_0, c^2\boldsymbol{\Omega}_0^{-1} + c^2\boldsymbol{\Lambda}_0^{-1})$ where $c^2 = 5$.

Here, $\boldsymbol{\mu}_0$, $\boldsymbol{\Omega}_0$, and $\boldsymbol{\Lambda}_0$ are set as follows. First, let $\boldsymbol{\mu}_0$ be a 10-dimensional zero vector. Next, $\boldsymbol{\Omega}_0$ is set so that most of its off-diagonal elements are zero while satisfying positive definite. Specifically, we use the following procedure applying Cholesky decomposition: (1) prepare a lower triangular matrix $\mathbf{B}_\Omega$ whose elements are all 1 including diagonal elements, (2) replace 20% of off-diagonal elements in $\mathbf{B}_\Omega$ with a value of zero, (3) standardize so that the diagonal elements of $\mathbf{B}_\Omega\mathbf{B}_\Omega^T$ are 1, and (4) let $\tilde{\mathbf{B}}_\Omega\tilde{\mathbf{B}}_\Omega^T$ be the result of procedure in (3), and then let $\boldsymbol{\Omega}_0$ be $(1-\alpha)^{-1}\tilde{\mathbf{B}}_\Omega\tilde{\mathbf{B}}_\Omega^T$ weighted with α ($0 < \alpha < 1$). Let $\boldsymbol{\Lambda}_0$ be $\alpha^{-1}\tilde{\mathbf{B}}_\Lambda\tilde{\mathbf{B}}_\Lambda^T$ according to the same procedure. Note that we call "uncorrelated noise" when the ratio of non-zero elements to the all non-diagonal elements of $\tilde{\mathbf{B}}_\Lambda$ is 0%, and call "correlated noise" when is 100%. From above, if we define $\boldsymbol{\Sigma}_0$ as $\boldsymbol{\Sigma}_0 = \boldsymbol{\Omega}_0^{-1} + \boldsymbol{\Lambda}_0^{-1}$, the following relationship holds:

$$\boldsymbol{\Sigma}_0 = (1-\alpha)\left(\tilde{\mathbf{B}}_\Omega\tilde{\mathbf{B}}_\Omega^T\right)^{-1} + \alpha\left(\tilde{\mathbf{B}}_\Lambda\tilde{\mathbf{B}}_\Lambda^T\right)^{-1}. \tag{7}$$

From equation (7), it can be seen that α ($0 < \alpha < 1$) is a weight for covariance matrix of noise. As α approaches 0, the effect of noise decreases, and as it approaches 1, the effect of noise increases.

4.2 Evaluation Objects

In this experiment, Proposed MT, Predictive MT, and Interpretive MT are compared. First, Proposed MT means our proposed procedure. Next, Predictive MT means MT method with emphasis on predictive performance. In this case, it is MT method using sample mean vector and sample covariance matrix

calculated from the observation values of unit space. Finally, Interpretive MT means MT method which emphasizes that it can be interpreted from the viewpoint of intrinsic technology. In this case, it is MT method using the sample mean vector and sample covariance matrix estimated from the observed values of unit space, under the known graph structure of $\mathbf{\Omega}_0$ (i.e., the position of the nonzero element of $\mathbf{\Omega}_0$). We note that the MT method using true parameters is called Ideal MT. Then, it is judged that the closer to the performance of Ideal MT, the better the anomaly detection procedure.

4.3 Evaluation criteria

In this experiment, the performance of the evaluation objects shown in Section 4.2 is evaluated from the viewpoint of prediction performance and consistency with the intrinsic technology. Here, the consistency with the intrinsic technology in this experiment is evaluated based on whether the generative model of $\boldsymbol{x}$ in the design space can be accurately estimated. The following describes specific evaluation criteria.

First, the evaluation criterion for measuring the prediction performance is the positive discrimination rate of test anomaly samples (hereinafter referred to simply as the positive discrimination rate) of each procedure. The positive discrimination rate takes a value from 0 to 100 and is better if close to 100. The threshold value is set at 1% of the negative discrimination rate of test normal samples (hereinafter referred to simply as the negative discrimination rate).

Next, the evaluation criterion to measure the consistency with the intrinsic technology is the KL divergence between the distribution $s(\boldsymbol{x})$ of the true measured value $\boldsymbol{x}$ and its predicted distribution $\hat{s}(\boldsymbol{x}) = \mathcal{N}(\boldsymbol{x} \mid \hat{\boldsymbol{\mu}}_0, \hat{\mathbf{\Omega}}_0^{-1})$. Here, Predictive MT and Interpretive MT in the previous section use $\hat{\mathbf{\Sigma}}_0^{-1}$ as $\hat{\mathbf{\Omega}}_0$.

4.4 Experimental Result

The results of this experiment are shown in Figure 4 and Figure 5. Note that the proposed procedure is conducted by a maximum likelihood estimation supposed that the structure of $\mathbf{\Omega}_0$ is given when we estimate parameters from the design space in order to align conditions with Interpretive MT.

Figure 4 shows graphs comparing the evaluation objects from the viewpoint of prediction performance. The left side corresponds to the experimental results for uncorrelated noise and the right side corresponds to the experimental results for correlated noise. The vertical axis of each figure is the positive discrimination rate, and the horizontal axis is the weight α for the covariance matrix of noise.

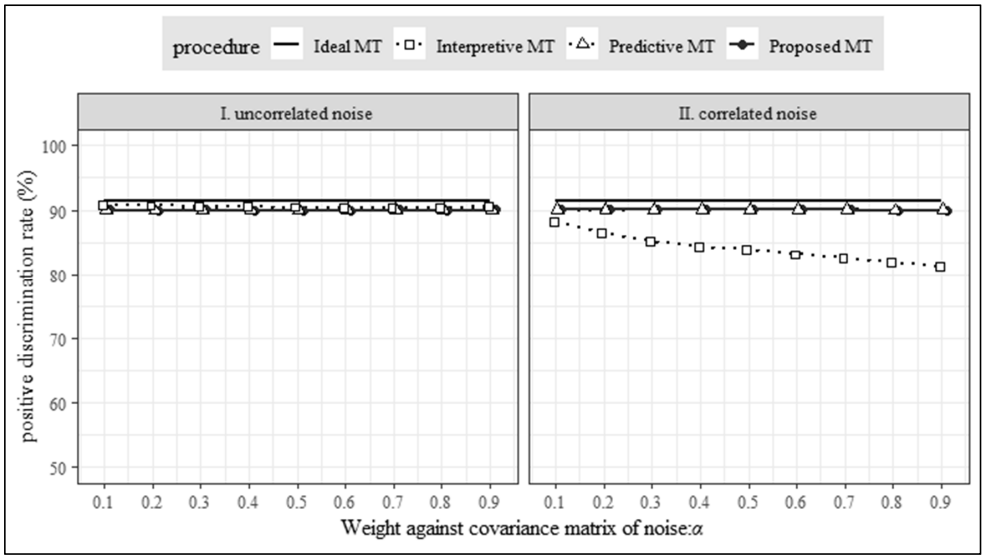

Figure 4 – Evaluation of Predictive Performance In Each Procedure

From Figure 4, in the case of correlated noise, it can be observed that the positive discrimination rate of Interpretive MT decreases as α increases. In the case of correlated noise, this is caused by the difference between the graph structure of $\boldsymbol{x}$ and $\boldsymbol{y}$, occurring due to the correlation between the variables of noise. Because the prior knowledge used by Interpretive MT is a graph structure of $\boldsymbol{x}$, the influence of incorrect model setting becomes stronger as the influence of noise becomes stronger. By contrast, although Proposed MT also uses the graph structure of $\boldsymbol{x}$, appropriate analysis is possible because the model is corrected from data in the unit space.

Figure 5 shows graphs comparing the evaluation objects from the viewpoint of consistency with the intrinsic technology. The left side corresponds to the experimental results for uncorrelated noise and the right side corresponds to the experimental results for correlated noise. The vertical axis of each figure is the KL divergence, and the horizontal axis is the weight α for the covariance matrix of noise.

It can be observed from Figure 4 that in the case of correlated noise, Proposed MT has obtained a predicted distribution with smaller KL divergence than Predictive MT and Interpretive MT. The cause of this phenomenon is that in Predictive MT and Interpretive MT, estimation is conducted on the population distribution of $\boldsymbol{y}$, not $\boldsymbol{x}$. In the case of correlated noise, the correlation structure of $\boldsymbol{y}$ becomes different from the correlation structure of $\boldsymbol{x}$ as the influence of the noise becomes stronger. Therefore, in the case of correlated noise, it can be said that it becomes difficult to estimate the generative model of $\boldsymbol{x}$. By contrast, Proposed MT can learn the generative model of unit space while maintaining the information of the generative model of $\boldsymbol{x}$.

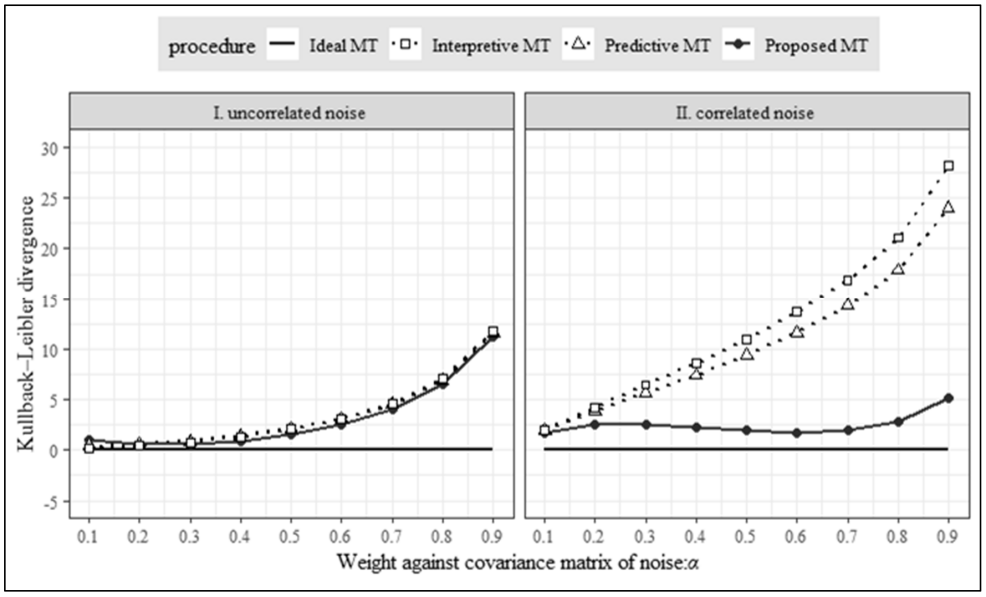

Figure 5 – Evaluation of Consistency with Intrinsic Technology in Each Procedure

5 DISCUSSION

In this research, a novel anomaly detection procedure was proposed for noisy data. Specifically, the anomaly detection procedure used in the MT method, which is a representative methodology based on the Taguchi method, has been improved such that noisy data can be properly analysed. Through each stage, namely the measurement and accumulation, analysis, and feedback of data, quality-management engineers can conduct anomaly detection while reflecting on their technical knowledge in an analysis. Although the proposal covers all three stages of data utilisation, a two-step estimation of the statistical model in the analysis stage is useful in the sense that it fills in the gap between the engineer's technical knowledge and advanced anomaly detection. The usefulness of the proposed procedure was thus shown through a theoretical examination and numerical experiment using a Monte Carlo simulation.

In recent years, the utilisation of IoT-related technologies in the manufacturing industry has become increasingly popular. Because IoT-related technology has also been reviewed as an important factor in establishing a competitive advantage in the manufacturing sector (Porter and Heppelmann, 2014), active discussions should be made on how to use IoT-related technologies in all aspects of corporate activities, including quality management. In fact, the impact of IoT on quality management has been considered from various perspectives, not only in industry but also in the academic field (e.g., Foidl and Felderer, 2015; Park et al., 2017; Shin et al., 2018). Thus, how to use data for quality management has been recognised as a common key factor.

This proposal should lead to a data utilisation design when considering the aspect of quality acquisition. In recent years, although many arguments have been made regarding the use of so-called big data collected from sensors and smart devices, it is difficult to obtain technical knowledge from such data because much of the data are noisy. By contrast, data targeted using a conventional statistical quality control and the Taguchi method are small but their noise is controlled. Thus, from such small data, various findings for a quality acquisition can be determined. This proposal will lead to the provisioning of a novel framework for conducting a data analysis while taking advantage of both big and small data.

The proposed procedure enables the learning of an anomaly detection model to achieve a high performance from the unit space corresponding to big data while obtaining technical knowledge from the design space corresponding to small data. Even in the field of anomaly detection, high-performance machine learning methodologies have been proposed, e.g., the local outlier factor developed by Breunig et al. (2000) and the one-class support vector machine developed by Schölkopf et al. (2001). However, many of these methodologies are black-boxed algorithms. By contrast, although conventional procedures of MT systems have clear algorithms, it is difficult to achieve as high a performance as a machine learning methodology. The proposed procedure is an example of a methodology for achieving a high performance while ensuring validity from the viewpoint of intrinsic technology, whereas the assumptions for statistical models are severe.

Finally, the contributions and limitations of this research are summarised as follows. The main contributions of this study are to establish a novel anomaly detection procedure for noisy data and improve the feasibility of new services, such as condition-based maintenance of the equipment using sensor data. A limitation of this study is its inability to quantitatively express the degree of abnormality for each occurrence factor. In the field of multivariate control charting, originating from the T^2 chart proposed by Hotelling (1947), when an anomaly occurs, the anomaly cause, which is a breakdown of the essential correlation structure or fluctuation from noise, is expressed as individual statistics (e.g., Jackson and Mudholkar, 1979). Similar proposals have also been made in the field of MT systems (Ohkubo and Nagata, 2018). From this perspective, a future study will be applied to improve the proposed procedure. It is also necessary to apply the proposed procedure to real cases such as condition-based equipment maintenance.

ACKNOWLEDGEMENTS

This work was supported by the Japan Society for the JSPS KAKENHI Grant Numbers JP18K11202 and JP18K13953.

REFERENCES

Bishop, C.M., 2006. *Pattern Recognition and Machine Learning*. New York: Springer.

Breunig, M.M., Kriegel, H.P., Ng, R.T. and Sander, J., 2000. LOF: Identifying density-based local outliers. *ACM SIGMOND Record*, [e-journal] 29(2), pp.93-104. DOI: 10.1145/342009.335388.

Foidl, H. and Felderer, M. 2015. Research challenges of industry 4.0 for quality management. In: M. Felderer, F. Piazolo, W. Ortner, L. Brehm, HJ. Hof, eds., *Proceedings of the International Conference on Enterprise Resource Planning Systems*. Munich, Germany. 16-17 November. pp.121-137. DOI: 10.1007/978-3-319-32799-0_10.

Hotelling, H., 1947. Multivariate quality control - illustrated by the air testing of sample bombsights. In: C. Eisenhart, M.W. Hastay and W.A. Wallis, eds. 1947. *Techniques of Statistical Analysis*. New York: McGraw-Hill. pp.113-184.

Ide, T., Lozano, A.C., Abe, N. and Liu, Y., 2009. Proximity-based anomaly detection using sparse structure learning. In: Ch. Apte, H. Park, K. Wang and M.J. Zaki, *Proceedings of the 2009 Society for Industrial and Applied Mathematics (SIAM) International Conference on Data Mining*. Sparks, Nevada. 30 April – 02 May. pp.97-108. DOI: 10.1137/1.9781611972795.9.

Inoh, J., Nagata, Y., Horita, K. and Mori, A., 2012. Prediction accuracies of improved Taguchi's T methods compared to those of multiple regression analysis. *Journal of The Japanese Society for Quality Control*, [e-journal] 42(2), pp.265-277. DOI: 10.20684/quality.42.2_265.

Jackson, J.E. and Mudholkar, G.S., 1979. Control procedures for residuals associated with principal component analysis. *Technometrics*, [e-journal] 21(3), pp.341-349. DOI: 10.1080/00401706.1979.10489779.

Jin, X. and Chow, T.W., 2013. Anomaly detection of cooling fan and fault classification of induction motor using Mahalanobis–Taguchi system. *Expert Systems with Applications*, [e-journal] 40(15), pp.5787-5795. DOI: 10.1016/j.eswa.2013.04.024.

Lauritzen, S.L., 1996. *Graphical Models*. Oxford: Oxford University Press.

Ohkubo, M. and Nagata, Y., 2017. The Mahalanobis–Taguchi method based on graphical modeling. *Japanese Journal of Applied Statistics*, [e-journal] 46(1), pp.13-26. DOI: 10.5023/jappstat.46.13.

Ohkubo, M. and Nagata, Y., 2018. Anomaly detection in high-dimensional data with the Mahalanobis–Taguchi system. *Total Quality Management & Business Excellence*, [e-journal] 29(9-10), pp.1213-1227. DOI: 10.1080/14783363.2018.1487615.

Park, S.H., Shin, W.S., Park, Y.H. and Lee, Y., 2017. Building a new culture for quality management in the era of the Fourth Industrial Revolution. *Total Quality Management & Business Excellence*, [e-journal] 28(9-10), pp.934-945. DOI: 10.1080/14783363.2017.1310703.

Peng, C.F., Ho, L.H., Tsai, S.B., Hsiao, Y.C., Zhai, Y., Chen, Q., Chang, L.C. and Shang, Z., 2017. Applying the Mahalanobis–Taguchi system to improve tablet PC production processes. *Sustainability*, [e-journal] 9(9), 1557. DOI: 10.3390/su9091557.

Porter, M.E. and Heppelmann, J.E., 2014. How smart, connected products are transforming competition. *Harvard Business Review*, 92(11), pp.64-88.

Schölkopf, B., Platt, J.C., Shawe-Taylor, J., Smola, A.J. and Williamson, R.C., 2001. Estimating the support of a high-dimensional distribution. *Neural Computation*, 13(7), pp.1443-1471.

Schwarz, G., 1978. Estimating the dimension of a model. *The Annals of Statistics*, [e-journal] 6(2), pp.461-464. DOI: 10.1162/089976601750264965.

Shin, W.S., Dahlgaard, J.J., Dahlgaard-Park, S.M. and Kim, M.G., 2018. A quality scorecard for the era of industry 4.0. *Total Quality Management & Business Excellence*, [e-journal] 29(9-10), pp.959-976. DOI: 10.1080/14783363.2018.1486536.

Taguchi, G., Chowdhury, S., and Wu, Y., 2005. *Taguchi's Quality Engineering Handbook*. Hoboken, NJ: Wiley.

Taguchi, G. and Jugulum, R., 2002. *The Mahalanobis–Taguchi Strategy: A Pattern Technology System*. New York: John Wiley & Sons.

Takahama, M. and Mikami, N., 2012. Detection of Abnormal Signs for Gas Turbine Power Plant. *Journal of Quality Engineering Society*, [e-journal] 20(4), pp.437-443. DOI: 10.18890/qes.20.4_45.

Woodall, W.H., Koudelik, R., Tsui, K.L., Kim, S.B., Stoumbos, Z.G. and Carvounis, C.P., 2003. A review and analysis of the Mahalanobis—Taguchi system. *Technometrics*, [e-journal] 45(1), pp.1-15. DOI: 10.1198/004017002188618626.

Yamanishi, K. and Takeuchi, J.I., 2002. A unifying framework for detecting outliers and change points from non-stationary time series data. In: O.R. Zaïane and R. Goebel, *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Edmonton Alberta, Canada, July, 2002. New York, NY: Association for Computing Machinery. pp.676-681. DOI: 10.1145/775047.775148.

ABOUT AUTHORS

Masato Ohkubo – (M.O.) Toyo University, Tokyo, Japan, Department of Business Administration, Faculty of Business Administration, e-mail: m.ohkubo@toyo.jp, Author's ORCID: 0000-0002-1885-6945.

Yasushi Nagata – (Y.N.) Waseda University, Tokyo, Japan, Department of Industrial and Management Systems Engineering, School of Creative Science and Engineering, e-mail: ynagata@waseda.jp, Author's ORCID: 0000-0002-4947-1204.

AUTHOR CONTRIBUTIONS

M.O. – conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing—original draft preparation, visualization; Y.N. – writing—review and editing, supervision; M.O. and Y.N. – project administration, funding acquisition.

CONFLICTS OF INTEREST

The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Framework to Evaluate Continuous Improvement Process Efficacy: A Case Study of a Capital Goods Company

DOI: 10.12776/QIP.V24I2.1436

Gorka Unzueta, Aritz Esnaola, José A. Eguren

Received: 2020-03-29 Accepted: 2020-04-25 Published: 2020-07-31

ABSTRACT

Purpose: This document describes a continuous improvement process assessment system (CIPAS). A continuous improvement process (CIP) was developed to progress through the levels of continuous improvement (CI) defined by Bessant, Caffyn and Gallagher (2001), and the CIPAS was developed to measure this evolution. The CIP and the CIPAS were tested in a mature industrial small and medium-sized enterprise (SME) cooperative company (Basque Country, Spain) that works in the capital goods sector.

Methodology/Approach: The study was developed according to an 'action research' strategy (Coughlan and Coghlan, 2002) over a period of two years. The action research team includes the authors and managers of several areas of the studied company.

Findings: The assessment identified critical elements and related routines for the effective execution of the CIP in this company. In addition, the evaluation system allowed for a visualisation of the company's CI maturity level progression.

Research Limitation/implication: The assessment system was designed in an ad hoc manner for this CIP and this industrial company, but it may be possible to adapt these to other types of companies by using the steps followed and indicators defined as an example.

Originality/Value of paper: The CIPAS is used to identify the key CI elements, to measure the evolution of CI routines, and to identify a CI maturity level of the company in which the CIP is applied. It can be applied to any type of company and serves to define future actions for its evolution.

Category: Case study

Keywords: continuous improvement assessment; continuous improvement process; industrial case study; action research

1 INTRODUCTION

The results of implementing continuous improvement (CI) have been well documented in the literature (Bhuiyan and Baghel, 2005; Marin-Garcia, Pardo del Val and Martín, 2008; Singh and Singh, 2015). However, in many cases, once CI techniques have been applied, the initial results are not maintained and become much less effective as time passes (Dale, 2015). For this reason, many authors insist that CI systems must be adapted to individual organisations (Singh and Singh, 2015). Organisation must define a CI deployment strategy to select the appropriate CI method and tools, and they must monitor and follow up in a structured manner with the CI system to develop a CI culture and CI routines and behaviours within the organisation (Bateman and Rich, 2003; Wu and Chen, 2006).

According to Jorgensen, Boer and Laugen (2006), an increased CI maturity level – in terms of the level of adoption of CI routines and behaviours – corresponds to an increased level of performance in the organisation. Several models have been developed to implement CI systems and explain CI sustainability (Upton, 1996; Bessant, Caffyn and Gallagher, 2001; Ljungstrom, 2005; Wu and Chen, 2006; Kumar, Antony and Tiwari, 2011; Egiguren, 2012; McLean, Antony and Dahlgaard, 2017). Taking as a reference previous models, the research team developed a continuous improvement model (CIM) to deploy improvement routines and develop a CI organisational culture, with the objective of increasing CI maturity. Furthermore, a continuous improvement process (CIP) was developed according to this model to define the steps and activities to follow to achieve this CI culture. To evaluate the evolution of CI organisational culture, the present study developed a continuous improvement process assessment system (CIPAS). The current paper presents the aforementioned work and how over a period of two years it was applied in an industrial small and medium sized enterprise (SME) company, here implementing the CIM to overcome the 2nd level and to establish the bases of the 3rd level of CI maturity.

The current paper is organised as follows. In section 2, the research methodology is presented. In section 3, the CIM and CIP are presented, and in section 4, the CIPAS is shown. In section 5, results are drawn from the application of the CIPAS, and finally, section 6 provides the conclusions.

2 METHODOLOGY AND SCOPE OF APPLICATION

The methodology was based on a case study (CS) investigation and embedded multiple units of analysis (Yin, 2013), in the same context was analysed eight units. The research was carried out in a single organisation or context, where the same process was replicated in multiple areas – each one considered a unit of analysis (UA) – to achieve a global understanding of the situation and the change caused in the organisation as a result of the changes carried out in each area. The

CIPAS was developed to assess the implementation and evolution of the CIP. The UAs were divided into three groups based on their common characteristics.

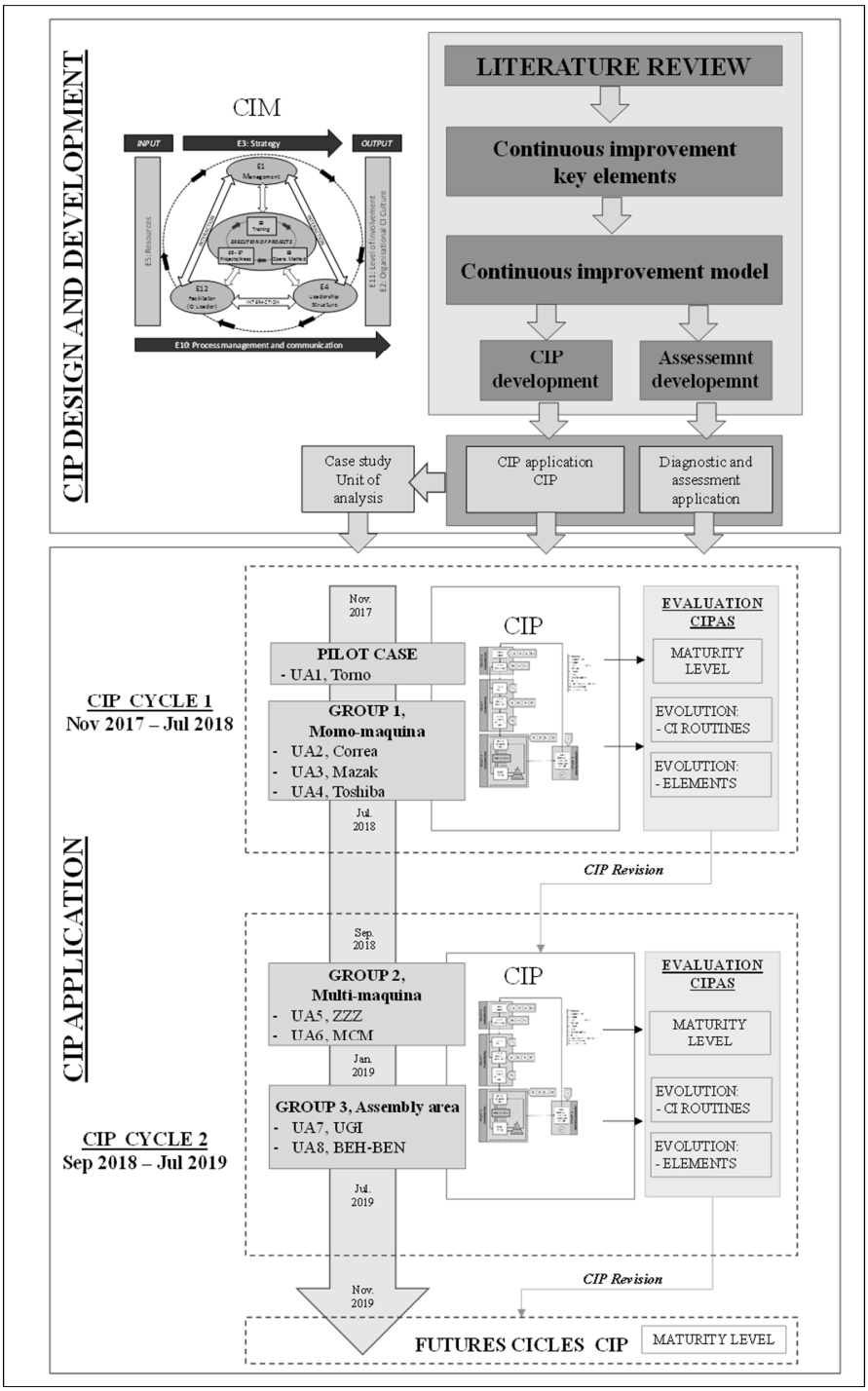

Figure 1 – Research Methodology Steps

Figure 1 shows in detail the deployment of the methodology, it can be seen that a review of the literature was carried out to identify the elements (E). Based on the elements, the CIM was developed (Figure 2), and based on this model, a CIP and an evaluation system (CIPAS) were proposed. The CIP is an annual cyclic process consisting of four stages, as shown in Figure 2. The current paper shows the results of the execution of two cycles from November 2017 to November 2019, in which eight units, which were grouped into three groups, were analysed.

The main objective of this work is to describe the CIPAS developed to evaluate the implementation and evolution of the CIP. The CIP was applied to an industrial SME cooperative in the capital goods sector. The research team worked directly in the company, applying the action research (AR) methodology. The AR cycles coincided with the CIP review and improvement cycles.

3 CONTINUOUS IMPROVEMENT

CI is the planned, organised and systematic process of ongoing, incremental and companywide change in existing practices and aims to improve company performance (Jorgensen, Boer and Laugen, 2006). Numerous studies have identified the elements and critical success factors (CSF) that must be taken into account when successfully applying CI, and most of them coincide with the key elements. However, various studies have highlighted the importance of various elements depending on the focus of the study and its scope of application (large or small and medium enterprises, public or private, etc.). To define the key elements of the model, a study carried out in a similar context (mature industrial companies in the Basque Country) (Egiguren, 2012) was taken into consideration as a reference. This study was developed in a large automotive supply company, while our CS focused on a capital goods sector SME. For this reason, a review of the elements was reinforced with the most recent contributions, as well as reviews published by several authors regarding the CSFs for the sustainability of CIP. The most important elements to develop a CIM and the principal authors are summarised as follows:

E1: Management. The involvement and commitment of the management are necessary (Readman and Bessant, 2007; Garcia-Sabater, Marin-Garcia and Perello-Marin, 2012; Gonzalez Aleu and Van Aken, 2015; Costa et al., 2019; Stankalla, Koval and Chromjakova, 2018).

E2: Company culture. New behaviours and routines must be developed among all members of the organisation (Bateman, 2005; Egiguren, 2012; McLean and Antony, 2017).

E3: Strategy. The CIP must be a strategic element of the operating plan, and the strategic goals and CI project goals must be aligned (Bessant, Caffyn and Gallagher, 2001; Jurburg et al., 2019).

E4: Leadership and structure. The organisation must have a CI organisational structure (Wu and Chen, 2006; Fryer, Ogden and Anthony, 2013; Lodgaard et al., 2016; Stankalla, Koval and Chromjakova, 2018).

E5: Resources. The organisation must make the necessary resources available: economic, time and training (Bateman and Rich, 2003; Wu and Chen, 2006).

E6: Projects and *E7: Areas*. Improvement projects should be clear, realistic and focused on critical areas or processes (Egiguren, 2012; Lodgaard et al., 2016), where the need for improvement should be clearly seen (Bateman and Rich, 2003).

E8: Operating method and improvement tools. It is necessary to adapt the operating method to the situation within each organisation (Bhuiyan and Baghel, 2005; Dale, 2015) to enhance the participation of the people in the CI activities (Jurburg et al., 2019). The tools must align with the maturity of the organisation in terms of CI and production (Wu and Chen, 2006; Kosieradzka and Ciechańska, 2018).

E9: Training. Training should be based on the operating method (Gonzalez Aleu and Van Aken, 2015; McLean and Antony, 2017; Costa et al., 2019) and aligned to improvement projects.

E10: Monitoring and communication. A CIP follow-up process must be established. This should measure two aspects: the effectiveness of the results from the improvement project and the evolution of the CIP (Bessant, Caffyn and Gallagher, 2001; Gonzalez Aleu and Van Aken, 2015).

E11: Level of involvement. The involvement of management and all employees is essential to achieve long-term improvement (Garcia-Sabater, Marin-Garcia and Perello-Marin, 2012; Costa et al., 2019). A systematic process and the application of improvement tools help to involve the participants (Jurburg et al., 2019).

E12: Facilitator (CI leader). The CI leader must take on the role of facilitator, guiding and coaching the promoter and project teams before and during projects (Garcia-Sabater, Marin-Garcia and Perello-Marin, 2012; Heavey, Ledwith and Murphy, 2014; Gonzalez Aleu and Van Aken, 2015). However, until the CI leader acquires the necessary skills, the facilitator can be a person outside the organisation.

CI refers not only to the outcomes, but also to the process by which these can be achieved (Bessant, Caffyn and Gallagher, 2001). Bessant, Caffyn and Gallagher (2001) proposed an evolutionary CI maturity model that is divided into five levels, in which organisations can be placed according to the maturity of the organisation in terms of CI. The evolutionary model promotes the development of an organisational culture that enables the organisation to move towards excellence in CI by adopting eight routines at the organisational level, R1 ‘understanding CI’, R2 ‘getting the CI habit’, R3 ‘leading the way’, R4 ‘focusing

CI', R5 'aligning CI', R6 'shared problem solving', R7 'CI of CI' and R8 'the learning organisation'. To progress between the levels, the organisation must consolidate and assume as the natural course of operations the routines generated while still creating new routines that take the organisation to the next level. Table 1 shows the description of each maturity level.

3.1 Continuous Improvement Model

The relation of the selected elements to each other was defined to create the CIM. As shown in Figure 2, the execution of projects is the central point of the model. The selected operative method (E8), training (E9) and projects and areas (E6, E7) are necessary 'tools' to develop the activities that generate the CI routines. To use these 'tools' correctly, it is necessary to have the support of the management (E1), to have a defined CI organisational structure (E4) including a CI leader as a facilitator (E12). The management, the CI leader and the defined CI structure are responsible for defining a strategy to deploy the CIM, correctly manage the resources and control the CIP and each project to ensure the involvement of people and develop a CI organisational culture.

According to Bessant, Caffyn and Gallagher (2001), the assimilation and evolution of the eight improvement routines increases the maturity level of the organisation. Fryer, Ogden and Anthony (2013), in the same line, stated that the evolution of several key elements to develop CI also increases the level of maturity. The development and evolution of CI routines and the evolution of CI maturity demonstrate the constant development of a CI organisational culture (Bessant and Caffyn, 1997; Bessant, Caffyn and Gallagher, 2001).

3.2 Continuous Improvement Process

To create and assimilate new CI routines, it is necessary to apply improvement tools and methods through a systematic and structured process. CIP refers to the process for carrying out the CIM and was configured in four stages. The phases to be executed in each stage were identified, along with the elements that exert a significant influence in each stage (Figure 2). CIP is a process that repeats itself cyclically with a certain periodicity. In the CS, the determined period is annual. The improvement stage (Stage 3) serves as an input for the diagnosis of the next cycle.

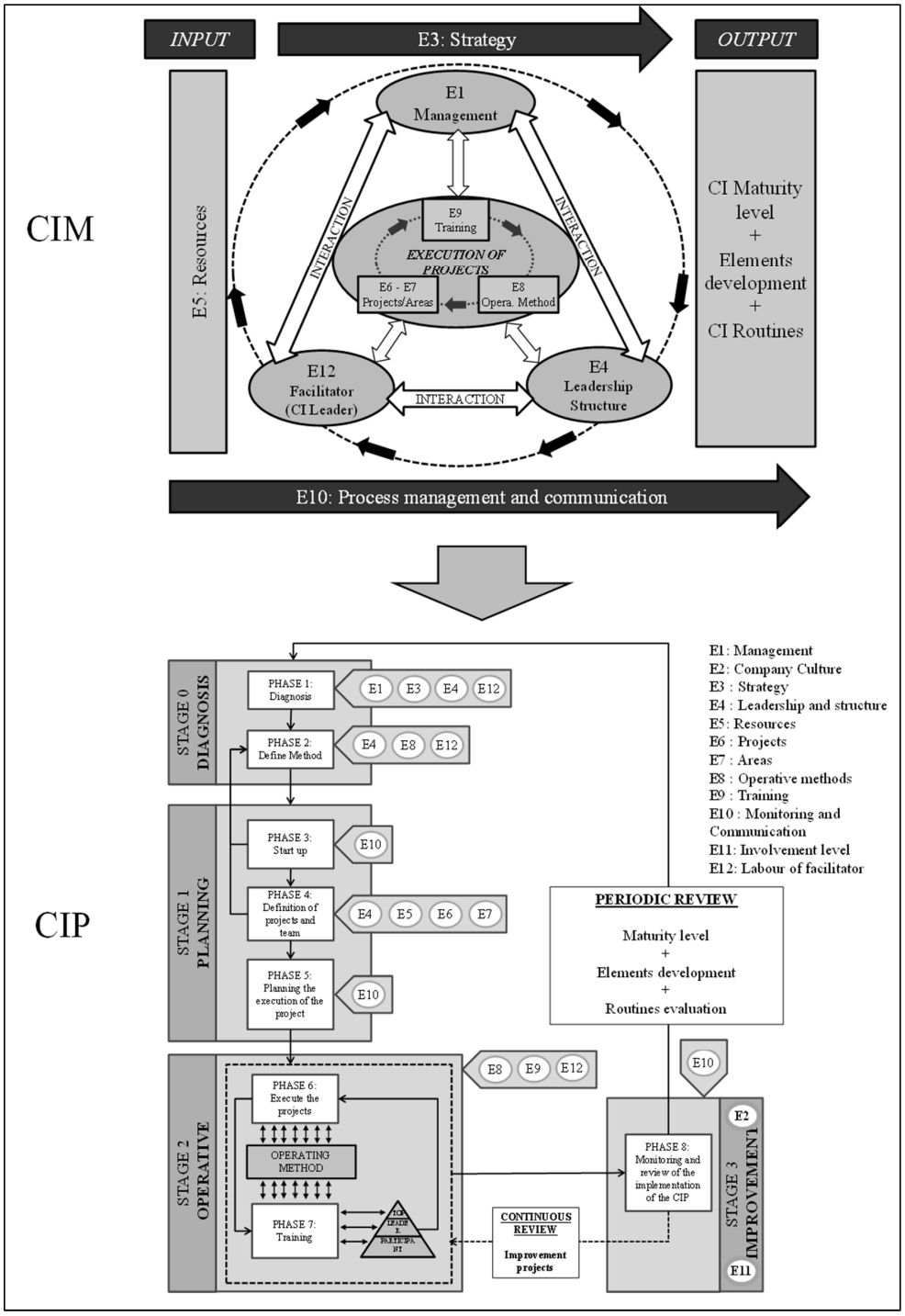

Figure 2 – CIM and CIP

3.2.1 STAGE 0: Diagnosis

The management must make the diagnosis and select the appropriate operating method to develop the CI system. The diagnostic stage was divided into two phases: diagnosis (phase 1) and definition of the operating method (phase 2).

Phase 1 (diagnosis) was based on the evaluation of the CI maturity level of the organisation. The maturity was measured using a questionnaire (presented in Appendix, Table A1), where in each level, the respondent have to answer 'true' or 'false' for several statements. Each level have ten items, and depending the number of 'trues', it is possible to identify the CI maturity status. The design of the questionnaire ensures that responding as 'true' to the items of a level is not possible if most of the items of the lower level have not been answered as 'true'.

In *Phase 2* (define method), while considering the CI maturity, production maturity levels (Kosieradzka and Ciechańska, 2018) and improvement tools used previously, among other things, management must identify the appropriate operating method. Table 1 summarises the principal tools and methods applied at each level.

Table 1 – Methods and Tools Used for Each Maturity Level (Source: Based on Bessant, Caffyn and Gallagher (2001), Kosieradzka and Ciechańska (2018))

	CI maturity level	**Production maturity level**	**Methods and tools used at a given level**
Level 1	*Pre-CI:* • Interest in the concept of CI. Nevertheless, implementation is on an ad hoc basis. No formal efforts or structure for improving the organisation.	*Performed production processes:* • Processes are not iterative or predictable. Impossible to control progress.	• Using employees' tacit knowledge • General control
Level 2	*Structured CI:* • There is a formal commitment to building a CI system. CI initiatives have been introduced.	*Manage production processes:* • Targets are met as a result of implementing a plan. Progress is monitored.	• 5S • Standard operating procedures • Autonomous maintenance • Provisioning Kanban • Quality goals and standards • Seven quality tools • OEE • Training plan
Level 3	*Goal-oriented CI:* • Commitment to linking CI behaviours to the strategic concerns of the organisation. Formal deployment of strategic goals. Monitoring and	*Defined production processes:* • Targets are met in processes defined in line with the process–approach parameters.	• Production and supply process maps • Value stream mapping (VSM) • Waste identification and elimination

	CI maturity level	Production maturity level	Methods and tools used at a given level
	measuring CI against these goals. CI activities are part of the main business activities. Most of the staff participate in CI activities.		• Work station layout adjusted to the process requirements • Collecting data on quality • Maintenance system (TPM, RCM) • Kaizen events (Blitz)
Level 4	*Proactive CI:* • There is an attempt to devolve autonomy and to empower individuals and groups to manage and direct their own processes. High level of experimentation.	*Quantitatively managed production processes:* • Quantitative and qualitative targets and performance control tools defined for processes and operations.	• Quality and productivity measures established • SPC identification of special causes • DMAIC (Six Sigma) • SMED • Production Kanban • Presentation of productivity and quality performance
Level 5	*Full CI capability:* • Approximates to a framework of a 'learning organisation'. Extensive learning behaviours, systematic finding and solving of problems and capturing and sharing of learning.	*Optimised production processes:* • Processes are continuously improved and adapted to the changing environment and corporate strategy.	• Process re-engineering • Kaizen in the whole organisation • SPC identification of common causes • DMAIC (Six Sigma) • Teamwork, culture of 0 defects, 0 equipment failure, 0 time waste

Stage 0 (diagnosis) ensured that the proposed operational methods and projects were aligned with the organisation's strategy and focused on the customer (external and internal).

In the CS, after regular meetings between the research team and top management to analyse the organisation's situation and after using the maturity level questionnaire to develop the interviews, the first diagnosis found that the organisation was at level 1. The company did not have an organisational structure to develop CI, the company did not have a defined CI leader, problems were solved as they arose using an employee's tacit knowledge without a strategy to deploy CI activities, the proposed solutions had a short path, and generally, management and employees were not concerned about CI activities. The studied company set itself the target of surpassing maturity level 2 within a period of two years. To this end, the research team together with the management defined a strategy to deploy the system, proposed defining an organisational structure with responsibilities related to improvement activities, apply an operating method to define the bases of the CI system and train employees and middle management in basic improvement tools.

The improvement tool selected to start defining the basis for CI was 5S. The 5S methodology is a lean tool developed by Japanese manufacturing companies comprising five stages: sort (*seiri*), set in order (*seiton*), shine (*seiso*), standardise (*seikats*u) and sustain (*shitsuke*) (Khan et al., 2019). The methodology is simple to apply, facilitates the participation of employees in the area and promotes teamwork. Many authors have pointed to 5S as a suitable tool for this purpose (Randhawa and Ahuja, 2017; Khan et al., 2019).

3.2.2 STAGE 1: Planning

In the planning stage, after identifying the signs and behaviours that ensure the support of those responsible (department managers) for the various areas within the organisation – which was done through specific training – the chosen operative method was adapted to the reality of the organisation. In addition, the channels and activities that facilitate communicating the characteristics and benefits of the CIP to all personnel were developed (Gonzalez and Martins, 2016; Stankalla, Koval and Chromjakova, 2018). Finally, the planning of each project was developed.

3.2.3 STAGE 2: Operative

In the operative stage, two phases were developed in parallel: the execution of the projects and the training of the participants. Execution was developed according to the project plan, and training was adapted to the operative method and the various organisational structure levels (see Table).

Table 2 – Developed Training Summary

Level	**Participants**	**Content**	**STAGE**
Top management	• Chief executive officer • Industrial director • Quality director	• Awareness talk about CI • General structure CIM, CIP • General planning	STAGE 0
Middle management	• Quality director (management representative as a CI leader) • Several department managers: production, planning, process engineering, quality, etc.	• Awareness talk about CI • CI levels, CIM, CIP • General planning • Brief training on the selected methodology (5S)	STAGE 1
Project team members	• Quality director (CI leader) • Project leaders • Project team members	• Formal training on selected method in depth: awareness and technical (5S)	STAGE 2

3.2.4 STAGE 3: Improvement

Organisations should analyse CI activity periodically to understand its weaknesses and implement corrections. Self-examination is the most effective way to achieve successful CI (Jorgensen, Boer and Laugen, 2006; Wu and Chen,

2006). As Wu and Chen (2006) suggested, the evaluation system and specific metrics used should be adapted to each CI level. Although the general evaluation system's structure can be the same, depending on the CI maturity level and the operative method applied, the metrics that are used can change. The CIPAS was developed according to these principles.

4 CONTINUOUS IMPROVEMENT PROCESS ASSESSMENT SYSTEM

The objective of the CIPAS is to analyse the evolution of CI organisational culture development, based on the evolution of the maturity level of CI, the assimilation and evolution of the CI routines and the evolution of the CIM key elements over time. The CIPAS measures three main aspects:

- CI maturity level
- CI routines development and assimilation
- CIM key elements development

4.1 CI Maturity Level

The maturity level was measured at the beginning of each CIP cycle. The assessment of the maturity level was based on a questionnaire (Appendix, Table A1) that must be completed by the promoter team, which was led by the CI leader, at the beginning of each cycle.

4.2 Evaluation of CI routines

According to Bessant, Caffyn and Gallagher (2001), to move up maturity levels, it is necessary to assimilate CI routines through the acquisition of skills and behaviours related to these routines. The evolution of these routines was evaluated using a questionnaire (Appendix, Table A3) that is based on the 'constituent behaviours', as defined by Bessant, for each of the routines (Bessant, Caffyn and Gallagher, 2001). The promoter team led by the CI leader must complete the questionnaire at the end of each CIP cycle.

4.3 Evaluation of CIM Key Elements

To evaluate the key elements of the CIM, two strategies were defined. First, the elements that must be kept constant during CIP implementation were defined. To measure the evolution of the other elements, indicators were defined for each of them.

Elements to keep constant: At the beginning of CIP implementation, several rules were defined to keep these elements constant.

Table 3 shows these elements, their description, why they should be kept constant and the rules defined to keep them constant. This reflection is necessary in each CIP cycle.

Valued elements: These elements were valued by means of indicators. The CI leader, the project leaders and the research team collected the information through direct observation and a survey completed by the employees. Three types of elements were evaluated: XV: Elements assessed in which the research team did not act directly, but which influenced their development; XA: Elements in which the research team acted directly for their correct development; Y: Output elements, resulting from the development of input elements XV and XA. Table 4 shows the valued elements, their description and the corresponding metrics of each element.

Table 3 – Elements that Must be kept Constant

Elements to observe	**Description**	**Why do you want to keep constant**	**Guidelines followed to keep constant**
E1:Management *E3: Strategy*	The CIP must be under the responsibility of the management. The organisation must deploy the CIP with a strategic vision, must select the operating method, and must find the areas and projects that will respond to the strategic objectives. The management must keep the activities aligned with the strategic objectives of the organisation.	Changes in strategy or management's vision during CIP implementation can confuse the team and employees and hinder implementation.	First, meetings with the management to define the strategic vision, the objectives to be achieved and the strategy for the deployment of the CIP. Establish project and CIP monitoring guidelines. Training meetings with management to carry out the diagnosis and establish the operational method to be applied. Participation of a management representative in follow-up and communication meetings.
E10: Follow-up and communication	The CIP is a process of change through the modification and establishment of routines and behaviours. Therefore, it is necessary to have a monitoring and communication system.	Changes in the monitoring (follow-up) system during CIP implementation can confuse the team and employees and be a barrier to their participation.	Rules to follow up were established to evaluate the improvement projects launched. Transparent and common communication channels were established for all projects.

Table 4 – Elements to Measure Its Evolution

Element Valued	Description	Metrics
E4: Leadership and strategy	Organisational structure developed for the CI, divided into three levels (Management, promoter team, and project teams). Leadership is divided into two levels – management and project leaders – and each type of leadership has its responsibilities.	**XV-E4-1:** Nº of changes in the team structure (0%–100%) **XV-E4-2:** Nº of meeting/month **XV-E4-3:** % Attendance at meetings
E5: Resources	Management must display the resources necessary to execute the CIP. Time to execute the improvements, release of the people involved in the CIP and economic resources to address the investments derived from the proposed improvements.	**XV-E5-1:** Hours dedicated to developing project **XV-E5-2:** Hours dedicated to follow up
E6–E7: Projects and areas	Management must select the projects and areas that are the most critical for the organisation. For selection must be considered: • The complexity of the area • The saturation level of the workers • The size of the area	**XV-E67-1:** Complexity of the areas (initial audit of the area) **XV-E67-2:** Operator saturation level **XV-E67-3:** Area size and quantity of means
E8: Operative method	The operating method must be adapted to the organisation. Each project must be managed efficiently, following the defined phases. The operating method must be evaluated by the participants.	**XA-E8-1:** % of phases executed effectively **XA-E8-2:** Score obtained in the survey/maximum possible score
E9: Training	A training plan must be defined. The training must be applied in the real environment (project). The training must be evaluated by the participants (survey).	**XA-E9-1:** % attendance to regulated formations **XA-E9-2:** Score obtained in the survey/maximum possible score
E11: Level of involvement	The leader must devote time (determined beforehand by the promoter team) to the control and follow-up of the project. The motivational work done by the leader must be evaluated by the project team (survey). The number of improvement proposals made by the participants must be gathered. Sensation of the participants applying the improvement methodology must be gathered (survey rating).	**Y-E11-1:** Hours dedicated to follow up/month **Y-E11-2:** Score obtained in the survey/maximum possible score **Y-E11-3:** Number of suggestions made by the participants of the area/first three months **Y-E11-4:** Score obtained in the survey/maximum possible score

Element Valued	Description	Metrics
E12: Labour of facilitator	The facilitator, as an expert in CI, develops the following functions: • Train the participants of the CI structure (management, leaders, employees). • Track the deployment of the CIP. • Support project leaders in the execution of the operational method and in the development of meetings and presentations. The labour of the facilitator must be evaluated by the participants (survey).	**XA-E12-2:** Score obtained in the survey/maximum possible score

The valuation of the metrics presented in Table 4 are represented in percentages. These percentages were calculated by taking 100% as the reference: the highest value of the item, the highest possible value on the Likert scale that was used in the survey or the level of compliance with the defined standard (e.g., the execution of defined meetings or the level of attendance at meetings).

5 RESULTS

The results of the evaluation confirm that through the implementation and application of the CIP, the organisation increased its CI maturity level. Management, middle management and employees developed several improvement routines that were further deployed and assimilated in the second cycle. Like the CIPAS, the results are divided into three aspects: the evaluation of the maturity level, the evaluation of the CI routines and the evaluation of the key elements of the CIM.

5.1 Maturity Level Evolution

The maturity level was measured three times at the beginning of each cycle. In the first cycle, with the direct support of the research team, management completed the evaluation. This evaluation was carried out after training the management team on issues related to CI. The second and third evaluation was completed by the promoter team, which was led by the CI leader with the support of the research team (Figure 3).

The periodic evaluation shows an increase in the level of maturity of the organisation. The first diagnosis indicated that the organisation was at level 1, with a value of 0.5, which increased as the CIP was deployed, until it surpassed level 2 with a value of 2.2. Among other things, the definition of an organisational structure dedicated to CI (defined specific roles to each participant), the definition of the follow-up rules and the application of an

operational method in a systematic and structured manner justifies this increase in maturity.

5.2 CIM Key Elements Evolution

The CIM key elements were measured at the end of each cycle on two occasions. Figure 3 shows how all elements, except for E12 (facilitator), positively evolved as CIP implementation progressed. After the first cycle, the poorest results were obtained for elements E4 (leadership and structure), E5 (resources), E6–7 (projects and areas) and E11 (involvement). In the second cycle, actions were taken to reinforce these elements. E4 was strengthened by officially introducing the meetings of the promoter team in the calendar of monthly periodic meetings. The CI leader led these meetings, where each project leader reported the evolution of his project. The periodic meetings of the promoter team encouraged the execution of the weekly meetings of the improvement teams. To improve E5, the management, through the promoter team, allocated more time to the implementation of improvement activities, especially to leaders and project participants. The CI leader increased his participation in project follow-up meetings, boosting their execution and ensuring the participation of team members. With the actions proposed, the level of employee involvement (E11) and participation increased, increasing the number of suggestions for improvement.

Through a survey, employees evaluated the work carried out by the facilitator. The principal investigator played the role of external facilitator, decreasing his participation in the second cycle as the CI leader increased his participation and took on the role of feacilitator.

5.3 CI Routines Evolution

On two occasions, the assimilation of routines was measured at the end of each cycle. In the CS, the promoter team led by the CI leader completed the questionnaire presented in Appendix (Table A3), with the support of the research team. Garcia-Sabater, Marin-Garcia and Perello-Marin (2012) structured the necessary routines to be developed at each maturity level. Based on this study, the researchers could quantify, the results, and the necessary level of assimilation of each routine for 2nd maturity level are shown in Figure 3.

As can be seen in Figure 3, the assimilation of routines evolved positively although in all routines, the objective set for maturity level 2 was not achieved. The company should continue to work to encourage employee participation in CI by applying new improvement tools. In areas where suggestion management systems, manufacturing process measurement and improvement activities or the supply Kanban method was implemented, the results of the ‘getting the CI habit’ and ‘leading the way’ routines were better. It is also necessary to focus on CI by aligning projects with the organisation’s strategic goals. During the first two cycles, the projects were oriented towards the definition of operating standards

and the standardisation of workplaces rather than towards the improvement of critical processes that can be focused on by achieving specific strategic objectives. Once the improvement dynamics were established during the first two cycles, in the following cycle, the improvement teams created were multidisciplinary, involving employees at different organisational levels, with the aim of developing more specific projects (changing layouts, improving manufacturing processes, developing self-control in critical processes, etc.). The execution of these new projects allows for the development of the 'shared problem solving' and 'focusing CI' routines.

However, Jorgensen, Boer and Laugen (2006) declared that the maturity model does not have to follow a linear progression. The results of the present CS support this affirmation. As can be seen in Figure 3 there are routines, such as 'aligning CI' and 'CI of CI' -which in theory correspond to higher maturity levels- that are more assimilated than others like 'CI focusing'.

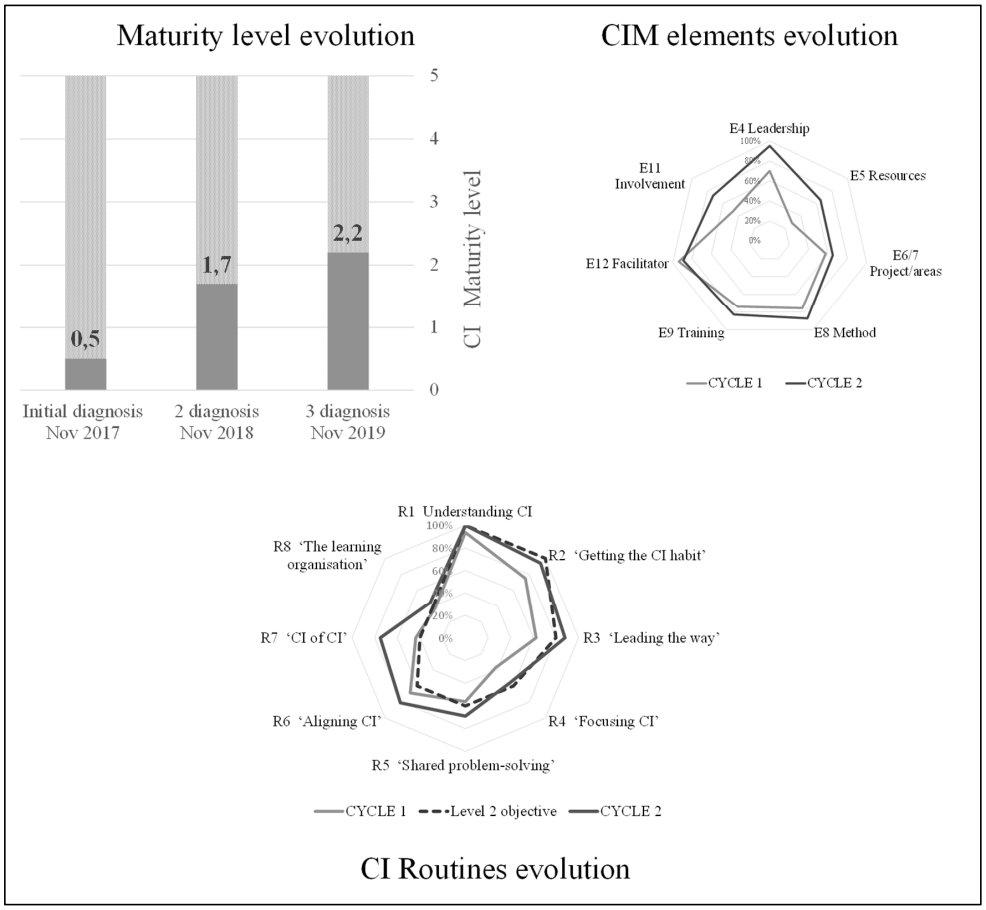

Figure 3 – Maturity Level, CIM Key Elements and CI Routines Evolution

6 CONCLUSION

The current field research shows that the improvement process (CIP) that was developed and implemented helped develop improvement routines, advance the CI maturity level and develop the basis of an organisational CI culture. The CIP served to meet the objective set by the research team and management, overcoming level 2 of CI maturity. The application of the CIP allowed for acquiring and assimilating improvement routines and reinforcing the key elements of the CIM. In the first two cycles, the bases of the CI system were established. An organisational structure adapted to the organisation itself was defined, based on a promoter team and led by a CI leader. During the first cycle, the roles and rules of the team's operations were established, and in the second cycle, the team members assimilated the roles and rules. Reinforcing the teamwork and involvement of the promoter team made it possible to define and launch new projects (self-control system, advanced product quality planning (APQP), picking logistics systems and automated storage, etc.) and create improvement teams to respond to the problems identified in the evaluation and diagnosis. To articulate the relations between the promoter team and the improvement teams, the participation of management through the CI leader was key.

On the other hand, the proposed measurement system (CIPAS) served to measure the evolution of the CIM that was implemented in the organisation. The CIPAS made it possible to assess the organisation's level of maturity on the Bessant scale, to assess the evolution of improvement routines and to identify the key elements of the CIM that needed to be improved in each cycle. In addition, the CIPAS facilitated the identification of activities that reinforced the maturity of the organisation. The diagnostic and evaluation tools facilitated the visualisation of the progress of the CI system in a visual way. The greatest challenge in implementing the CIPAS was to make management, especially the promoter team, realise that to advance in CI maturity levels, it is necessary to measure the evolution of the CIP. To this end, it was necessary to work with the promoter team on the design of the CIPAS and to demonstrate that based on the evaluations, they were able to propose actions to improve CIP itself, in addition to launching new improvement projects.

In light of the findings, future research in the organisation should be conducted to discover what key elements of CI are still lacking or should be strengthened in higher maturity levels.

Regarding limitations, even if it is only an organization-specific implementation and is not possible to generalize the results, the step-by-step approach of the case study can be a reference to applied to other organisations. The developed questionnaires provide a practical approach to assess the maturity level and to check the development of the organization's routines. In addition, the CIP and the CIPAS can be used as a theoretical basis to adapt the CIM, and apply these in another type of organisations that use other business models.

REFERENCES

Bateman, N., 2005. Sustainability: The elusive element of process improvement. *International Journal of Operations and Production Management*, [e-journal] 25(3), pp.261-276. DOI: 10.1108/01443570510581862.

Bateman, N. and Rich, N., 2003. Companies' perceptions of inhibitors and enablers for process improvement activities. *International Journal of Operations & Production Management*, [e-journal] 23(2), pp.185-199. DOI: 10.1108/01443570310458447.

Bessant, J. and Caffyn, S., 1997. High involvement innovation through continuous improvement. *International Journal of Technology Management*, [e-journal] 14(1), pp.7-28. DOI: 10.1504/IJTM.1997.001705.

Bessant, J., Caffyn, S. and Gallagher, M., 2001. An evolutionary model of continuous improvement behaviour. *Technovation*, [e-journal] 21(2), pp.67-77. DOI: 10.1016/S0166-4972(00)00023-7.

Bhuiyan, N. and Baghel, A., 2005. An overview of continuous improvement: from the past to the present. *Management Decision*, [e-journal] 43(5), pp.761-771. DOI: 10.1108/00251740510597761.

Costa, F., Lispi, L., Staudacher, A.P., Rossini, M., Kundu, K. and Cifone, F.D., 2019. How to foster Sustainable Continuous Improvement: A cause-effect relations map of Lean soft practices. *Operations Research Perspectives*, 6, p.100091. DOI: 10.1016/j.orp.2018.100091.

Coughlan, P. and Coghlan, D., 2002. Action research for operations management. *International Journal of Operations & Production Management*, [e-journal] 22(2), pp.220-240. DOI: 10.1108/01443570210417515.

Dale, B., 2015. *Total quality management*. New York: John Wiley & Sons.

Egiguren, J.A.E., 2012. *Desarrollo de un modelo para abordar proyectos de mejora continua de procesos productivos de forma eficaz y eficiente*. PhD. Spain: Mondragon University.

Fryer, K., Ogden, S. and Anthony, J., 2013. Bessant's continuous improvement model: Revisiting and revising. *International Journal of Public Sector Management*, [e-journal] 26(6), pp.481-494. DOI: 10.1108/IJPSM-05-2012-0052.

Garcia-Sabater, J.P., Marin-Garcia, J.A. and Perello-Marin, M.R., 2012. Is Implementation of Continuous Improvement Possible? An Evolutionary Model of Enablers and Inhibitors. *Human Factors and Ergonomics in Manufacturing*, [e-journal] 22(2), pp.99-112. DOI: 10.1002/hfm.20262.

Gonzalez Aleu, F. and Van Aken, E.M., 2015. Systematic literature review of critical success factors for continuous improvement projects. *International Journal of Lean Six Sigma*, [e-journal] 7(3), pp.214-232. DOI: 10.1108/IJLSS-06-2015-0025.

Gonzalez, R.V.D. and Martins, M.F., 2016. Capability for continuous improvement, Analysis of companies from automotive and capital goods industries. *The TQM Journal*, [e-journal] 28(2), pp.250-274. DOI: 10.1108/TQM-07-2014-0059.

Heavey, C., Ledwith, A. and Murphy, E., 2014. Introducing a new continuous improvement framework for increased organisational return on investment. *The TQM Journal*, [e-journal] 26(6), pp.594-609. DOI: 10.1108/TQM-06-2013-0065.

Jorgensen, F., Boer, H. and Laugen, B.T., 2006. CI Implementation: An Empirical Test of the CI Maturity Model. *Creativity and Innovation Management*, 15(4), pp.328-337. DOI: 10.1111/j.1467-8691.2006.00404.x.

Jurburg, D., Viles, E., Tanco, M., Mateo, R. and Lleó, Á., 2019. Understanding the main organisational antecedents of employee participation in continuous improvement. *The TQM Journal*, [e-journal] 31(3), pp.359-376. DOI: 10.1108/TQM-10-2018-0135.

Khan, S.A., Kaviani, M.A., Galli, B.J. and Ishtiaq, P., 2019. Application of continuous improvement techniques to improve organization performance: A case study. *International Journal of Lean Six Sigma*, [e-journal] 10(2), pp.542-565. DOI: 10.1108/IJLSS-05-2017-0048.

Kosieradzka, A. and Ciechańska, O., 2018. Impact of enterprise maturity on the implementation of six sigma concept. *Management and Production Engineering Review*, [e-journal] 9(3), pp.59-70. DOI: 10.24425/119535.

Kumar, M., Antony, J. and Tiwari, M.K., 2011. Six Sigma implementation framework for SMEs-a roadmap to manage and sustain the change. *International Journal of Production Research*, [e-journal] 49(18), pp.5449-5467. DOI: 10.1080/00207543.2011.563836.

Ljungstrom, M., 2005. A model for starting up and implementing continuous improvements and work development in practice. *The TQM Magazine*, [e-journal] 17(5), pp.385-405. DOI: 10.1108/09544780510615915.

Lodgaard, E., Ingvaldsen, J.A., Aschehoug, S. and Gamme, I., 2016. Barriers to Continuous Improvement:Perceptions of Top Managers, Middle Managers and Workers. *Procedia CIRP*, [e-journal] 41, pp.1119-1124. DOI: 10.1016/j.procir.2016.01.012.

Marin-Garcia, J.A., Pardo del Val, M. and Martín, T.B., 2008. Longitudinal study of the results of continuous improvement in an industrial company. *Team Performance Management*, [e-journal] 14(1/2), pp.56-69. DOI: 10.1108/13527590810860203.

McLean, R.S. and Antony, J., 2017. A conceptual continuous improvement implementation framework for UK manufacturing companies. *International Journal of Quality & Reliability Management*, [e-journal] 34(7), pp.1015-1033. DOI: 10.1108/IJQRM-02-2016-0022.

McLean, R.S., Antony, J. and Dahlgaard, J.J., 2017. Failure of Continuous Improvement initiatives in manufacturing environments:a systematic review of the evidence. *Total Quality Management and Business Excellence*, [e-journal] 28(3-4), pp.219-237. DOI: 10.1080/14783363.2015.1063414.

Randhawa, J.S. and Ahuja, I.S., 2017. 5S implementation methodologies: literature review and directions. *International Journal of Productivity and Quality Management*, [e-journal] 20(1), pp.48-74. DOI: 10.1504/IJPQM.2017.080692.

Readman, J. and Bessant, J., 2007. What challenges lie ahead for improvement programmes in the UK? Lessons from the CINet Continuous Improvement Survey 2003. *International Journal of Technology Management*, [e-journal] 37(3/4), pp.290-305. DOI: 10.1504/IJTM.2007.012264.

Singh, J. and Singh, H., 2015. Continuous improvement philosophy:literature review and directions. *Benchmarking*, [e-journal] 22(1), pp.75-119. DOI: 10.1108/BIJ-06-2012-0038.

Stankalla, R., Koval, O. and Chromjakova, F., 2018. A review of critical success factors for the successful implementation of Lean Six Sigma and Six Sigma in manufacturing small and medium sized enterprises. *Quality Engineering*, [e-journal] 30(3), pp.453-468. DOI: 10.1080/08982112.2018.1448933.

Upton, D., 1996. Mechanisms for Building and Sustaining Operations Improvement A Brief Review of Operations Improvement Techniques. *European Management Journal*, 14(3), p.215-228.

Wu, C.W. and Chen, C.L., 2006. An integrated structural model toward successful continuous improvement activity. *Technovation*, [e-journal] 26(5-6), pp.697-707. DOI: 10.1016/j.technovation.2005.05.002.

Yin, R.K., 2013. *Case study research: Design and methods*. California: Sage Publications.

ABOUT AUTHORS

Gorka Unzueta – (G.U.) Mondragon University, Spain, Mechanical and Industrial Production department, e-mail: gunzueta@mondragon.edu, Author's ORCID: 0000-0002-6889-3293.

Aritz Esnaola – (A.E.) Mondragon University, Spain, Mechanical and Industrial Production department, e-mail: aesnaola@dumondragon.edu, Author's ORCID: 0000-0003-0955-7545.

José Alberto Eguren – (J.A.E.) Mondragon University, Spain, Mechanical and Industrial Production department, e-mail: jaeguren@mondragon.edu, Author's ORCID: 0000-0002-4127-8776.

AUTHOR CONTRIBUTIONS

All authors contributed equally to this article G.U., A.E and J.A.E. – conceptualization, resources, validation, writing review and editing.

CONFLICTS OF INTEREST

The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

APPENDIX

Each statement must be evaluated as a "true" or "false".

Table A1 – Maturity Level Questionnaire (1)

MATURITY LEVEL
LEVEL1
1. Problems are resolved as they arise (randomly).
2. Departmental teams are created to solve specific problems.
3. There is organisational structure to develop the CI.
4. There is a CI leader.
5. Improvement projects are occasional (a system exists to 'look for problems').
6. Employees participate in improvement projects.
7. The solutions posed to the problems are middle to long term.
8. There is a structured system to manage improvement suggestions.
9. The solutions implemented have an impact on the strategy.
10. There is a recognition system (economic, promotion, etc.).
LEVEL 2
1. There is an organisational structure for the CI and a CI leader (partial of full availability).
2. Improvement teams are created to solve specific problems.
3. There is a CI process, and it is applied.
4. There are process improvement teams.
5. Teams apply improvement tools (troubleshooting).
6. Employees have been trained in basic tools for improvement and problem solving (5S, seven quality basics tools).
7. 50% or more employees participate in CI activities.
8. There is an idea management system (suggestions box, improvement meetings, improvement suggestions analysis, etc.).
9. There is a recognition system (economic, promotion, etc.).
10. CI activities are part of the daily (operational) activities of the organisation.
LEVEL 3
1. The strategic objectives are deployed at all levels of the organisation in a formal way.
2. The improvement projects are monitored against the strategic objectives deployed.
3. There is a CI leader with total availability to exercise his role.
4. There is a coordinated system to launch new improvement teams.
5. 75% or more employees participate in CI activities.

MATURITY LEVEL
6. CI activities are an important part of the daily (operational) activities of the organisation.
7. The CI system includes interdepartmental (interfunctional) improvement teams.
8. The CI teams include personnel from different organisational levels (employees, middle managers, department managers, etc.).
9. CI teams include personnel from other organisations (client, supplier).
10. Improvement projects have a focus on solving interdepartmental problems (global vision, not departmental).
LEVEL 4
1. The strategic objectives are deployed at all levels of the organisation in a formal way.
2. The improvement projects are monitored against the strategic objectives deployed.
3. The CI leader (full time) and the CI organisational structure are seated and exercise their role (search for improvement opportunities).
4. Improvement teams are autonomous and can define problems, establish specific objectives and plan their actions.
5. More than 75% of employees participate in CI activities.
6. CI activities are part of the main activities of the organisation.
7. The improvement teams are interdepartmental and include personnel from different organisational levels (employees, middle managers, department managers).
8. CI teams include personnel from other organisations (client, supplier, etc.).
9. Improvement projects have a focus on solving interdepartmental problems (global vision, not departmental).
10. The level of experimentation (test of different solutions) is high.
LEVEL 5
1. The strategic objectives are deployed at all levels of the organisation in a formal manner, and the projects respond to these objectives.
2. The CI leader and the members of the promoter team have a total availability to exercise their role (search for opportunities, coordination of improvement projects, etc.).
3. The promoter team tracks the improvement projects, ensuring that they respond to the strategic objectives.
4. There is a system to capture learning and share new knowledge (the CI leader and the promoter team play this role).
5. The promoter team and improvement teams constantly identify the need for learning at all levels of the organisation.
6. Improvement teams are autonomous and can define problems, establish specific objectives and plan their actions.
7. More than 75% of employees participate in CI activities.
8. The improvement teams are interdepartmental and include personnel from different organisational levels and include other organisations (Customer supplier).
9. Improvement teams apply a system to find and solve problems.
10. Experimentation is widespread and autonomous but controlled by management (through an organisational structure).

Depending on the number of 'true' answers obtained, the maturity level of the CI can be defined.

$$CI\ maturity\ level = \frac{\boldsymbol{number\ of\ 'true'\ answers}}{10} \quad (1)$$

Table A2 – Number of 'True' Answers

LEVEL	Number of 'true' answers
LEVEL 1	0 – 10
LEVEL 2	11 – 20
LEVEL 3	21 – 30
LEVEL 4	31 – 40
LEVEL 5	41 - 50

The routines evolution questionnaire has been developed based on constituent behaviours as defined by Bessant, Caffyn and Gallagher (2001). Each statement must be evaluated on a Likert-type scale ranging from 1 (disagree) to 4 (strongly agree).

Table A3 – Routines Evolution Questionnaire (2)

Routines evolution		Likert
R1 Understanding CI' - the ability to articulate the that basic values of CI		
1. People at all levels demonstrate a shared belief in the value of small steps, and everyone can contribute by being actively involved in making and recognising incremental improvements.	Are people involved in developing and implementing small improvements in their jobs?	-
2. When something goes wrong, the natural reaction of people at all levels is to look for reasons why rather than to blame individual(s).	Faced with problems, are solutions sought before the guilty?	-
3. People make use of some formal problem finding and solving cycle.	Are problems and solutions discussed in appropriate discussion forums using troubleshooting tools?	-
R2 'Getting the CI habit' - the ability to generate sustained involvement in CI		
4. People use appropriate tools and techniques to support CI.	Are appropriate techniques and tools used to solve the problems?	-
5. People use measurement to shape the improvement process.	Are measured the improvements made?	-
6. People (as individuals and/or groups) initiate and carry through CI activities – they participate in the process.	Do workers propose improvements? Do workers participate in CI activities?	-
7. Closing the loop – ideas are responded to in a clearly defined and timely fashion and are either implemented or otherwise dealt with.	Are management members (leaders) adequately responding to improvement suggestions?	-

Routines evolution		Likert
R3 'Leading the way' - the ability to lead, direct and other support the creation and sustaining of CI behaviours		
8. Managers support the CI process through allocation of time, money, space and other resources.	Are continuous improvement activities supported with necessary resources (time, economical, education, training, etc.)?	-
9. Managers recognise in formal (but not necessarily financial) ways the contribution of employees to CI.	Is the contribution of workers making improvements or suggestions for improvement formally recognised?	-
10. Managers lead by example, becoming actively involved in the design and implementation of CI.	Do the leaders lead the improvement activities to set an example?	-
11. Managers support experiment by not punishing mistakes but by encouraging learning from them.	Do managers (leaders) encourage employees to experiment (without penalising the error) to find the right solutions?	-
R4 'Focusing CI' - the ability to link CI activities to the strategic goals of the company		
12. Individuals and groups use the organisation's strategic goals and objectives to focus and prioritise improvements, and everyone understands (i.e., is able to explain) what the company's or department's strategy, goals and objectives are.	Does the company present the strategic objectives in a clear and general way? Do workers know what the company's objectives are?	-
13. Individuals and groups (e.g., departments, CI teams) assess their proposed changes (before embarking on initial investigation and before implementing a solution) against departmental or company objectives to ensure they are consistent with them.	Do workers evaluate/contrast their suggestions for improvement against the company's objectives?	-
14. Individuals and groups monitor/measure the results of their improvement activity and the impact it has on strategic or departmental objectives.	Are the results of the improvements made measured and their impact on the company's objectives contrasted?	-
15. CI activities are an integral part of the individual or groups work, not a parallel activity.	Are MC activities part of daily work? Or are they an extra job?	-
R5 'Shared problem-solving' - the ability to move CI activity across organisational boundaries		
16. People co-operate across internal divisions (e.g., cross-functional groups) in CI, as well as working in their own areas.	Are there multidisciplinary teams to execute the CI activities?	-
17. People understand and share a holistic view (process understanding and ownership).	Do workers have a global vision of the CI system?	-
18. People are oriented towards internal and external customers in their CI activity.	Is the CI system customer oriented (internal/external)?	-
19. Specific CI projects with outside agencies (customers, suppliers, etc.) are taking place.	Are CI projects developed with agents outside the organisation?	-
20. Relevant CI activities involve representatives from different organisational levels.	Do CI activities involve workers of different organisational levels?	-

Routines evolution		Likert
R6 'Aligning CI' - the ability to create consistency between CI values and behaviour and the organisational context (structures, procedures, etc.)		
21. Ongoing assessment ensures that the organisation's structure and infrastructure and the CI system consistently support and reinforce each other.	Is there an organisational structure that supports the CI activities?	-
22. The individual/group responsible for designing the CI system design it to fit within the current structure and infrastructure.	Does the organisational structure of the CI conform to the current organisational structure?	-
23. Individuals with responsibility for particular company processes/systems hold ongoing reviews to assess whether these processes/systems and the CI system remain compatible.	Are the company's processes compatible with the CI system?	-
24. People with responsibility for the CI system ensure that when a major organisational change is planned, its potential impact on the CI system is assessed and adjustments are made as necessary.	When there are changes in the organisation, is it analysed if the changes affect the CI system/structure?	-
R7 'Continuous improvement of continuous improvement' - the ability to strategically manage the development of CI		
25. The CI system is continually monitored and developed; a designated individual or group monitors the CI system and measures the incidence (i.e., frequency and location) of CI activity and the results of CI activity.	Is the continuous monitoring of the CI system and the results of the CI activities?	-
26. There is a cyclical planning process whereby (a) the CI system is regularly reviewed and, if necessary, amended (single-loop learning).	Is the CI system reviewed regularly (annual frequency) and modified if necessary?	-
27. There is periodic review of the CI system in relation to the organisation as a whole, which may lead to a major regeneration (double-loop learning).	Is the CI system reviewed and its relationship with the organisation (analysing if there are changes in the organisation itself), and is it adapted if necessary?	-
28. Senior management makes available sufficient resources (time, money, personnel) to support the ongoing development of the CI system.	Does the management support and give the necessary resources (time, people, money) to develop the CI system?	-
R8 'The learning organisation' - generating the ability to enable learning to take place and be captured at all levels		
29. People learn from their experiences, both positive and negative.	Do workers learn from their experiences (positive or negative)?	-
30. Individuals seek out opportunities for learning/personal development (e.g., actively all levels experiment, set their own learning objectives).	Do people look for opportunities to learn and develop personally?	-
31. Individuals and groups share (make available) their learning from all work experiences.	Do workers share their knowledge with others naturally?	-

Routines evolution		Likert
R8 'The learning organisation' - generating the ability to enable learning to take place and be captured at all levels		
32. The organisation articulates and consolidates (captures and shares) the learning of individuals and groups.	Are there internal training plans to socialise knowledge?	-
33. Managers accept and, where necessary, act on all the learning that takes place.	Does management accept the training developed? Does management participate in the trainings?	-
34. People and teams ensure that their learning is captured by making use of the mechanisms provided for doing so.	Do workers ensure that their knowledge is documented?	-
35. Designated individual(s) use organisational mechanisms to deploy the learning captured across the organisation.	Are there mechanisms in the organisation to share knowledge across the organisation?	-

Are the 'Illnesses' of Traditional Likert Scales Treatable?

DOI: 10.12776/QIP.V24I2.1439

Zsuzsanna E. Tóth, Gábor Árva, Rita V. Dénes

Received: 2020-03-31 Accepted: 2020-06-11 Published: 2020-07-31

ABSTRACT

Purpose: The main aim of this paper is to introduce the development and the application of a fuzzy rating scale in measuring customer satisfaction which are to be demonstrated through a healthcare example in order to illustrate how the proposed methodology is able to enhance the reliability of traditional Likert scale-based evaluations.

Methodology/Approach: The proposed methodology is built on fuzzy sets the membership function of which is composed of two sigmoid functions by applying Dombi's conjunction operator. The possible 'values' of the linguistic variable expressing customer satisfaction are to be expressed by these functions which can also be linked to the level of organizational performance allowing the illustration of the mainly nonlinear relationship between the provided and perceived service performance.

Findings: The application of the proposed fuzzy rating scale confirms its ability to reflect the unambiguity of human ratings as well as the context-dependency of ratings resulting in a more precise representation of human judgements.

Research Limitation/implication: The presented methodology may be viewed as a viable approach in any kind of service quality evaluations where Likert-type scales are traditionally applied to handle its weaknesses.

Originality/Value of paper: The proposed methodology is not only able to reflect the satisfaction of customers and the organizational performance simultaneously, but the expectations of customers related to the desired level of performance can also be incorporated into the establishment of the scale yielding to more reliably supported managerial decisions.

Category: Research paper

Keywords: fuzzy number; Likert scale; healthcare; service quality; patient satisfaction

1 INTRODUCTION

Measurement is a key management activity in the service industry where subjectivity is an unassignable part of the evaluation process. In order to measure the level of service quality and to ascertain reliably whether the needs and requirements of the customers are met, it is fundamental that their expectations and perceptions are properly measured and correctly understood (Lupo, 2013). In order to serve that purpose, there is a need to design suitable and reliable evaluation methodologies and to identify appropriate measure units to highlight the achieved service performance level (Lupo, 2013; Battisti, Nicolini and Salini, 2010). However, the intangibility, inseparability and heterogeneity embedded in services make it difficult to measure and evaluate service quality (SQ) in a way upon which sole managerial decision could be made.

The vast majority of surveys developed to gain information related to customer perceptions and satisfaction associated with SQ utilizes primarily Likert scales, however, these scales are often criticized on the grounds that they cannot depict human judgement reliably so that the purpose of this paper is to provide more reliable methodological solutions. Likert scales generally use crisp values to present the feelings and subjective perceptions related to specific SQ dimensions. As the evaluation process obviously encompasses intangible and subjective information, crisp values are inadequate to reflect the real ratings of customers, which means that non-negligible difficulties arise when the differences and uncertainties in human semantic expressions are to be understood (Hu, Lee and Yen, 2010). The relevant literature calls attention to the application of the fuzzy rating scale as a viable alternative approach to Likert scale-based measurement (see e.g. Hesketh, Pryor, Gleitzman and Hesketh, 1988; de la Rosa de Sáa, Gil, González-Rodríguez, López, and Lubiano, 2015) since if the evaluation is the outcome of the evaluator's subjective interpretation of linguistic variables, it must be conducted in an uncertain, fuzzy environment. Therefore, this study aims to introduce a conceptual model to assess perceived SQ utilizing the fuzzy concept and also presents the advantages arising from the more effective evaluation of patient feedbacks associated with healthcare SQ. Additionally, we are also to compare the results brought by traditional Likert scales and the proposed approach in order to highlight the methodological benefits that can contribute to more reliable managerial decisions.

The paper is structured as follows. Section 2 discusses the features of traditional Likert scales. Section 3 includes the methodological background. Section 4 presents a case study demonstrating the application of the proposed methodology as well as its main benefits. Section 5 discusses conclusions and gives an overview of future research directions.

2 LITERATURE REVIEW

Measuring service quality, the satisfaction of stakeholders and the relative importance of the various service features are mainly realized through the application of Likert scales by utilizing either the original SERVQUAL methodology (Parasuraman, Zeithaml and Berry, 1988) or its numerous modifications specifically developed for the various service industries (in healthcare context see e.g. Vandamme and Leunis, 1993; Ramsaran-Fowdar, 2008; Al-Borie and Damanhouri, 2013; Garrard and Narayan, 2013). The widespread application and high popularity of Likert scales are primarily owing to their ease of use and the simple interpration of results. However, the subject of service quality is frequently burdened by fuzzy terms such as attitude, perception, satisfaction etc. as respondents may subjectively fill out the questionnaire based on their unique experience and perceptions of the given service performance. This subjective assessment is intrinsically imprecise and ambiguous, possibly inhibiting service quality (Liou and Chen, 2006).

The application of the traditional Likert scale raises diversified issues. Rating service quality features in a questionnaire is a complex task as customers make multiple decisions under uncertainty. First, the number of 'values' to choose from is usually small (Gil and González-Rodríguez, 2012) which means that the variability, diversity and subjectivity associated with an accurate rating are usually lost. Second, when Likert-type data are analysed for statistical purposes, the number of techniques that can be applied are quite limited since statistical conclusions addressed to ordinal data could only be reliable and relevant information could be lost (Lubiano et al., 2016). Third, the weighting of aspects is a key issue as the different SQ features are naturally not equally important for the respondents. As a consequence, an additional concern arises as the respondents' attitudes towards the rated features are not homogeneous as time goes on (see e.g. Tóth, Surman and Árva, 2017). Moreover, if raters' preferences are heterogeneous, it does matter how and to what extent it influences the overall evaluation of a given SQ feature. 'Average' scores are supposed to hide the real situation, namely, the performance of the given feature (e.g. Kuzmanovic, Savic, Popovic and Martic, 2013). As a conclusion, crisp values are inadequate to present the evaluations of customers properly due to intangible and subjective information embedded in the evaluation process.

The reliability could be enhanced by increasing the number of responses (Lozano, García-Cueto and Muñiz, 2008; de la Rosa de Sáa, Gil, González-Rodríguez, López, and Lubiano, 2015), however, it cannot be achieved by using a natural language (Sowa, 2013). In order to embed human perceptions more precisely, fuzzy set theory is increasingly applied in these situations. Fuzzy set theory is believed to be able to improve successfully the reliability of service process measurements and evaluations (Li, 2013; Lupo, 2016) by handling uncertainty in the case of weakly defined measurements (Benoit, 2013) and can be applied to overcome the limitations of standard scales by modelling the imprecision of human rating evaluations (Calcagní and Lombardi, 2014) as well.

This fuzzy assessment of service quality attributes is much closer to human thinking than the methods based on crisp numbers (Lin and Wu, 2008) yielding to more reliably supported managerial decisions as well.

3 METHODOLOGY

The purpose of the introduced methodology is to overcome the weaknesses of traditional Likert scale-based evaluations as well as to assess the service performance and customer satisfaction simultaneously. Tóth, Jónás and Dénes (2018) have already proved that in a fuzzy environment, one may be able to deal with the vagueness arising either from the uncertainty or subjectivity of the respondent or from the variation of the performance over time. The proposed fuzzy number-based evaluation may be further enhanced by linking the satisfaction levels of patients to the experienced performance level.

With the purpose of evaluating SQ and customer satisfaction simultaneously, the possible values 'completely dissatisfied', 'dissatisfied', 'rather dissatisfied', 'rather satisfied', 'satisfied', 'completely satisfied' of the linguistic variable expressing patient satisfaction are considered to be fuzzy sets and they are assigned to the experienced performance level. The membership function, which expresses the degree to which a given level of performance belongs to a certain fuzzy set, is composed as a Dombi-intersection of two sigmoid-shaped functions.

Definition (1), the sigmoid function $\sigma_a^{(\lambda)}(x)$ *w*ith parameter a and λ, is given by:

$$\sigma_a^{(\lambda)}(x) = \frac{1}{1+e^{-\lambda(x-a)}}, \tag{1}$$

where $x, a, \lambda \in \mathbb{R}$ and λ is nonzero (Dombi, 2009).

The main properties of the sigmoid function are thoroughly discussed in Jónás, Tóth and Árva (2018). Let $\sigma_{a_l}^{(\lambda_l)}(x)$ be an increasing sigmoid function given by the parameters a_l, λ_l and $\sigma_{a_r}^{(\lambda_r)}(x)$ having the parameters a_r and λ_r be a decreasing sigmoid function. Conjoining these two sigmoid functions by Dombi's intersection operator denoted by $*_{(D)}$ (Dombi, 2009), one may get the following membership function:

$$\sigma_{a_l}^{(\lambda_l)}(x) *_{(D)} \sigma_{a_r}^{(\lambda_r)}(x) = \frac{1}{1+e^{-\lambda_l(x-a_l)}+e^{-\lambda_r(x-a_r)}}. \tag{2}$$

In order to determine the parameters a and λ of the sigmoid function given in Definition (1), one should choose two pints of the function curve. If one seeks to be consistent with human thinking and traditional evaluations, the parameters a and λ should be determined based on the parameter triplet l, m, r. That is, the rater is asked to express his or her opinion on the experienced level of performance by the parameter triplet l, m, r, where m is the value which seems most likely to express the assessor's judgement in the examined dimension, and

in addition to that l and r denote the worst and the best possible value which the assessor would give when evaluating the given statement, respectively.

It should be highlighted here that in the case of a traditional Likert scale-based evaluations only the value m can be given, that is, the respondent is forced to reduce his or her opinion to a single value which is most likely to express his or her judgement. On a fuzzy Likert scale, the additional two parameters l and r should be chosen so that they are proportional to the perceived variability of the performance or to the uncertainty or subjectivity of the respondent. In other words, the higher the uncertainty in the evaluation or the more variation of performance is experienced, the larger should be the difference between l and r. Since the sigmoid function neither takes the values of 0 or 1 (these are only the limits of this function), one should choose a small positive number denoted by ε, for example, $\varepsilon = 0.001$. After that, based on the values l and m, the parameters a_l and λ_l can be determined as:

$$a_l = \frac{l+m}{2}, \tag{3}$$

$$\lambda_l = \frac{2}{m-l} \ln\left(\frac{1-\varepsilon}{\varepsilon}\right). \tag{4}$$

The parameters a_r and λ_r are determined based on the values m and r as follows:

$$a_r = \frac{r+m}{2}, \tag{5}$$

$$\lambda_r = \frac{2}{m-r} \ln\left(\frac{1-\varepsilon}{\varepsilon}\right). \tag{6}$$

Having identified the parameters of the increasing sigmoid function $\sigma_{a_l}^{(\lambda_l)}(x)$ according to Eq. (3) – (4) and the parameters of the decreasing sigmoid function $\sigma_{a_r}^{(\lambda_r)}(x)$ based on Eq. (5) – (6), the two sigmoid functions are conjunct by applying Dombi's intersection operator in (2). Setting the parameters as given in (3) – (6) results in a function value of ε at the points l and r and $1-\varepsilon$ at the point m. Henceforth, the variable x represents the performance level, whereas the function values of $\sigma_{a_l}^{(\lambda_l)}(x)$ or $\sigma_{a_r}^{(\lambda_r)}(x)$ express the truth of the statement that a certain level of performance belongs to a given fuzzy set which depicts the patients' verbal judgement on the performance, that is, their satisfaction with the experienced performance level.

From this point, the Dombi's intersection of two sigmid-shaped membership functions is utilized as a membership function of a fuzzy set expressing the verbal judgement of patients with a given level of performance experienced at a healthcare institution.

With the purpose of establishing a fuzzy Likert scale, the following considerations are taken into account:

- to each linguistic variable, a sigmoid-shaped membership function is assigned;
- to the linguistic variable 'completely dissatisfied', a sigmoid function with $l = -\infty$, and similarly, to the linguistic variable 'completely satisfied' a membership function with $\mathrm{r} = \infty$ is assigned;
- in each point of the scale, two linguistic variables are defined with a membership value being greater than ε;
- in our approach, the value m of a certain membership function is assumed to be equal to the value of r of the previous membership function and to the value of l of the following membership function.

In their work, Hu, Lee and Yen (2010) utilized a similar rating scale for evaluating hospital out-patient services by a fuzzy linguistic SERVQUAL model. In contrast to their work, the current research does not specify the parameters of the linguistic terms in advance. Instead, a group of patients is asked to give the parameter m for each linguistic term in each SQ dimension to be evaluated. That is, in each dimension, patients should choose a value which expresses the performance of the healthcare institution if their satisfaction is characterized by a particular linguistic label. Establishing the rating scale based on surveying the patients is beneficial since it allows not only the mapping of the relationship between the healthcare institutions' performance and the patients' satisfaction level but also the involvement of customers' expectations related to the performance level into the formulation of the rating scale. During the 'calibration' phase of the rating scale, patients are asked to answer the following questions in each of the SQ dimensions to be evaluated:

- *What is the performance level under which you would be 'completely dissatisfied' with the performance of the healthcare institution?*
- *What is the performance level which seems most likely to express the performance if you are 'dissatisfied', 'rather dissatisfied', 'rather satisfied' and 'satisfied' with the performance of the healthcare institution, respectively?*
- *What is the performance level above which you would be 'completely satisfied' with the performance of the healthcare institution?*

By answering the above listed questions, the values of m can be determined for each membership function. In our approach, the value of m of a certain membership function is assumed to be equal to the value of r of the foregoing membership function and to the value of l of the forthcoming membership function, as a result of which, for each linguistic term, the corresponding membership function can be calculated unambiguously by utilizing the equations (3) – (6) and then, the two sigmoid functions can be conjoined by Dombi's

intersection operator in (2). As such, for each of the possible values of the linguistic variable expressing patient satisfaction, the membership function of the corresponding fuzzy set can be determined.

4 HEALTHCARE EXAMPLE DEMONSTRATING THE APPLICATION OF THE METHODOLOGY

In recent years, healthcare has become one of the extremely complex and consistently growing industries in the world (Bertolini, Bevilacqua, Ciarapica and Giacchetta, 2011). Due to the complexity of healthcare services and to the great number of stakeholders, service quality (SQ) in the healthcare sector is fairly variable. Patient perceptions have attracted considerable attention and have been increasingly emphasized as an important element of healthcare SQ evaluation for several reasons (Iversen, Holmboe and Bjertnæs, 2012; Carlucci, Renna and Schiuma, 2013). First, a high level of SQ has a relationship with patient satisfaction, willingness to re-use the services (e.g. Arab, Tabatabaei, Rashidian, Rahimi and Zarei, 2012). Second, patient feedbacks are an integral part of any accreditation and evaluation programs. Therefore, patient feedbacks are considered as an essential element in planning and policy making that enhance the more effective management of the services provided by healthcare institutions (Carlucci, Renna and Schiuma, 2013) by providing the opportunity to organizational learning and development and identifying the shortcomings in the service provision.

Based on the literature, a great variety of approaches is used for data collection to increase the quality awareness of the healthcare system. As a mean the most important stakeholders', namely, patients' experience is measured and evaluated to improve healthcare quality on different levels of the system. The most popular method is to collect patient satisfaction and perception data to assess the quality of healthcare services (e.g. Alhashem, Alquraini and Chowdhury, 2011; Naidu, 2009) which are primarily based on the application of Likert scales. At the same time, the results of fuzzy set theory have been increasingly utilized in healthcare service quality evaluations as well (e.g. Woldegebriel, Kitaw and Rafele, 2015; Akdag, Kalaycı, Karagöz, Zülfikar and Giz, 2014; Singh and Prasher, 2017; Tsai, Chang and Lin, 2010; Büyüközkan, Çifçi and Güleryüz, 2011; Hu, Lee and Yen, 2010; Lupo, 2016; Behdioğlu, Acar and Burhan, 2017).

In the light of the state of the art (e.g. Naidu, 2009; Kessler and Mylod, 2011; Haque, Sarwar, Yasmin and Anwar, 2012; Yesilada and Director, 2010; Roberge, Tremblay, Turgeon and Berbiche, 2013; Grondahl, Wilde-Larsson, Karlsson and Hall-Lord, 2013), our study considers 7 major aspects of patient satisfaction, namely, processes (main processes and supporting processes of caring); outcomes (reputation, image, efficiency, effectiveness); care characteristics (personalized attention, availability of healthcare workers, timing and organizing of caring, catering); accessibility (availability, accessibility for disabled people with reduced mobility); communication (communication and interaction with

healthcare workers, clarity and timing of communication and information); responsiveness (attitude and empathy of healthcare workers, respect and courtesy); tangibles and environment (physical facilities, equipment, appearance of contact personnel, modernity, hygiene, safety).

Since Lubiano et al. (2016) point out that in the case of fuzzy Likert scales, detailed instructions for the respondents are usually needed on how to answer the questionnaire, a pilot study has been launched among university students. In order to gain experience with the application of the proposed methodology, 219 engineering students were asked to assess specific quality-related attributes based upon their last experience in a healthcare institution. The experience gained during this pilot study could serve as a basis for the establishment of a final survey which is planned to be launched among out-patients of a specific Hungarian healthcare institution specialized in rehabilitation. These students were asked to assess the performance level under which he or she would be 'completely dissatisfied' with the healthcare institution's performance and similarly, to address the performance level above which he or she would be 'completely satisfied 'with the experienced level of performance for all the aspects listed above. Besides that, for the 'scale points' located in the middle of the scale, respondents were asked to give a performance level which is most likely to express the performance of the healthcare institution if he or she is 'dissatisfied', 'rather dissatisfied', 'rather satisfied' or 'satisfied' with the care received, respectively. Since these questions should be answered in each of the investigated service quality dimensions, the different expectations associated with the healthcare institution's performance related to the studied dimensions could be taken into account as well.

Figure 1 depicts the healthcare institutions' performance (on the x-axis) and the membership functions of the linguistic terms 'completely dissatisfied', dissatisfied', 'rather dissatisfied', 'rather satisfied', 'satisfied' and 'completely satisfied' assigned to the performance level in each investigated service quality dimensions.

Figure 1 suggests that the linguistic terms representing the patients' satisfaction do depend on which quality-related aspect is to be evaluated. While in the case of outcomes, the performance level should be higher than 55.1 to avoid 'completely dissatisfied' customers, in the case of accessibility, the performance level of 36.2 is already enough if healthcare institutions seek to avoid 'completely dissatisfied' patients. The same conclusions may be drawn if one examines the performance level above which patients are 'completely satisfied': in the case of outcomes, institutions should achieve as high performance level as 99, while in the case of communication, a performance level of 90.3 is already considered as one 'completely satisfying' the patients. One may be tempted to conclude that the higher the importance of a particular service dimension is, the higher the expectations related to the performance in this dimension are.

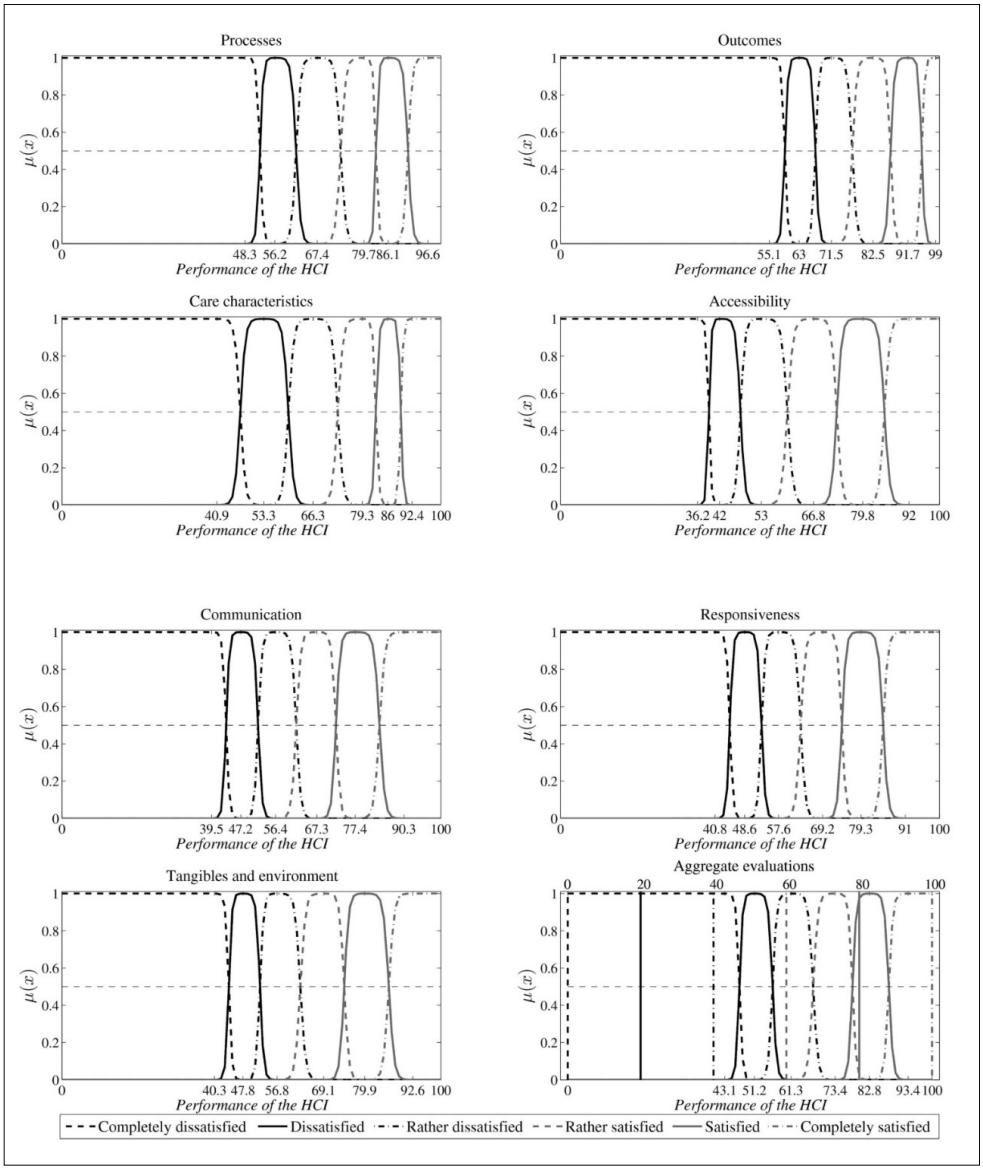

Figure 1 – Membership Functions of the Linguistic Terms Assigned to the Healthacre Institutions Performance Level in the Investigated Dimensions as Well as That of the Agregate Evaluation

Figure 1 also demonstrates that up to the one-third, in some cases almost up to the midpoint of the performance scale, patients are usually 'completely dissatisfied' with the provided performance level of the healthcare institution. A performance of 33 (at around the one-third of the performance scale) is considered to be 'completely dissatisfying' in all examined dimensions, while the performance of 50, at the half of the performance scale, results either in 'dissatisfied' or 'rather dissatisfied' patients, except for the attribute titled as

outcomes. In this special case the performance level is thought to be 'completely dissatisfying'. On the contrary, on a traditional Likert scale, a performance level around the one-third of the scale is already considered to be assessed by the second or the third linguistic term representing the patients' satisfaction. As a result of different expectations related to the performance level, the distance between the consecutive 'scale points' is not constant and also depends on which quality attribute is studied. Investigating the evaluations given in the quality attribute titled as care characteristics, based on Figure 1 one may also conclude that the 'distance' between the scale points which seems most likely to express the performance if the patients are 'rather dissatisfied' and 'rather satisfied' is 13 units, while the consecutive scale point is only 7.2 units away.

The different expectations in the examined quality attributes lead to the fact that the same performance is judged differently in various dimensions. While in the case of processes, a performance level of 56.2 is judged as 'dissatisfying', the same performance is more likely to be evaluated as 'rather dissatisfying' if the responsiveness of healthcare institutions is examined.

Jónás, Tóth and Árva (2018) suggest a methodology which is also based on Dombi's Pliant Inequality Model to aggregate sigmoid-shaped membership functions. The subplot in the bottom right-hand corner of Figure 1 shows the aggregate rating scale depicting the overall judgement related to the healthcare institutions' performance. The upper x-axis of this subplot belongs to a scale on which the consecutive linguistic terms 'completely dissatisfied (0)', 'dissatisfied (20)', 'rather dissatisfied (40)', 'rather satisfied (60)', 'satisfied (80)', 'completely satisfied (100)' are distributed equally apart from each other, exactly as often practiced when carrying out Likert scale-based evlautions, whereas the lower x-axis denotes their corresponding fuzzy counterparts.

Based on Figure 1, it can be concluded that the assumption of equally-distributed scale points in all dimensions usually does not match the patients' expectations. That is, traditional Likert scales not only lack the ability to deal with uncertainty, vagueness, imprecision or take the variation of performance into account, but the assumption of the same 'distance' between the consecutive scale points is not consistent with human judgement resulting in the fact that either the performance or the satisfaction is improperly interpreted and evaluated. The inadequate or biased judgement may lead to weakly supported managerial decisions and as a consequence, may set organizational competitiveness back as well. Fuzzy Likert scales, on the contrary, are able to depict patient evaluations in a manner which is consistent with human thinking due to the fact that this methodology is not only able to take into account the subjectivity, imprecision and variation of performance but could also model the often nonlinear relationship between organizational performance and customer satisfaction. Unlike prior studies in the literature (e.g. Hu, Lee and Yen, 2010), the proposed scale is established based on patient expectations related to the desired level of performance. This feature of the proposed rating scale further enhances its applicability owing to the fact that not only the perceived performance level and the satisfaction with the

experienced performance but also the expectations concerning the desired level of performance can be incorporated into the development of the proposed rating scale. The ability of measuring these three aspects of service quality simultaneously can be considered as the main benefit of the suggested methodology.

The proposed methodology has two notable limitations which should be highlighted here. One of them is the need to answer six questions during the 'calibration' phase of the scale in each of the quality aspects to be evaluated. The other major constraint of the suggested methodology arises from the fact that detailed explanations and instructions are needed on how to use the scale. One reason for asking university students to test the scale was the fact that due to their prior methodological studies they have been familiar with fuzzy logic-based issues. The majority of students does not experience any trouble when filling out the survey. Taking these considerations into account, after some minor modifications, the proposed methodology might be applied to gain information from out-patients of healthcare institutions but that has to be prepared carefully.

5 CONCLUSION AND FUTURE RESEARCH DIRECTIONS

In this paper, a fuzzy rating scale-based methodology is developed to overcome the weaknesses of Likert scale-based evaluations with the aim of assessing specific dimensions of healthcare service quality. The proposed methodology is considered to deal with the inherent uncertainty, subjectivity and vagueness characterizing stakeholders ratings by expressing their own judgements associated with quality attributes. Our results are in line with that of Yeh and Kuo (2003), Calcagní and Lombardi (2014) or Liou and Chen (2006) who pointed out that fuzzy evaluation of quality attributes is much closer to human judgement than traditional, 'crispy' evaluation. Tóth, Jónás and Dénes (2018) have already presented a flexible fuzzy number-based evaluation of institutional performance, their methodology, however, is not able to take into account the satisfaction of patients with the perceived level of performance. The evaluation framework proposed by Hu, Lee and Yen (2010) deals with the satisfaction of patients, yet, the linguistic terms representing customer satisfaction are determined by the researcher in advance.

In this study, the membership functions of the fuzzy sets expressing the judgement of patients on a given level of performance are determined based on patient expectations leading to a more precise and reliable depiction of patient judgements. Examining the established rating scales, one may conclude that the 'distance' between the consecutive scale point is not constant and there is no crisp boundary among the 'scale values' as the traditional Likert scale assumes. In addition to that, the proposed methodology is able to handle the fact that patient expectations do depend on which quality attribute is to be evaluated. As a result, the discussed methodology is able to map the often strongly nonlinear relationship between quality attributes and customer satisfaction. By offering the

way to evaluate the organizational performance and patient satisfaction simultaneously, the proposed fuzzy evaluation environment aims at supporting healthcare decision makers in order to facilitate effective and efficient strategies related to quality improvements by identifying which quality dimensions require more consideration.

Lubiano et al. (2016) argue that respondents need a special training before using fuzzy rating scales which was the reason for asking university students to test the proposed approach. Based on the encouraging results and the fact that the majority of the students has not experienced any difficulty, after some minor modifications a similar questionnaire is to be launched to collect feedbacks from out-patients of a given Hungarian healthcare institution.

In a further research, the fuzzy AHP process may be utilized to determine the importance of various service quality dimensions based on patients' viewpoints. Based on the importance 'scores', one may be able to weight the statements in a manner that expresses clearly how important a particular service dimension is (e.g. Büyüközkan, Çifçi and Güleryüz, 2011). Another possible future research direction is to investigate how the different levels of healthcare (GP, outpatient care, hospitals, rehabilitation institutions) or even the patients' personal characteristics influence the patient expectations related to healthcare quality issues. Since the proposed fuzzy rating scale is built on patient expectations with the provided performance level, one may assume, that different expectations at the distinct levels of provision may lead to different rating scales.

What is more, the applicability of the proposed scale is not limited merely to the assessment of patient satisfaction. Several other fields in which the evaluation is subjective in nature may be investigated by the proposed methodology. Besides examining stakeholder expectations in healthcare context (e.g. policy maker or employee expectations), it may also support managerial decisions in a couple of other areas where evaluation of service quality is at the forefront of organizational excellence. Based on the results demonstrated through the case study, the fuzzy rating scale introduced in this paper offers a viable alternative technique for these evaluation goals.

REFERENCES

Akdag, H., Kalaycı, T., Karagöz, S., Zülfikar, H. and Giz, D., 2014. The evaluation of hospital service quality by fuzzy MCDM. *Applied Soft Computing*, [e-journal] 23, pp.239-248. DOI: 1016/j.asoc.2014.06.033.

Al-Borie, H.M. and Damanhouri, A.M.S., 2013. Patient's satisfaction of service quality in Saudi hospitals: A SERVQUAL analysis. *International Journal of Health Care Quality Assurance*, [e-journal] 26(1), pp.20-30. DOI: 10.1108/09526861311288613.

Alhashem, M.A., Alquraini, H. and Chowdhury, R.I., 2011. Factors influencing patient satisfaction in primary healthcare clinics in Kuwait. *International Journal of Health Care Quality Assurance*, [e-journal] 24(3), pp.249-262. DOI: 10.1108/09526861111116688.

Arab, M., Tabatabaei, M., Rashidian, A., Rahimi, A. and Zarei, E., 2012. The effect of service quality on patient loyalty: a study of private hospitals in Tehran, Iran. *Iranian Journal of Public Health*, 41(9), pp.71-77.

Battisti, F., Nicolini, G. and Salini, S., 2010. The Rasch model in customer satisfaction survey data. *Quality Technology & Quantitative Management*, [e-journal] 7(1), pp.15-34. DOI: 10.1080/16843703.2010.11673216.

Behdioğlu, S., Acar, E. and Burhan, H.A., 2017. Evaluating service quality by fuzzy SERVQUAL: a case study in a physiotherapy and rehabilitation hospital. *Total Quality Management & Business Excellence*, [e-journal] 30(3/4), pp.1-19. DOI: 10.1080/14783363.2017.1302796.

Benoit, E., 2013. Expression of uncertainty in fuzzy scales based measurements. *Measurement*, [e-journal] 46(9), pp.3778-3782. DOI: 10.1016/j.measurement.2013.04.006.

Bertolini, M., Bevilacqua, M., Ciarapica, F.E. and Giacchetta, G., 2011. Business process re□engineering in healthcare management: a case study. *Business Process Management Journal*, [e-journal] 17(1), pp.42-66. DOI: 10.1108/14637151111105571.

Büyüközkan, G., Çifçi, G. and Güleryüz, S., 2011. Strategic analysis of healthcare service quality using fuzzy AHP methodology. *Expert Systems with Applications,* [e-journal] 38(8), pp.9407-9424. DOI: 10.1016/j.eswa.2011.01.103.

Calcagnì, A. and Lombardi, L., 2014. Dynamic Fuzzy Rating Tracker (DYFRAT): a novel methodology for modeling real-time dynamic cognitive processes in rating scales. *Applied Soft Computing*, [e-journal] 24, pp.948-961. DOI: 10.1016/j.asoc.2014.08.049.

Carlucci, D., Renna, P. and Schiuma, G., 2013. Evaluating service quality dimensions as antecedents to outpatient satisfaction using back propagation neural network. *Health Care Management Science*, [e-journal] 16(1), pp.37-44. DOI: 10.1007/s10729-012-9211-1.

de la Rosa de Sáa, S., Gil, M.Á., González-Rodríguez, G., López, M.T. and Lubiano, M.A., 2015. Fuzzy rating scale-based questionnaires and their statistical analysis. *IEEE Transactions on Fuzzy Systems*, [e-journal] 23(1), pp.111-126. DOI: 10.1109/TFUZZ.2014.2307895.

Dombi, J., 2009. Pliant Arithmetics and Pliant Arithmetic Operations. *Acta Polytechnica Hungarica*, 6(5), pp.19-49.

Garrard, F. and Narayan, H., 2013. Assessing obstetric patient experience: a SERVQUAL questionnaire. *International Journal of Health Care Quality Assurance*, [e-journal] 26(7), pp.582-592. DOI: 10.1108/IJHCQA-08-2011-0049.

Gil, M.Á. and González-Rodríguez, G., 2012. Fuzzy vs. Likert scale in statistics. In: E. Trillas, P.P. Bonissone, L. Magdalena and J. Kacprzyk, eds. 2012. *Combining experimentation and theory*. Berlin, Heidelberg: Springer. pp.407-420. DOI: 10.1007/978-3-642-24666-1_27.

Grondahl, V.A., Wilde-Larsson, B., Karlsson, I. and Hall-Lord, M.L., 2013. Patients' experiences of care quality and satisfaction during hospital stay: a qualitative study. *European Journal for Person Centered Healthcare*, [e-journal] 1(1), pp.185-192. DOI: 10.5750/ejpch.v1i1.650.

Haque, A., Sarwar, A.A.M., Yasmin, F. and Anwar, A., 2012. The impact of customer perceived service quality on customer satisfaction for private health centre in Malaysia: a structural equation modeling approach. *Information Management and Business Review*, 4(5), pp.257-267.

Hesketh, B., Pryor, R., Gleitzman, M. and Hesketh, T., 1988. Practical applications and psychometric evaluation of a computerized fuzzy graphic rating scale. *Advances in Psychology*, [e-journal] 56, pp.425-454. DOI: 10.1016/S0166-4115(08)60493-8.

Hu, H-Y., Lee, C-H. and Yen, T-M., 2010. Service quality gaps analysis based on Fuzzy linguistic SERVQUAL with a case study in hospital out□patient services. *The TQM Journal*, [e-journal] 22(5), pp.499-515. DOI: 10.1108/17542731011072847.

Iversen, H.H., Holmboe, O. and Bjertnæs, Ø.A., 2012. The Cancer Patient Experiences Questionnaire (CPEQ): reliability and construct validity following a national survey to assess hospital cancer care from the patient perspective. *British Medical Journal*, [e-journal] 2(5), p.e001437. DOI: 10.1136/bmjopen-2012-001437.

Jónás, T., Tóth, ZS.E. and Árva, G., 2018. Applying a fuzzy questionnaire in a peer review process. *Total Quality Management & Business Excellence*, [e-journal] 29(9/10), pp.1228-1245. DOI: 10.1080/14783363.2018.1487616.

Kessler, D.P. and Mylod, D., 2011. Does patient satisfaction affect patient loyalty?. *International Journal of Health Care Quality Assurance*, [e-journal] 24(4), pp.266-273. DOI: 10.1108/09526861111125570.

Kuzmanovic, M., Savic, G., Popovic, M. and Martic, M., 2013. A new approach to evaluation of university teaching considering heterogeneity of students' preferences. *Higher Education*, [e-journal] 66(2), pp.153-171. DOI: 10.1007/s10734-012-9596-2.

Li, Q., 2013. A novel Likert scale based on fuzzy sets theory. *Expert Systems with Applications*, [e-journal] 40(5), pp.1609-1618. DOI: 10.1016/j.eswa.2012.09.015.

Lin, C.J. and Wu, W.W., 2008. A causal analytical method for group decision-making under fuzzy environment. *Expert Systems with Applications*, [e-journal] 34(1), pp.205-213. DOI: 10.1016/j.eswa.2006.08.012.

Liou, T.S. and Chen, C.W., 2006. Subjective appraisal of service quality using fuzzy linguistic assessment. *International Journal of Quality & Reliability Management*, [e-journal] 23(8), pp.928-943. DOI: 10.1108/02656710610688149.

Lozano, L.M., García-Cueto, E. and Muñiz, J., 2008. Effect of the number of response categories on the reliability and validity of rating scales. *Methodology*, [e-journal] 4(2), pp.73-79. DOI: 10.1027/1614-2241.4.2.73.

Lubiano, M.A., de Sáa, S.D.L.R., Montenegro, M., Sinova, B. and Gil, M.Á., 2016. Descriptive analysis of responses to items in questionnaires. Why not using a fuzzy rating scale?. *Information Sciences*, [e-journal] 360, pp.131-148. DOI: 10.1016/j.ins.2016.04.029.

Lupo, T., 2013. A fuzzy ServQual based method for reliable measurements of education quality in Italian higher education area. *Expert Systems with Applications*, [e-journal] 40(17), pp.7096-7110. DOI: 10.1016/j.eswa.2013.06.045.

Lupo, T., 2016. A fuzzy framework to evaluate service quality in the healthcare industry: An empirical case of public hospital service evaluation in Sicily. *Applied Soft Computing*, [e-journal] 40, pp.468-478. DOI: 10.1016/j.asoc.2015.12.010.

Naidu, A., 2009. Factors affecting patient satisfaction and healthcare quality. *International Journal of Health Care Quality Assurance*, [e-journal] 22(4), pp.366-381. DOI: 10.1108/09526860910964834.

Parasuraman, A, Zeithaml, V.A. and Berry, L.L., 1988. SERVQUAL: A multiple item scale for measuring customer perceptions of service quality. *Journal of Retailing*, 64(1), pp.12-40.

Ramsaran-Fowdar, R.R., 2008. The relative importance of service dimensions in a healthcare setting. *International Journal of Health Care Quality Assurance*, [e-journal] 21(1), pp.104-124. DOI: 10.1108/09526860810841192.

Roberge, D., Tremblay, D., Turgeon, M.È. and Berbiche, D., 2013. Patients' and professionals' evaluations of quality of care in oncology outpatient clinics. *Supportive Care in Cancer*, [e-journal] 21(11), pp.2983-2990. DOI: 10.1007/s00520-013-1872-x.

Singh, A. and Prasher, A., 2019. Measuring healthcare service quality from patients' perspective: using Fuzzy AHP application. *Total Quality Management & Business Excellence,* [e-journal] 30(3/4), pp.1-18. DOI: 10.1080/14783363.2017.1302794.

Sowa, J.F., 2013. What Is the Source of Fuzziness?. In: R. Seiging, E. Trillas, C. Moraga and S. Termini, eds. 2013. *On Fuzziness*. Berlin, Heidelberg: Springer. pp.645-652. DOI: 10.1007/978-3-642-35644-5_31.

Tóth, ZS.E., Jónás, T. and Dénes, R., 2018. Service quality evaluations in healthcare based on flexible fuzzy numbers - Patients and employees in the focus. In: S.M. Dahlgaard-Park and J.J. Dahlgaard, eds. 2018. *21st QMOD International Conference on Quality and Service Sciences*. Wales, UK, 22-24 August 2018. Lund: Lund University Library Press. pp. 972-983.

Tóth, ZS.E., Surman, V. and Árva, G., 2017. Challenges in course evaluations at Budapest University of Technology and Economics. In: B. Zafer, Y.M. Melis and X.T. Roslind, eds. 2017. *8th ICEEPSY - International Conference on Education and Educational Psychology*. Porto, Portugal, 11-14 October 2017. Porto: Future Academy. pp.629-641. DOI: 10.15405/epsbs.2017.10.60

Tsai, H-Y., Chang, C-W. and Lin, H-L., 2010. Fuzzy hierarchy sensitive with Delphi method to evaluate hospital organization performance. *Expert Systems with Applications,* [e-journal] 37, pp.5533-5541. DOI: 10.1016/j.eswa.2010.02.099.

Vandamme, R. and Leunis, J., 1993. Development of a multiple-item scale for measuring hospital service quality. *International Journal of Service Industry Management*, [e-journal] 4(3), pp.30-49. DOI: 10.1108/09564239310041661.

Woldegebriel, S., Kitaw, D. and Rafele, C., 2015. Application of Fuzzy Logic for Prioritizing Service Quality Improvement in Healthcare. *International Journal of Scientific and Engineering Research,* 6(5), pp.530-537.

Yeh, C.H. and Kuo, Y.L., 2003. Evaluating passenger services of Asia-Pacific international airports. *Transportation Research Part E: Logistics and Transportation Review*, [e-journal] 39(1), pp.35-48. DOI: 10.1016/S1366-5545(02)00017-0.

Yesilada, F. and Direktor, E., 2010. Health service quality: a comparison of public and private hospitals. *African Journal of Business Management*, 4(6), pp.962-971.

ABOUT AUTHORS

Zsuzsanna Eszter Tóth – (ZS.E.T.) Eötvös Loránd University, Budapest, Hungary, Department of Management and Business Law, Assoc. Prof., e-mail: tothzs@gti.elte.hu, Author's ORCID: 0000-0002-0628-1320.

Gábor Árva – (G.A.) Budapest University of Technology and Economics, Budapest, Hungary, Department of Management and Business Economics, Assist. Lecturer, e-mail: arva@mvt.bme.hu, Author's ORCID: 0000-0002-4330-1319.

Rita Veronika Dénes – (R.V.D.) Eötvös Loránd University, Budapest, Hungary, Department of Management and Business Law, Assist. Lecturer, e-mail: denes@gti.elte.hu, Author's ORCID: 0000-0003-1870-9942.

AUTHOR CONTRIBUTIONS

ZS.E.T. – conceptualization, writing—review and editing, supervision; G.A. – methodology, software, visualization; R.V.D. – investigation, writing—original draft preparation; ZS.E.T. and G.A. – formal analysis.

CONFLICTS OF INTEREST

The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Determinants of Courier Service Quality in e-Commerce from Customers' Perspective

DOI: 10.12776/QIP.V24I2.1438

Aleksandra Gulc

Received: 2020-03-30 Accepted: 2020-06-24 Published: 2020-07-31

ABSTRACT

Purpose: In recent years, the increasing popularity of e-commerce has become a driving force behind the development of courier service both in Poland and worldwide. In order to face growing competition on courier market and meet clients' expectations, one of the crucial strategic goal of courier operators is to provide high quality service. The aim of the article is to identify and classify the key factors which determine the perception of courier service quality by customers shopping online.

Methodology/Approach: The first part of paper is focused on the literature review concerning the determinants of courier service quality from customers' perspective. Next, a survey method was used to collect data among customers using courier service while shopping online in Poland. Finally, the exploratory factor analysis was used to indicate the key dimensions and factors effecting the courier service quality perceived by customers.

Findings: The article presents the complex review of literature concerning determinants of courier service and the original scale, which can be used to measure courier service quality. The results of author's empirical research indicated that, the key dimensions effecting courier service quality from the customers point of view were: Reliability, Visual Identification, Service Complexity, Relational Capital, Social Responsibility, Responsiveness, Technical Quality.

Research Limitation/implication: As the research aim was to develop the scale of courier service quality, in future research the CFA method (Confirmatory Factor Analysis) should be used to verify the scale. The main limitation of this research concerned a perspective of one stakeholder participating in the process of courier service – e-clients. The further research would be focused on identification of determinants of courier service quality in opinion of online

shops and courier companies and development of models presenting the relations between them.

Originality/Value of paper: This paper presents the original research that provides new knowledge about determinants of courier service quality in e-commerce. The literature analysis has shown that the previous research were often fragmentary and situational, but also did not take into account the specificity of e-commerce branch.

Category: Research paper

Keywords: courier service quality; determinants; e-commerce

1 INTRODUCTION

Courier service, which is a special kind of service within TSL sector, has recently become a critical link in logistics supply chains of many trading, manufacturing and service companies (Jarocka and Wang, 2018). Recently courier, express and parcel market (CEP) has developed dynamically worldwide and this trend will be maintained in the coming years. The value of the global courier market ammounted to 306 billion dollars in 2018, and according to forecasts it will reach 400 billion dollars by 2024, which means an increase of 8-10% annually (Statista, 2019). The value of European CEP sector has been growing recently by 7% annually while the volume of shipment by 8%. As far as the market value is concerned, Germany, Great Britain and France are the leaders. Moreover, 80% of total revenues from courier service in Europe is generated in eight countries – Germany, Great Britain, France, Spain, Italy, the Netherlands, Belgium and Poland (Dieke et al., 2019). Comparing to other European countries, the dynamics of CEP market value in Poland is one of the highest and it has increased by 11% annually. The growth rate of Polish courier market is three times faster than GPD in Poland. Nevertheless, the share of Polish CEP market has reached only 3% of European market value and the volume of shipments is eight times smaller comparing to Western countries. The experts predict that the growth of courier market in Poland will be developing much faster than in other European countries, which place Poland as a leader, next to Holland and Romania (Gawryluk, 2019).

Increasing popularity of online shopping has become one of the key driving force behind the development of courier service in recent years. A few years ago, courier companies were not interested in cooperation with e-commerce branch, because of the low scale of sales via Internet, the high fragmentation of deliveries and exceptional needs of individual customers. Since the opening of the World Wide Web for commercial use in 1995, the global value of e-commerce sales has continued to grow rapidly supported by increasing Internet connectivity, expanding global trade and more sophisticated shipping technology. In 2018, the value of global e-commerce market amounted to 2.3 trillion dollars, which meant 50% increase comparing to 2015. It is presumed that the global value of the

e-commerce market will grow dynamically over the next six years, reaching about 40% of the global market share in 2026. The growth rate of online sales will be more than four times faster than the total retail sales and B2C service will continue to thrive the most among the segments of e-commerce sector (ITF, 2019).

In condition of rising popularity of online shopping, courier service is an crucial element effecting e-customers satisfaction and future intentions, but also e-retailers success (Jun, Yang and Kim, 2004; Kawa, Pieranski and Zdrenka, 2018; Liu et al, 2008; Micu, Aivaz and Capatina, 2013; Valaei, Rezaei and Shahijan, 2016; Yu et al., 2012). E-shoppers appreciate convenience of delivery, so they expect services that will suit their needs. Therefore, most of the largest courier operators prepare the offer dedicated to e-commerce branch. The options like multiple delivery choices, real time delivery information, possibility to reschedule the delivery or same-day delivery are among the interest of e-shoppers (Gulc, 2017a). As far as the delivery location is concerned, European e-customers chose home as the most popular one, while parcel shops were their second choice, which confirmed that they prefer convenience of delivery. Regular e-shoppers had positive experience with courier service – 78% of them considered the delivery of their latest online purchase was easy and 60% indicated that the latest returns experience was positive (DPD Group, 2019). With regard to the technology development, the courier service offer has evolved from traditional door-to-door delivery to a self-service based on non-human interfaces with the use of information technologies and modern logistics solutions (PWC, 2019). The final supervisor, which verify and evaluate the service offer, are clients. Thus, the courier service should not be treated as an offer of provider, but as a result of close cooperation with the customers. In client-oriented company, the customers should be perceived as one of resource bringing the essential information about their experiences, needs and expectations towards service quality (Prahalad and Ramaswamy, 2004; Grönroos, 2007; Grönroos and Ravald, 2011; Vargo and Lusch, 2004). Therefore, the further research concerning the service quality from the customers point of view, seems to be crucial and desired in condition of the evolution of courier service driven by e-commerce development and technology progress.

The aim of the article is to identify and classify the key factors which determine courier service quality perceived by clients shopping online. The main result of the research is the development of new scale which can be used as a measuring tool of courier service quality from the perspective of e-customers. The remainder of the paper consists of five sections. Section 2 presents the literature review on the issues relating to the aspects of courier service quality, especially the factors which determine the service quality and methods of measuring service quality. The analysed research were focused on the assessment of courier service quality from the perspective of business or individual clients. The results of the literature review served as a basis for the formulation of the list of potential factors determining the courier service quality in e-commerce. Section 3

describes the research methodology, while the results and discussion are described in section 4 and 5. Finally, section 6 summarizes the findings and the conclusions briefly explaining the limitations of the research and implications for future research efforts.

2 LITERATURE REVIEW

This section presents the critical review of chosen publications concerning courier service quality especially the determinants and methods of measuring service quality. Although, CEP market has been constantly developing and the service quality is one of the key goals for courier operators, there are only a few articles concerning this issue both in the foreign and Polish literature.

Most of analysed research were based on SERVQUAL method in order to assess the service quality dimensions (tangibility, reliability, assurance, responsiveness and empathy) (Frąś, 2014; Yee and Daud, 2011; Tabassum and Badiuddin, 2014; Yu et al., 2012). Yee and Daud (2011) examined the influence of dimensions of service quality on customers' satisfaction of parcel delivery. The research revealed that tangibility, reliability and assurance had an impact on customer satisfaction, while empathy and responsiveness had no significant effect on customer satisfaction. Similarly, Tabassum and Badiuddin (2014) based their research on SERVQUAL method in its original form, however they adjusted particular items to the specificity of courier service. The main objective of this study was to determine the gap between expectations and perception of service quality by customers. The highest gap was observed in case of reliability and responsiveness, so these issues should be strengthen by courier operators. The original theoretical approach presented Yu et al. (2012), who deliberated the method of courier quality service improvement called Two-Stage Quality Functional Deployment in order to transform express service demand into the express service resources. This research was focused on eliminating internal factors, which adversely affected service quality within the enterprise, but it did not verify their influence on service quality perceived by clients. Although, this method was verified in express industry, the authors suggest that it can be adopted to other branches of service sector.

Although the SERVQUAL is the most popular method of service quality measuring, it is also widely criticised because of its general character, so it requires modification to the specificity of particular service type. Moreover, the method does not distinguish the difference between satisfaction and service quality itself. The critics also undermined the main assumption of this method concerning the overall service quality which is calculated as the difference between expected and perceived quality. In fact, the clients usually assess perceived service quality without the comparison with their expectations (Gulc, 2017b).Therefore, some authors used less complex method of measuring courier service quality perceived by clients. A few previous studies considered the specific features of courier service (Gulc, 2017c). Some authors used logistic

service quality scale (Liu and Liu, 2014; Ho et al., 2012), while others try to develop an original set of dimensions of courier service (Valaei, Rezaei and Shahijan, 2016; Gulc, 2017c). The study of Liu and Liu was focused on the assessment of the express logistics service quality on the basis of SERVQUAL method but with the use of Logistic Service Quality (LSQ) scale including the following dimensions: reliability, protection, security, empathy and perception. The research results revealed that customers were not satisfied with the protection, reliability and empathy while express service (Liu and Liu, 2014). Ho et al. (2012) used the LSQ scale in order to determine the most important factor of courier service quality effecting customers' satisfaction. The scale included the following dimensions: timeliness, condition/accuracy of order, quality of information and availability/quality of personnel. A multiple regression analysis indicated that timeliness, which was usually one of the most important dimension of courier service quality in other research, was replaced by condition/accuracy of order. The aim of study of Valaei, Rezaei and Shahijan (2016) was to determine the dimentions of courier service quality and the impact of perceived service quality on overall service quality. The study was based on SERVQUAL method, however the authors formulated the scale in the context of courier service titled as 'CouQual' including the following dimensions: promptness, convenience, accuracy, safety, and tangibles. The research results implied that the most important factors were promptness, safety and convenience, while accuracy and tangibility did not positively contribute to perceived service quality. The authors suggested that the model should be implied in different countries in order to check if the culture of different country could effect the perception of courier service quality.

In Polish literature, there are only a few papers concerning courier service quality, some of them were concentred on B2B service (Chodak, Latus and Prałat, 2010; Chodak, 2013; Dmowski, Śmiechowska and Zelmańska, 2013) while others on B2C service (Frąś, 2014) or both B2B and B2C (Dyczkowska, 2011; Gulc, 2017c). The research of Frąś aimed to evaluate the perceived quality according to SERVQUAL scale. The research results showed that the lowest rated dimension was responsiveness, and the highest – tangibility (Frąś, 2014). Chodak, Latus and Prałat's research (2010) was concentrated on the evaluation of the cooperation between courier or postal operators and online shops. The clients assessed service quality taking into account timeliness, delivery damage and loss, but also courier attitude and time of settlement in case of cash payment. The respondents were more satisfied with courier service than traditional postal service. Chodak (2013) conducted further research on courier service in e-commerce in Poland. The results of the research confirmed that over 80% of the online shops assessed courier service quality very highly. Respondents indicated the following factors which influenced the choice of courier companies: timeliness of services, service price, competence and culture of couriers, time of reaction and tracking the shipment. Less important factors were: diversity of service, availability of a program for printing waybills and labels, the time of money return and the least important – flexibility of delivery time and the

functioning of one nationwide phone number that guarantees quick contact (Chodak, 2013). Dmowski et al. tried to indicate the most important factors of courier service quality from the perspective of business clients. The research focused on the following dimensions of service quality: knowledge and competence of staff, level of customer service, time of response and time of delivery, communication between courier companies and clients, the willingness and engagement to solve problems. The results revealed that a high overall quality assessment depended on functional dimensions of service quality (Dmowski, Śmiechowska and Zelmańska, 2013). The research conducted by Dyczkowska (2011) was focused mainly on the identification of factors determining the choice of courier company, but also the assessment of service quality by customers and the reasons of their complaints. The results indicated the factors affecting courier service quality perceived by customers, in particular the time of order completion, service price, lack of damage, the offer complexity, low complaint level and the quality of customer service. Inspired by the Franceschini's concept of aging of quality indicators, Gulc carried out a pilot research concerning the present and future expectations of clients using courier service. On the basis of literature and current trends on courier market, the author prepared a set of thirteen factors determining the choice of courier service. The respondents – individual and business clients of courier companies in Poland, assessed the importance of each criteria nowadays and in in the perspective of the next 5-10 years. The results showed that the most important criteria for clients would be time of delivery, trust, flexibility and tele-technologies, while the price would be less important (Gulc, 2017c).

The literature review carried out on the courier service quality led to the following conclusions:

- The authors of analysed research have not developed a measuring scale of courier service quality in the context of e-commerce in B2C relation.
- The research conducted so far in relation to courier service were often fragmentary and situational – the analysed papers were focused on a narrow group of respondents, a selected region of country or only one of the quality criteria.
- The empirical research concerning CEP market in the context of e-commerce in Poland is very limited.

The review of the literature allowed to formulate the following research questions:

RQ1. Which factors significantly affect the courier service quality in e-commerce from the perspective of clients shopping online?

RQ2. How can the factors affecting courier service quality be classified?

RQ3. Which aspects of courier service are the most important for e-customers in the context of service quality?

3 METHODOLOGY

In order to explain the phenomena, which is courier service quality and develop an original measuring scale, an exploratory factor analysis (EFA) was used. Such approach is particularly appropriate when the phenomena is poorly recognized in the literature and the researcher does not assume a priori hypothesis about factors or patterns of measured variable. The analysed research did not reflect the specific features of courier service in e-commerce sector so that the author decided to use the EFA method in order identify a set of latent constructs (dimensions) determining courier service quality and develop a measurement scale. The method is used by researchers to underlay a battery of measured variables, uncover the underlying structure of a relatively large set of variables and identify relationships between measured variables (Norris and Lecavalier, 2009; Fabrigar et al., 1999). The EFA method is often used in the field of logistics (Ryciuk, 2016; Leończuk, 2019) and management (Ejdys, 2018; Wipulanusat, Panuwatwanich and Stewart, 2017).

The empirical research process included three stages: the preparation of initial list of factors on the basis of literature review, the survey conducted among e-clients using courier service in Poland and statistical analysis of obtained data with the use of exploratory factor analysis.

On the basis of literature review presented in section 2 and available documents of courier operators, the author prepared the list of factors which could determine the courier service quality perceived by individual customers shopping online. The final questionnaire included 38 factors. The empirical research was conducted with the use of CAWI method among individual clients using courier service while shopping online in Poland. The research was conducted in the period of November 2018 till January 2019 and the final number of respondents, who took part in the research and completed the survey amounted 594 persons. In order to identify the determinants of courier service quality, the respondents were asked the following question: *What is the impact of the following factors on the courier service quality in your opinion?* To answer the question, the 7-point Likert scale was used (in which 1 – *very weak* and 7 – *very strong impact on courier service quality*).

The research sample included the similar proportion of women to men, with slight predominance of women (52%) than men (48%). Taking into account the age of respondents, the largest group included respondents at the age of 36-55 years old (281 persons), which amounted approximately 47% of the total number of respondents. The second largest group consisted of e-customers aged 26-35 (22%). The share of people between 18-25 years old (14%) and those over 55 years old (16%) was similar. The least numerous group included respondents below 18 years old (0.34%). As far as the education level was concerned, the majority of respondents were persons with higher education – approximately 81%, less numerous was the group with medium education (17%), while the

minority were those having primary and vocational level of education – approximately 2%.

The respondents participating in the survey represented all 16 voivodships and their distribution was in correspondence with the structure of the general population of e-clients in Poland. The largest number of respondents were inhabitants of more populated voivodships, where more online shops were located – mazowieckie (91 respondents), śląskie (57 respondents), małopolskie (48 respondents). The least number of respondents was observed in case of less populated voivodeships – lubuskie (22 respondents), opolskie and świętokrzyskie.

Most of respondents, who participated in the survey, used courier service very frequently – more often that once a month (almost 60%), so they are the most reliable source of information. Over 30% used it once a quarter and only about 7% of respondents used courier service very rare – once a year or even less.

In the first stage of the analysis, the assumptions for using the factor analysis were checked: requirement of test sample (minimum 200 respondents) and the number of variables (3-4 variables describing the potential construct). Both conditions in this research were met. Next, the correlation between the examined variables was analysed and it was stated that each variable significantly correlated with several other variables, which was also proved by the determinant of the correlation matrix amounting 0.00000000002424. The Bartlett sphericity test also indicted that the correlation matrix contained significant available correlations. Moreover, the high value of Kaiser-Mayer-Olkin criterion (approximately 1) was also the premise for exploratory factor analysis.

To determine the number of factors (dimensions), the principal components analysis was used. Using the Kaiser's criterion, which is called eigenvalue-greater-than-one rule, seven factors were distinguished, which explained over 66.12% of the output variance of the variables. Similar results were obtained using the Cattel method, based on a graph showing seven factors with eigenvalues greater than 1. For a seven-factor solution, a factor analysis was carried out using the principal axis method with Varimax normalized orthogonal rotation. The principal axis method was chosen as it requires no distributional assumptions and may be used if data are not normally distributed (Fabrigar et al., 1999). Orthogonal rotations are appropriate when the purpose for the factor analysis is to generate factor scores or when the theoretical hypotheses concern uncorrelated dimensions. Among the orthogonal types of rotations, Varimax is generally regarded as best and is most often used (Fabrigar et al., 1999). Later, the factor loadings were analysed. According to Bedyńska and Cypryańska (2013), if the factor loading is greater than 0.4, it means that this variable is strongly loaded with the factor. However in practice, higher values are taken as often as possible. In order to indicate the variables which have strong contribution into particular construct, only the factors with the value of factor loading over 0.5 were chosen to the analysis. Using the factor loading matrix,

insignificant variables with factor values less than 0.5 were removed, among others: *V1 – Modern transport fleet, V2 – Modern and functional ICT technologies, V9 – Attractive prices and discounts, V14 – Transparent procedures, documents and service standards, V15 – Simplicity of placing an order, V21 – Cultural and polite behaviour of employees, V22 – Security of transactions, V31 – Individualisation of the service, V32 – Paying full attention to customer, V33 – Ensuring that the interests of the client are realised.* All analysis presented in this paper were done with the use of Statistica software.

4 RESULTS

In result of factor analysis, 27 variables were classified into seven hidden variables (constructs/factors). In the Table 1, there are presented factor loadings, which are numerical values that indicate the strength and direction of a factor on a measured variable, so they indicate how strongly the factor influences the measured variable.

In result of exploratory factor analysis, the new scale was developed including seven dimensions and variables determining courier service quality (Table 1). The validity of the questionnaire was measured with the use of Cronbach's alpha coefficient. The coefficient values reached high level (above 0.7) in case of all distinguished dimensions. Therefore, it can be concluded that the developed scale is a reliable measurement tool (Sagan, 2004).

Table 1 – Results of Exploratory Factor Analysis and Reliability of Scale

Dimension	**Variable**		**Factor loadings**	**Cronbach's alpha coefficient**
F1: Reliability	V10	Timeliness of delivery	0.773	0.902
	V11	Successful delivery attempt	0.782	
	V12	Compliance and completeness of delivery	0.728	
	V13	Lack of damage to the package	0.785	
	V27	Efficient and fast handling of order	0.604	
F2: Visual identification	V6	Aesthetic and neat appearance of courier	0.716	0.,784
	V7	Characteristic trade mark and uniform colour	0.690	
	V8	Aesthetic and functional company branches and pick-up points	0.637	
F3: Service complexity	V36	Wide range of additional service	0.785	0.809
	V38	Service diversity	0.712	
F4: Relational capital	V16	Positive experience with courier service	0.576	0.868
	V17	Positive feedback from other customers	0.547	
	V18	Positive image and brand of courier company	0.715	

Dimension	Variable		Factor loadings	Cronbach's alpha coefficient
	V19	Experience and credibility of courier company	0.680	
	V20	Knowledge and competence of employees	0.510	
	V25	Trust in the courier company	0.643	
F5: Social responsibility	V4	Ecological technical solutions	0.693	0.708
	V34	Courier company's involvement in social actions	0.574	
F6: Responsiveness	V23	Easy contact with the courier company	0.611	0.893
	V24	Efficient communication with courier company employees	0.773	
	V26	Accurate and clear information on the conditions of service provision	0.531	
	V28	Efficient handling of returns	0.604	
	V29	Readiness to react quickly to reported problems and interference	0.736	
	V30	Flexibility (choice and change of date and place of service)	0.543	
F7: Technical quality	V3	Modern ecological solutions	0.513	0.762
	V37	Availability of the service	0.579	
	V39	Choice of the delivery method	0.529	

The research result revealed that among all examined dimensions effecting the courier service quality from the perspective of e-customers, *Reliability* was the most important (mean value – 6.67), followed by *Responsiveness* (mean value – 6.15). *Technical quality* (mean value – 5.82) and *Relational capital* (mean value – 5.46) were also highly rated. On the other hand, *Service complexity* was assessed as neutral dimension (mean value – 4.40), while the least important was *Visual identification* (mean value – 3.69) and *Social responsibility* (mean value – 3.24) (Figure 1).

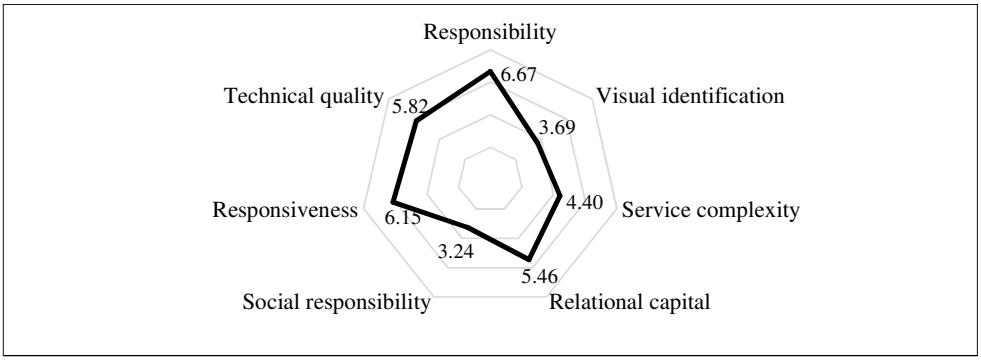

Figure 1 – Distribution of Means for All Dimensions of Courier Service Quality from e-clients' Perspective

5 DISCUSSION

The research results indicate that courier service quality in e-commerce, which is analysed in the perspective of e-customers, is determined by the following dimensions:

F1: Reliability – contains the most important factors that affect courier service quality from the perspective of receiver, such as timeliness of delivery, successful attempt of delivery, compliance and completeness of the order, but also the lack of damage to the parcel, efficient and quick order processing;

F2: Visual identification – includes elements that create the image of courier company like aesthetic and neat appearance of courier, characteristic trademark and uniform colours, as well as aesthetic and functional company branches or Pick Up-Drop-Off locations (PUDO);

F3: Service complexity – is directly related to the offer of courier service including a wide range of additional service as well as the diverse scope and range of service;

F4: Relational capital – contains elements that contribute to long-term relationships between the courier company and clients, including the positive experience of clients using the courier service, opinions of other clients, as well as trust in the courier company, positive image and brand, experience and credibility, knowledge and competence of employees;

F5: Social responsibility – contains selected elements related to the social responsibility of courier companies, such as the involvement of courier companies in social actions and the use of ecological technical solutions (for example cars, electric drones, bicycles or ecological packaging);

F6: Responsiveness – is primarily associated with easy and efficient communication with the employees of the courier company, but also the willingness of employees to respond quickly to the problems and disruptions reported by the customers, efficient handling of return of ordered goods, providing accurate and clear information on the service terms and flexibility by ensuring that the date and place of delivery are selected or changed;

F7: Technical quality – includes selected elements related to the material aspects of the service like modern technical solutions dedicated for customers including the network of PUDO and parcel lockers and drones, the option of choosing the method of sending or delivery and the availability of service by providing a convenient location for PUDO and working hours.

The result of empirical research and statistical analysis indicated the determinants of courier service quality in e-commerce perceived by e-customers, which had not been mentioned in the analysed literature, especially relational capital and social responsibility. According to the obtained results, the most important dimension affecting courier service quality was reliability manifesting in timeliness, successful delivery attempt, completeness of delivery and lack of

damage to the parcel. This tendency was also confirmed by other research (Ho et al., 2012; Valaei, Rezaei and Shahijan, 2016; Dyczkowska, 2011; Chodak, 2013). Similarly to other authors' studies, the research revealed that customers appreciated the responsiveness of staff while using courier service (Dmowski, Śmiechowska and Zelmańska, 2013). Moreover, the technical aspects of courier service were very important for e-customers, which is in accordance with other research conducted in Poland (Frąś, 2014) and at the same time on the contrary to the foreign research (Valaei, Rezaei and Shahijan, 2016). Taking into account the previously analysed papers, the developed scale revealed new aspects of courier service quality, which are important from the e-customers point of view.

6 CONCLUSION

In the light of the literature review, the study confirmed that there has not been elaborated a universal set of factors determining courier service quality. Their choice depended on the type of provided service, the segment of customers, but also the geographical and cultural area. This paper addressed an important and current research area that has not been sufficiently developed in the literature so far. The results of empirical research and statistical analysis revealed the determinants of courier service quality in e-commerce perceived by e-customers. On the basis of empirical research and conducted statistical analysis, 27 from 39 factors were classified according to 7 dimensions namely: *Reliability, Visual Identification, Service Complexity, Relational Capital, Social Responsibility, Responsiveness* and *Technical Quality*. The developed scale including dimensions and factors affecting courier service quality, need to be tested in order to be used as measurement tool in further research.

The main limitations of this research concerned a perspective of only one group of stakeholders participating in the process of courier service – e-clients. Therefore, the achieved results could be useful, serving as a starting point for the future research. The author intend to use CFA method (Confirmatory Factor Analysis) in order to test the hypothesis that a relationship between the observed variables and their underlying latent factors exists. The further research will be focused on identification of determinants of courier service quality in opinion of online shops and courier companies. The final result of all research steps will be the model of courier service quality in e-commerce presenting relations between three stakeholders participating in courier service (e-shops, courier company and final receivers).

ACKNOWLEDGEMENTS

The study was carried out as part of Research Project No. 2017/25/N/HS4/02051, titled "Relational model of quality shaping system of courier services in e-commerce", which was financed by the National Science Centre, Poland.

REFERENCES

Bedyńska, S. and Cypryańska, M. eds., 2013. *Statystyczny drogowskaz 1. Praktyczne wprowadzenie do wnioskowania statystycznego*. Warszawa: Wydawnictwo Akademickie SEDNO.

Chodak, G., 2013. Metody dostarczania towarów przez polskie sklepy internetowe – wyniki badań. *Gospodarka Materiałowa i Logistyka*, 65(6), pp.21-32.

Chodak, G., Latus, Ł. and Prałat, E., 2010. Współpraca sklepów internetowych z przedsiębiorstwami kurierskimi i Pocztą Polską – wyniki badań. *Gospodarka Materiałowa i Logistyka*, 62(6), pp.18-26.

Dieke, A.K., Arnold, R., Bender, CH., Hillebrand, A., Niederprum, A. Taş, S., Tiele, S and Wielgosch, J., 2019. *Development of Cross-border E-commerce through Parcel Delivery*. [pdf] Luxembourg: Publications Office of the European Union. Availabe at: <https://www.wik.org/fileadmin/Studien/2019/ET0219218ENN_ParcelsStudy_Final.pdf> [Accessed 30 April 2020]. DOI: 10.2873/931558.

Dmowski, P., Śmiechowska, M. and Zelmańska M., 2013. Jakość jako czynnik budujący przewagę konkurencyjną na rynku usług kurierskich. In: G. Rosa and A. Smalec, eds. 2013. *Marketing przyszłości. Trendy. Strategie. Instrumenty*. Szczecin: Uniwersytet Szczeciński. pp.167-180.

DPD Group, 2019. *E-shoppers in Europe: 2019 Barometer*. [pdf] DPD Group. Available at: <https://www.dpd.com/group/wp-content/uploads/sites/77/2020/02/E-shoppers-in-Europe_2019-Barometer_EN.pdf> [Assessed 30 May 2020].

Dyczkowska, J., 2011. Klient na rynku usług TSL. *Zeszyty Naukowe Instytutu Ekonomii i Zarządzania, Politechnika Koszlińska*, 15, pp.119-134.

Ejdys, J., 2018. *Zaufanie do technologii w e-administracji*. Białystok: Oficyna Wydawnicza Politechniki Białostockiej.

Fabrigar, L.R., Wegener, D.T., MacCallum, R.C. and Strahan, E.J., 1999. Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, [e-journal] 4(3), pp.272-299. DOI: 10.1037/1082-989X.4.3.272.

Frąś, J., 2014. Wybrane instrumenty pomiaru jakości usług logistycznych. *Finanse, Rynki Finansowe, Ubezpieczenia, Zeszyty Naukowe Uniwersytetu Szczecińskiego*, 66(803), pp.297-317.

Gawryluk, M., 2019. *Rozwój rynku przesyłek kurierskich, ekspresowych i paczkowych (KEP) w Polsce od 2014 r. do 2023 r.* [online] Warszawa: Poczta Polska. Available at: <https://media.poczta-polska.pl/pr/465205/poczta-polska-w-2023-roku-rynek-kep-bedzie-mial-wartosc-12-mld-zl> [Accessed 12 May 2020].

Grönroos, CH., 2007. *Service Management and Marketing: Customer Management in Service Competition*. West Susex: John Willey & Sons.

Grönroos, CH. and Ravald A., 2011. Service as business logic: implications for value creation and marketing. *Journal of Service Management*, [e-journal] 22(1), pp.5-22. DOI: 10.1108/09564231111106893.

Gulc, A., 2017a. Charakterystyka usług kurierskich w Polsce. *Humanities and Social Sciences*, 22(24), pp.61-75.

Gulc, A., 2017b. Models and Methods of Measuring the Quality of Logistic Service. *Procedia Engineering*, 18, pp.255-264.

Gulc, A., 2017c. Courier service quality from the clients' perspective. *Engineering Management in Production and Services*, [e-journal] 9(1), pp.36-45. DOI: 10.1515/emj-2017-0004.

Ho, J.S.Y., Teik, D.O.L., Tiffany, F., Kok, L.F. and Teh, T.Y., 2012. Logistic Service Quality among Courier Services in Malaysia. *International Proceedings of Economics Development & Research*, 38(2012), pp.113-117.

ITF, 2019. *ITF Transport Outlook 2019*. [online] Paris: OECD Publishing. Available at: <https://www.oecd-ilibrary.org/transport/itf-transport-outlook-2019_transp_outlook-en-2019-en> [Assessed 04 May 2019].

Jarocka, M. and Wang, H., 2018. Definition and classification criteria of logistics services for elderly. *Engineering Management in Production and Services*, [e-journal] 10(4), pp.65-75. DOI: 10.2478/emj-2018-0023.

Jun, M., Yang, Z. and Kim, D., 2004. Customers' perceptions of online retailing service quality and their satisfaction. *International Journal of Quality & Reliability Management*, [e-journal] 21(8), pp.817-840. DOI: 10.1108/02656710410551728.

Kawa A., Pieranski B. and Zdrenka W., 2018. Dynamic Configuration of Same-Day Delivery in E-commerce. In: A. Sieminski et al., eds. *Modern Approaches for Intelligent Information and Database Systems. Studies in Computational Intelligence*. Cham: Springer. pp.305-315. DOI: 10.1007/978-3-319-76081-0_26.

Leończuk, D., 2019. *Ocena dokonań adaptacyjnego łańcucha dostaw*. Białystok: Oficyna Wydawnicza Politechniki Białostockiej.

Liu, L. and Liu, CH., 2014. Empirical Study of Express Logistics Service Quality – A Survey of Changdao County Express Sector. In: ICETIS (International Conference on Education Technology and Information System), *Proceedings of the 2014 International Conference on Education Technology and Information System (ICETIS 2014)*. Jinan, China, 26-27 April 2014. Atlantis Press. pp.542-546. DOI: 10.2991/icetis-14.2014.121.

Liu, X., He, M., Gao, F and Xie, P., 2008. An empirical study of online shopping customer satisfaction in China: a holistic perspective. *International Journal of Retail Distribution Management*, [e-journal] 36(11), pp.919-940. DOI: 10.1108/09590550810911683.

Micu, A., Aivaz, K. and Capatina, A., 2013. Implications of logistic service quality on the satisfaction level and retention rate of an e-commerce retailer's customers. *Economic Computation & Economic Cybernetics Studies & Research*, 47(2), pp.147-155.

Norris, M. and Lecavalier, L., 2009. Evaluating the Use of Exploratory Factor Analysis in Developmental Disability Psychological Research. *Journal of Autism and Developmental Disorders*, [e-journal] 40(1), pp.8-20. DOI: 10.1007/s10803-009-0816-2.

Prahalad, C.K. and Ramaswamy, V., 2004. Co-creation experiences: the new practice in value creation. *Journal of Interactive Marketing*, [e-journal] 18(3), pp.5-14. DOI: 10.1002/dir.20015.

PWC, 2019. *Five Forces Transforming Transport & Logistics, Trend Book.* [pdf] PwC. Available through: <https://www.pwc.pl/pl/pdf/publikacje/2018/transport-logistics-trendbook-2019-en.pdf> [Accessed 11 May 2020].

Ryciuk, U., 2016. *Zaufanie międzyorganizacyjne w łańcuchach dostaw w budownictwie.* Warszawa: Wydawnictwo Naukowe PWN SA.

Sagan, A., 2004. *Badania marketingowe. Podstawowe kierunki.* Kraków: Wydawnictwo Akademii Ekonomicznej w Krakowie.

Statista, 2019. *Retail e-commerce sales worldwide from 2014 to 2023 (in billion U.S. dollars).* [Industries > E-Commerce > B2C E-Commerce]. Statista. [online] Availabe through: <https://www.statista.com/statistics/379046/worldwide-retail-e-commerce-sales/> [Assessed 10 March 2019].

Tabassum, R. and Badiuddin, A., 2014. Measuring the service quality gap in courier industry. *IRJA-Indian Research Journal*, 1(5), pp.1-11.

Valaei, N., Rezaei, S. and Shahijan, M.K., 2016. CouQual: assessing overall service quality in courier service industry and the moderating impact of age, gender and ethnicity. *International Journal of Management Concepts and Philosophy*, [e-journal] 9(2), pp.144-169. DOI: 10.1504/IJMCP.2016.077770.

Vargo, S.L. and Lusch, R.F., 2004. Evolving to a New Dominant Logic for Marketing. *Journal of Marketing*, [e-journal] (68)1, pp.1-17. DOI: 10.1509/jmkg.68.1.1.24036.

Wipulanusat, W., Panuwatwanich, K. and Stewart, R.A., 2017. Exploring leadership styles for innovation: an exploratory factor analysis. *Engineering Management in Production and Services*, [e-journal] 9(1), pp.7-17. DOI: 10.1515/emj-2017-0001.

Yee, H.L. and Daud, D., 2011. Measuring Customer Satisfaction in the Parcel Service Delivery: a Pilot Study in Malaysia. *Business and Economic Research*, [e-journal] 1(1), pp.1-12. DOI: 10.5296/ber.v1i1.1125.

Yu, B., Zhang, S., Wu, S. and Xie, J.A., 2012. Study of Courier Service Quality Improvement Based on a Two-Stage QFD. In: Z. Zhang, R. Zhang and J. Zhang, eds., *LISS 2012: Proceedings of 2nd International Conference on Logistics, Informatics and Service Science*. Beijing, China, 12-15 July 2012. Berlin: Springer. pp.885-889. DOI: 10.1007/978-3-642-32054-5_123.

ABOUT THE AUTHOR

Aleksandra Gulc – Bialystok University of Technology, Poland, Faculty of Engineering Management, International Departament of Logistics and Service Engineering, e-mail: a.gulc@pb.edu.pl, Authors' ORCID: 0000-0002-6401-4765.

CONFLICTS OF INTEREST

The author declares no conflict of interest. The funder had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Co-creation as a Success Factor in the Development of Constructive Customer-Focused Dialogues

DOI: 10.12776/QIP.V24I2.1442

Anna Mårtensson, Pernilla Ingelsson, Ingela Bäckström

Received: 2020-04-02 Accepted: 2020-05-25 Published: 2020-07-31

ABSTRACT

Purpose: The purpose of this study is to describe the process of co-creating a dialogue model that aims to increase citizen value in a municipality organisation. In addition, the purpose is to present the results from the development process and to evaluate the chosen dialogue model.

Methodology/Approach: A dialogue model was developed through a co-creation process with a series of workshops, discussions and interactive tasks. The whole process was carried out in three steps. In the first step, the success factors of a constructive dialogue based on citizen value were identified. In the second step, several dialogue models were developed, tested and evaluated, and one dialogue model was chosen. In the third step, the chosen dialogue model was evaluated.

Findings: An evaluation of a real-life use of the dialogue model supports the finding that the process delivers a dialogue model that enables the required prerequisites for constructive dialogues: for example, the opportunity to prepare, to create structured and transparent documentation, and to enable a holistic view of the dialogue model. A co-creation that involved co-workers contributed to developing a dialogue model that could be adapted to the co-workers' own context.

Research Limitation/implication: This study is conducted in a single organization, hence no generalizable conclusions can be made.

Originality/Value of paper: Using a co-creative process when developing and realising a dialogue model enhances the possibility of adapting the model to an organisation's specific context.

Category: Case study

Keywords: co-creation; customer focus; dialogue model; A3; holistic view

1 INTRODUCTION

When organisations try to understand customers and stakeholders, the organisations find the customer and stakeholder needs and expectations to be the most challenging to understand (Westher, 2018). After understanding customer needs and expectations, the organisation must be managed to fulfil these. This process demands knowledge about customers and the capacity to develop organisations in line with customer demands. Customers' needs and situations or contexts must be understood to create customer value (Kristensson, Gustafsson, and Witell, 2014). Creating customer value is considered the most valuable aspect of organisations (Oh, 1999) and is considered the source of other values in an organisation (Hammer, 1996). Notably, customers can be found in all organisations, even though some organisations do not refer to them as customers (Bäckström, 2009).

Quality management is defined by Flynn, Schroeder and Sakakibara (1994) as '*an approach to achieving and sustaining high quality output*', which supports both a customer focus by understanding value creation from the customer's perspective and by creating the values, methods and tools to develop organisations in a customer-oriented direction (Bergman and Klefsjö, 2010). A focus on customers aligns with the values of the cornerstone model that form the basis of quality management: having committed management, focusing on processes, improving continuously, basing decisions on facts, and enabling everyone to be committed (Bergman and Klefsjö, 2010). The value 'focus on customers' is considered the purpose of quality management and is claimed to contribute to knowledge and understanding of customer value creation. The relationship between internal and external customers is that internal customers–employees–need to be satisfied to deliver an outcome in line with the satisfaction of external customers (Dahlgaard, Kristensen and Kanji, 2002). The satisfaction of customers can be equated to achieving customer value, which Carlsson and Wilmot (2006) explain as giving value to internal customers and making it possible to maximise the value of external customers.

Organisational values are fundamental for organisational culture, which is unique for each organisation (Campbell, 2004). Organisational culture is connected to organisational performance in two ways: culture energises employees by appealing to their higher ideals and undefined values, and it shapes and coordinates behaviours and decisions (Chatman and Eunyoung Cha, 2003; Grönfeldt and Strother, 2006). According to Wu, Zhang and Schroeder (2011), quality management culture is important because it influences how quality practices are customised to achieve higher performance outcomes. Ingelsson, Eriksson and Lilja (2012) describe a developed model that connects organisational values with a strategy for selecting people within the organisation who support the organisation's values. The model is inspired by Hellsten and Klefsjö (2000) and has the overall aim of increasing internal and external customer satisfaction with a reduced amount of resources. Managers and employees are the organisational resources that affect the organisational culture

and thereby the performance outcome. According to Bäckström (2009) and Lagrosen, Bäckström and Wiklund (2012), the 'leadership commitment' and 'participation of everybody' values are important for supporting sustainable health among co-workers when quality management is practised. Employees also have knowledge of the unique context, which is a crucial factor of success in quality management work (Asif et al., 2009).

By tapping into employees' capacity, their existing knowledge base can be used to reach newly developed products, services or processes (Carmeli, Gelbard and Reitner-Palmon, 2013). Other ways to become more innovative are to involve other actors in society, such as industry, academia, suppliers and customers (Gassmann and Enkel, 2004). Lee, Hwang and Choi (2012) express that innovation in the public sector is driven to improve service performance and add value through public benefit. Different definitions exist for the terms of development, change and innovation. Within each context, a label needs to be placed on what is being done. In this paper, the term development is used to mean taking steps forward to increase value for customers.

1.1 Co-creation for Development

Co-operation has a long history in the field of quality management. At the beginning of the 20th century, Frederick Winslow Taylor described the importance of co-operation between employers and employees for achieving expected outcomes to reduce waste on the production line (Taylor, 1998). In this section, the function of co-operation is understood as the dissemination of information that has been developed through–among other things–the cornerstone model, which is used to describe the focus on customers and their involvement and the stimulation of customer engagement (Bergman and Klefsjö, 2010). Research has identified the need to find models for co-operation within organisations regarding, for example, the avoidance of misunderstandings and increase in knowledge transfer. Cronemyr (2007) demonstrate positive results in these areas through the method of 'knowledge overlapping seminars', that is, seminars that aim to achieve common concepts, fewer misunderstandings, better quality, and more satisfied customers.

A more elaborated form of co-operation is co-creation. Co-creation can be viewed as a collaborative process used in business through which customers and organisations interact to identify needs and benefits from services and the goals (Leavy, 2012; Lee, Olson and Trimi, 2012). When applied in the context of increased citizen value, co-creation can be viewed as engendering participation and transferring knowledge. Lee, Olson and Trimi (2012) point to the significance of co-creation in assisting organisations with establishing shared values.

Co-creation can be carried out in many different ways. Dialogues present a process of increasing knowledge and highlighting a subject from different angles. According to Hamilton and Pinnegar (2015), a dialogue is defined as '*a process*

of coming-to-know through which meaning is made and on the strength of which we develop assertions for action or understanding' and contains '*inquiry, critique, evidence, reflection, and response'*. The interest in dialogues has increased in organisations with public relations and has played a key role in organisations that want to involve stakeholders from a broader perspective (Gutiérrez-García, Recalde and Piñera-Camacho, 2015). In addition to encompassing a broader stakeholder perspective, dialogues also support a management culture that is able to meet the various expectations of a wide range of individuals and groups (ibid.).

In appreciative inquiry (AI), appreciative interviews are one-on-one dialogues that focus on asking questions and sharing stories about highpoint experiences and valuable activities–in a positive spirit (Cooperrider and Whitney, 2005). The interviews generate results that assist organisations in growing in the requested direction (ibid.). AI is a theory, an approach and a mindset that leads to creativity and organisational learning (Watkins and Cooperrider, 2000). AI builds a vibrant, high-performing and customer-focused culture by systematically discovering what gives life to an organisation and when the organisation is the most effective and capable in economic, ecological and social terms (Cooperrider and Whitney, 2005). The culture of AI is a culture that Whitney and Trosten-Bloom (2010) describe as working with the strengths of the persons involved.

1.2 Increased Customer Benefit: Contextual Challenges

Municipalities in Sweden are required to follow the mandatory directive issued by the government and have the opportunity to engage in voluntary assignments. According to Statens författningssamling 2017:725 (SFS, 2017), a member of a municipality is defined as follows.

5 § '*A member of a municipality is the one who is registered in the municipality, owns real estate in the municipality or is to pay municipal tax there'*.

'A member of a municipality is also a citizen of one of the Member States of the European Union (Union citizen) who resides in the municipality but who, according to § 5, second paragraph, the Public Registration Act (1991)(481), shall not be registered there'.

The task of a municipality is to carry out assignments for its citizens. Laws, regulations and guidelines control the assignments, which can be complemented by local regulations such as e.g politically targeted investments. The political organisation of a municipality is governed by the national law SFS 2017:725 Kommunallag (SFS, 2017); however, each municipality is free to create a structure and working methods for performing its internal tasks (Sveriges Kommuner och Landsting, 2019). A municipality is a complex system with many stakeholders that needs to be dynamic, adapt to needs and interact with many other systems in society. The constituent parts of a municipality need to be coordinated and followed upon within the municipal structure; they then form the basis for continued work.

Municipalities are part of the public sector, and funding is based on taxes. According to Lee, Hwang and Choi (2012), drivers in the public sector are increasing customer value for target groups and reducing costs for taxpayers. Moore (1995) clarifies that public value is understood as the achievement of preferred outcomes by using public resources in the most effective manner available. Result-based management is a common practice in the public sector and occurs at all levels, from local, regional, and national to supranational (Van Thiel and Leeuw, 2002). Researchers Van Thiel and Leeuw (2002) further reflect on the relations between performance measurement indicators, analyses and the room for interpretation in different respects. Moreover, the authors describe that most products in the public sector are intangible. Performance indicators should reflect quality and reliability rather than 'hard' product attributes: *'Public services are not only about efficiency and effectiveness but also about justice, fairness, equity, and accountability'* (ibid.).

This case study focuses on co-creation as an important factor in the development process to increase citizen value when the management in a municipality analyses and evaluates performance. Committed co-workers and managers are a prerequisite for organisational performance outcomes. If the values are also shared and if appropriate methodologies and tools are applied, the conditions for a successful result are improved. Because all organisations are unique, the adoption of organisational values needs to be customised. Achieving this co-creation can be a way to make co-workers committed. Moreover, according to Seddon and Caulkin (2007), adopting a new way of thinking through action is easier than inducing new and counterintuitive ways of acting by thinking about the ways. Co-creation can also be a way to induce a breakthrough during changes in an organisation. Senge (1999) describes that people do not resist change but resist being changed.

The purpose of this paper is to describe the process of co-creating a dialogue model that aims to increase citizen value in a municipality organisation. In addition, the purpose is to present the results from the development process and evaluation of the chosen dialogue model.

2 METHODOLOGY

The municipality approached Mid Sweden University to obtain input after it identified a need to improve its ways of working to achieve a higher performance. The municipality is located in the middle of Sweden and serves approximately 100,000 citizens. In the autumn of 2017, the council established guidelines for a new model for controlling and managing the municipality that increased the focus on performance-based management. In short, these guidelines imply that control measures must be based on quality analyses of actual, achieved results among citizens/users and that the municipality desires to achieve better performance-based management through simplification, clarity and quality analyses. The municipality wanted to find a way of working that would focus on

creating citizen value and identifying dialogues about the results as an important tool.

A joint pre-study project started on 01 September 2018 and ended on 01 March 2019 for the purposes of developing a new method of working with performance analysis and dialogues. The project team had five members–two from the municipality and three from the University. The work was divided into two parts; Part 1 was to develop the dialogue model and consisted of two steps (*Step One* and *Step Two*), and Part 2 consisted of one step (*Step Three*) was to evaluate the use of the developed dialogue model.

To obtain an understanding of the current situation in the municipality, the project team started Part 1 by conducting interviews with the leaders, both elected representatives and officials. The project team created an interview guide that focused on the respondents describing the existing management model and their experiences using it. The project team members from the municipality chose a total of six relevant respondents. Members from the University conducted the interviews over Skype or the telephone. During the interviews, notes were taken, which were subsequently sent to each respondent for verification. In two cases, the notes were corrected.

In parallel to the interviews, the development of the dialogue model was initiated. In *Step One*, input for developing a prototype model for the dialogue was collected. This was done through a workshop, with the purpose of identifying the success factors from fruitful and constructive dialogues based on the effects on and the results of citizens–a total of 30 officials participated. Based on the results from the interviews with the leaders and the first workshop, *Step Two* in the development process started with a number of dialogue model prototypes were developed. These dialogue models were then tested and evaluated in a second workshop, where 21 officials and seven elected representatives participated. The results from the second workshop led to the decision to choose one dialogue model. Both workshops were facilitated by the researchers on the project.

Part 2 included testing and evaluating the final dialogue model. In the first round, the dialogue model was used in a total of six dialogues between the councillor and the municipal administration from different boards. Subsequently, six participating persons–two elected representatives and four officials–answered a written evaluation. In addition to the written evaluation, the project members from the municipality conducted an oral evaluation with nine participants, one elected representative and eight officials, who were partly the same participants who answered the written evaluation. During the oral evaluation, notes were taken by the project members from the municipality. Thus far, the dialogue model was used twice with the same configuration as in the first meeting, and after each round, an oral evaluation was performed by the project members from the municipality.

Even starting the project required that all of its parts were to be completed through the co-operation of the team members from the municipality and the University, because the co-operations was a prerequisite for the funding of the project. Throughout the project, dialogues continuously occurred within the project team to conduct analyses and evaluations and to make decisions.

3 RESULTS

This section is divided into two parts. Part 1 presents the development process of the dialogue model (*Step One*) and the testing of the prototypes (*Step Two*). Part 2 presents the evaluation of the dialogue model used in real life for the first time and a summary of the model's continued use (*Step Three*).

3.1 Part 1 – Development and Testing of the Model for Dialogue

According to the plan, two workshops with officials and elected representatives from the municipality were conducted. The purpose was to obtain as good input as possible for the actual performance-based dialogue models as well as anchoring the work in the organisation.

Step One

For the first workshop, the invitation was sent to the administration managers and the financial manager and to the internal quality network. The purpose of this workshop was to identify the success factors based on fruitful and constructive dialogues on the effects on and results of citizens. The purpose was also to identify concrete proposals for how the success factors could be realised. The identified success factors and proposals for the concrete implementation of the factors are presented in Table 1. The workshop started with one-on-one dialogues, and participants used an interview guide inspired by appreciative inquiry with questions that aimed to find the strengths and success factors. When an interview was completed, the pair of participants that performed the interview met another pair. Together, their mission was to find and agree on the five most inspiring success factors from the interviews and write them down on Post-it Notes. Each group of two pairs presented their results from the practice to the entire group. Subsequently, all Post-it Notes were rearranged to find patterns and make groups of related Post-its. The new Post-it groups received a descriptive heading. The next exercise required the participants to choose a heading and start a group task to generate an idea around the question: What concrete ideas do I want to send to the work group that is developing the prototypes? During this exercise, the participants had the opportunity to change groups and contribute to several headings.

Table 1 – Identified Success Factors and Proposals for the Concrete Implementation of the Factors from the First Workshop in the Development Process

Category	Prerequisites	Approach	Structure and process	Goal/Why?
Underlying concepts	Self-awareness	Be curious	Documentation	Common challenges
	Commitment	Be open	Moderator	Clear purpose
	Why─Creates commitment	Be listening	Facilitator	Solution-oriented
	Participation	Listen	Time for preparation	Common goal
	Understanding	Listen─understand the context, be humble	Enough time for dialogue	Clarity
	Understanding of different perspectives	Think outside the box─have courage	Appropriate "vocabulary"	Who? Participant, target group, customer
		Be respectful	Common basis	
			Structure	
			Systematics	
			Using tools for common understanding	
			Solid preparation and perseverance	
			Meeting in person	

The results from the first workshop, together with the results from the interviews, became the input for the project team's work to generate a number of model prototypes for a new way of carrying out performance-based dialogues. In total, three different prototypes for performance-based dialogues were generated and tested in two different forms during workshop two. The prototypes tested in workshop two were as follows: traditional meeting with an agenda, structured A3[1] and a dialogue canvas[2]. The two different forms in which each dialogue model was tested were traditional questions and questions with strength-based[3] elements. In addition to the dialogue models, three fictitious cases were

[1] A3 refers to the internationally agreed-upon paper size, which is viewed as a common, logical way of thinking; this design can be adapted to different situations. A3 connects a problem with cases, goals, and proposed measures to reach the target and provide the resources to measure success (Shook, 2011).
[2] In this case, the dialogue canvas is the A1 paper format with pre-printed fields with questions and a space to write down documentation. All participants are encouraged to write on the canvas during the dialogue.
[3] The strength-based questions are inspired by the appreciative inquiry approach (AI); see for instance Whitney and Trosten-Bloom (2010).

developed to focus on the process and not on real-life challenges. The three fictitious cases were Elementary School, Home Care and Origo Park. A general description of the development of Part 1 is visualised in Figure 1.

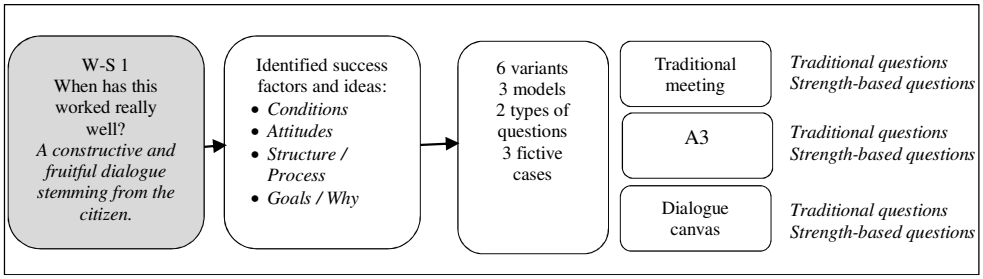

Figure 1 – General Description of Step One in the Overall Development of the Model for Performance-based Dialogues

Step Two

Before the second workshop, the invitation was sent to everyone who participated or wanted to participate in the first workshop and to elected representatives on the municipal board finance committees (ordinary and replacement). On this occasion, the participants were grouped, and each group had the opportunity to test two prototypes during the workshop. The different prototypes were tested according to a schedule, see Table 2. Before starting a test, each group received information and instructions from the observer, and if the prototype contained roles, the roles were distributed amongst the participants in the group. Each test was timed and had a specific time limit. After each test, participants answered an evaluation form with questions on whether everyone had the opportunity to speak and be listened to, how they experienced the documentation, whether a supporting structure existed, and what worked well and what worked less well.

Table 2 – Schedule of Tests of the Different Performance-based Dialogue Prototypes

Type of dialogue model	A3 (am)	A3 (pm)	Dialogue canvas (am)	Dialogue canvas (pm)	Agenda (am)	Agenda (pm)
Observer	Researcher A		Researcher B		Researcher C	
Case	Elementary school	Elementary school	Home care	Home care	Origo Park	Origo Park
Type of questions	Ordinary	Strenght-based	Ordinary	Strenght-based	Ordinary	Strenght-based
Group	1	3	2	1	3	2

When all prototypes were tested, the groups evaluated each prototype from four different perspectives: citizen, co-worker, process and business, and financials. The group had to split 100% of their meeting time, as they experienced how they focused on the parts. The participants also had to fill in the 'other' opportunity. At the end of the workshop, the participants worked in pairs and brainstormed around three themes: 'Needs to be!', 'Definitely not!', and 'Maybe!'. This session aimed to collect the participants' overall perceptions of the prototypes.

After workshop 2, all evaluations were analysed, which resulted in a dialogue model that starts with sending out preparatory materials to participants. The dialogue model also consists of meeting rules, a documentation template, a clear division of roles, and a focus on citizen values and having a holistic view. The prototype that best matched all of these criteria was an A3 model with strength-based inspired questions. Subsequently, this dialogue model was jointly discussed with the project group, the chairperson of the municipal board and the municipal director. The output from the discussion led to a further adaptation of the dialogue model before testing the model in real life.

3.2 Part 2 – Evaluation of the Performance-based Dialogue Model

Step Three

At this time, the developed model for performance-based dialogues was ready to be tested in real life. In this part, the dialogue model was evaluated through both written and oral evaluations with the participants.

The results from the written evaluation indicate that, in general, all respondents were positive about using the dialogue model. Four respondents found that the dialogue model contributed to an increased understanding and clarification of expectations and results. One respondent stated that it was difficult to compare because he or she was new to the role. Another respondent expressed that the large size of the business made it difficult to state that the dialogue model fully contributed to understanding and clarifying expectations and results; however, the dialogue model was better than the previous way of working. All respondents thought that all participants had the opportunity to speak and that they participated and were committed during the meeting. Five respondents also thought that all participants' opinions and views were taken into account, and one respondent did not answer that question. The opinions on the documentation were divided. Three respondents were positive, one respondent found it to be too short but thought that the documentation could still support the organisation, and two respondents did not like the format, or they had not seen it yet. The allocated meeting time was not a prioritised issue, according to the respondents.

In addition to these questions, the respondents submitted comments on what 'worked well' and 'what worked less well'. Regarding 'what worked well', there were some examples of comments regarding the area: a good structure that supported speaking about continuous improvement, good atmosphere, suitable group size, good discussion about both business and financial results, and a good

way of discussing quality and business. On the list of 'what worked less well' were the following: too much time was set aside, the documentation that was sent out before the meeting needed to be developed in terms of scope and content, there was a risk that the participants were too careful and nice during the meeting, and these kinds of dialogues are also needed between different administration parts.

The outcome from the oral evaluations with participants after the performance-based dialogues were completed was also analysed according to 'what worked well' and 'what worked less well'. Additionally, a third field, 'other', was added. Regarding 'what worked well' the comments were as follows: a good atmosphere during the dialogues and the A3 had a good structure–for instance, everyone around the table was able to talk–and the focus of the questions. Respondents also highlighted that 'what worked well' were the outcomes from the dialogue and that the focus was not only on the financials. One respondent expressed that participating in the performance-based dialogue was a learning experience. Regarding 'what worked less well', there were comments about the atmosphere, which was too kind and not challenging enough. The 'what worked less well' comments also included a dialogue model that was too structured, a request for more open conversations and the need to further develop preparatory material. Regarding 'other', there were reflections on the feedback, such as the following: 'What did the municipal board, municipal director and financial manager think?' 'What happens to the assignments that ended up at the municipal executive's office?' 'Why did the HR director who participated not participate?' Issues listed as 'other' included the number of participants who had influence, which was not too many and not too few.

During 2019, the dialogue model was used two more times in the same configuration. The purpose was the same, but the theme was adjusted to meet the internal business needs. The second time the dialogue model was used, the theme was economic deviation, and the third time the theme addressed next year's operational planning. The dialogue model was supplemented with an added point during the actual meeting; each meeting started with a review and follow-up of the decisions and measures that were the output of the previous meeting. Until this point, the evaluations showed that the outcome from using the dialogue model was appreciated by the participants. The highest management had increased knowledge about the organisation, and constructive actions were taken. The format of the model, to set aside time for recurring meetings, was highlighted. This way of working also contributed to continuity and provided a clearer way of addressing the assignments of the administrations and committees. A fourth meeting using the dialogue model is already booked, which has as its theme the results of the next year, and the plan is to continue with four meetings on a yearly basis. During 2020, the meetings will also be extended to include the municipal corporations.

4 ANALYSIS AND DISCUSSION

The purpose of this paper was *to describe the process of co-creating a dialogue model that aimed to increase citizen value in a municipality organisation. In addition, the purpose was to present the results from the development process and an evaluation of the chosen dialogue model.*

A dialogue model was developed through a co-creation process with the intention of the model being a useful methodology for the municipality board and management to use in their work on performance-based management. Co-creation was carried out in more than one way. The first way was to create conditions for the development process that enabled elected representatives, managers and co-workers to participate. The second way to engender a co-operation that became a co-creation among the members on the project team, who were from the municipality and the University.

According to Leavy (2012) and Lee, Olson and Trimi (2012), interaction that occurs through co-creation can create prerequisites for identifying the needs and benefits from services and goals. During the entire development process, the focus on citizen value was the main goal. In the first workshop, the focus was on identifying the success factors from fruitful and constructive dialogues that were based on the effects on and results of the citizen, which can be concretely implemented in quality management through the value 'focus on customers' from the cornerstone model of Bergman and Klefsjö (2010). In the second workshop, during prototype testing, the participants contributed to the development of the dialogue model by evaluating the citizen value when using each prototype. The participants also contributed by categorising different topics from the tests into 'Needs to be!', 'Definitely not!', and 'Maybe!' categories. The evaluations from using the dialogue model in real life also supported the concept that citizen value was considered. In the evaluation form, the fact that the participants in the dialogue were able to talk about quality and business was seen as a benefit, given that both are connected to creating value for the citizen. The dialogues indicated that the participants were able to focus on questions in the developed dialogue model, the A3, which were oriented towards value creation.

Co-creation has also been shown as fruitful in relation to the quality management values presented in Bergman and Klefsjö (2010): 'let everybody be committed' and 'management commitment'. By inviting co-workers and managers to the workshops and allowing them to contribute knowledge and comments, conditions for commitment were created. Inviting and involving committed co-workers followed the model proposed by Ingelsson, Eriksson and Lilja (2012), in which the step of involving employees who support the organisation's values added to the chain between valuation and choosing the right methodologies and tools to fulfil the desired result.

By implementing a new model for control and management with an increased focus on performance-based management in the municipality, the council aimed for changes in ways of acting. Seddon and Caulkin (2007) express that it is easier

to adopt a new way of thinking through action than by thinking about the new and unknown. The way that the dialogue model was developed can be seen as an example of acting to find new and counterintuitive ways of acting. Co-creation through the involvement of co-workers during the development of the dialogue model can facilitate change: as Senge (1999) describes, people do not resist change but resist being changed.

The developed dialogue model was used in dialogues by the municipality's highest management and by different administration parts of the organisation. The evaluations performed after the real-life test indicate that the dialogue model created conditions in the organisation for listening to each other and that the model increased the understanding of different needs by clarifying expectations and results. The developed dialogue model was found to have similarities with the method of 'knowledge overlapping seminars' presented by Cronemyr (2007), since the evaluations presented results in line with creating a common view and increased knowledge among the participants.

The municipality has a defined goal for its development of performance-based management; however, whether the municipality has a shared organisational culture and shared values among co-workers is unknown. By involving co-workers in the development process of the methods and tools, the prerequisites were created to allow their values to appear in the method that was developed. In the next step, the co-workers' participation may have effects on organisational performance, because the values are connected to organisational culture, which is connected to organisational performance (Chatman and Eunyoung Cha, 2003; Grönfeldt and Strother, 2006). Another possible effect of allowing co-workers to participate in the development process is that it could create conditions for health among them (Bäckström, 2009; Lagrosen, Bäckström and Wiklund, 2012). Asif et al. (2009) suggest that co-workers should be involved when quality management is implemented and that they should assist in the adoption of the unique context. This organisation has not expressed that they are implementing quality management, but the desired results are in line with quality management outcomes. However, it can be assumed that the co-workers have made it possible to develop the dialogue model for adaptation to their own context.

Gutiérrez-García, Recalde and Piñera-Camacho (2015) describe two perspectives on having dialogues in the public sector: creating an opportunity to involve stakeholders and supporting a management culture. On the management side, dialogues create conditions for meeting the various expectations of a wide range of individuals and groups. In the presented development process and the adaptation of the dialogue model, only one of these sides can be said to have been present, creating conditions for meeting the various expectations of a wide range of individuals and groups. This situation exists because at this stage, the dialogue model has been implemented only in the internal process of performance-based management but with a customer focus. The evaluations indicate that conditions to satisfy and create value for the internal customers occurred in the development process and by using the dialogue model. Internal

customer value creation increases the conditions that maximise the value for external customers (Carlsson and Wilmot, 2006).

5 CONCLUSION

The development process in this study can be used in other situations and organisations. The process is not limited to creating dialogue models in performance-based management because the co-creation process can be used for other key topics. The results from this study indicate that the co-creation process is suitable when an organisation wants to involve co-workers to improve their knowledge, contextualise a specific issue and develop methodologies for internal use. In this study, an internally developed dialogue model fulfilled the desired goal of finding a way of working that focuses on creating citizen value.

The results indicate that questions raised in the developed dialogue model (A3) helped the participants focus on value creation. Therefore, such a model can guide the dialogue towards desired areas and create a dialogue whereby the participants listen to each other more. The dialogue model also created a new way of acting for management. The results also indicate that the developed dialogue model created an opportunity for the municipality board to adopt a holistic view.

Areas for future research include following up on whether the dialogue model supports the long-term view of the agreed-upon conclusion and the activities in each performed dialogue and verifying the effects on created value for citizens from the citizens' perspective. The dialogue model may also need to be further developed so that all participants experience its use as effective, given the feedback on the time spent, and so that they can feel free to express themselves regardless of their opinions and challenge the prevailing norms and structures.

REFERENCES

Asif, M., Joost de Bruijn, E., Douglas, A. and Fisscher, O., 2009. Why quality management programs fail: A strategic and operations management perspective. *International Journal of Quality & Reliability Management*, [e-journal] 26(8), pp.778-794. DOI: 10.1108/02656710910984165.

Bäckström, I., 2009. *On the relationship between sustainable health and quality management: leadership and organizational behaviours from Swedish organizations*. PhD. Östersund, Mid Sweden University.

Bergman, B. and Klefsjö, B., 2010. *Quality: From customer needs to customer satisfaction*. Lund: Studentlitteratur.

Campbell, C.R., 2004. Longitudinal Study of One Organization's Culture: Do Values Endure?. *American Journal of Business*, [e-journal] 19(2), pp.41-52. DOI: 10.1108/19355181200400011.

Carlson, C.R. and Wilmot, W.W., 2006. *Innovation the Five Disciplines for Creating what Customers Want*. New York: Crown Publishing Group.

Carmeli, A., Gelbard, R. and Reitner-Palmon, R., 2013. Problem-solvning capacity, and creative performance: The importance of knowledge sharing. *Human Resource Management*, 52(1), pp.95-122.

Chatman, J.A. and Eunyoung Cha, S., 2003. Leading by Leveraging Culture. *California Management Review*, 45(4), pp.20-34.

Cooperrider, D. and Whitney, D., 2005. *Appreciative Inquiry: A positive revolution in change*. San Francisco: Berrett-Koehler Publishers.

Cronemyr, P., 2007. *Six Sigma Management: action research with some contributions to theories and methods*. PhD. Chalmers University of Technology. Available at: <http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-100856> [Accessed 08 July 2020].

Dahlgaard, J., Kristensen, K. and Kanji, G., 2002. *Fundamentals of Total Quality Management*. Cheltenham: Nelson Thornes.

Flynn, B.B., Schroeder, R.G. and Sakakibara, S., 1994. A framework for quality management research and an associated measurement instrument. *Journal of Operations management*, 11(4), pp.339-366.

Gassmann, O. and Enkel, E., 2004. Towards a theory of open innovation: Three core process archetypes. In: RADMA (Research and Developlent Management Association), *Proceedings of the R&D Management Conference.* Lissabon.

Grönfeldt, S. and Strother, J., 2006. *Service leadership: the quest for competitive advantage*. Thousand Oaks, Calif.: SAGE Publications.

Gutiérrez-García, E., Recalde, M. and Piñera-Camacho, A., 2015. Reinventing the wheel? A comparative overview of the concept of dialogue. *Public Relations Review*, [e-journal] 41(5), pp.744-753. DOI: 10.1016/j.pubrev.2015.06.006.

Hamilton, M.L. and Pinnegar, S., 2015. *Knowing, Becoming, doing as Teacher Educators: Identity, Intimate Scholarship, Inquiry*. Emerald Publishing Group Limited.

Hammer, M., 1996. *Beyond reengineering: how the process-centred organization is changing our work and our lives*. New York: Harper business.

Hellsten, U. and Klefsjö, B., 2000. TQM as a management system consisting of values, techniques and tools. *The TQM Magazine*, [e-journal] 12(4), pp.238-244. DOI: 10.1108/09544780010325822.

Ingelsson, P., Eriksson, M. and Lilja, J., 2012. Can selecting the right values help TQM implementation? A case study about organisational homogeneity at the Walt Disney Company. *Total Quality Management*, [e-journal] 23(1), pp.1-11. DOI: 10.1080/14783363.2011.637801.

Kristensson, P., Gustafsson, A. and Witell, L., 2014. *Tjänsteinnovation*. Lund: Studentlitteratur.

Lagrosen, Y., Bäckström, I. and Wiklund, H., 2012. Approach for measuring health-related quality management. *The TQM Journal*, [e-journal] 24(1), pp.59-71. DOI: 10.1108/17542731211191221.

Leavy, B., 2012. Collaborative innovation as the new imperative – design thinking, value co-creation and the power if "pull". *Strategy & Leadership*, [e-journal] 40(2), pp.25-34. DOI: 10.1108/10878571211209323.

Lee, S., Hwang, T. and Choi, D., 2012. Open innovation in the public sector of leading countries. *Management Decision*, [e-journal] 50(1), pp.147-162. DOI: 10.1108/00251741211194921.

Lee, S., Olson, D. and Trimi, S., 2012. Co-innovation: Convergenomics, collaboration, and co-creation for organizational values. *Management Decision*, [e-journal] 50(5), pp.817-831. DOI: 10.1108/00251741211227528.

Moore, M.H., 1995. *Creating public value: Strategic Management in Government*. Cambridge Massachusetts: Harvard University Press.

Oh, H., 1999. Service quality, customer satisfaction, and customer value: A holistic perspective. *International Journal of Hospitality Management*, [e-journal] 18(1), pp.67-82. DOI: 10.1016/S0278-4319(98)00047-4.

Seddon, J. and Caulkin, S., 2007. Systems Thinking, Lean Production and Action Learning. *Action Learning: Research and Practice*, 4(1), pp.9-24. DOI: 10.1080/14767330701231438.

Senge, P., 1999. *The dance of change: The challenge to sustaining momentum in learning organizations*. New York: Doubleday.

SFS (Statens Författningssamling), 2017. *Kommunallag (2017:725)*. Finansdepartementet K: SFS. Available at: <https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/kommunallag-2017725_sfs-2017-725> [Accessed 20 May 2019].

Shook, J., 2011. *Lean management: med hjälp av A3-analyses*. Malmö: Liber.

Sveriges Kommuner och Landsting, 2019. *Så styrs en kommun.* Bruxelles: Swedish Association of Local Authorities and Regions. Available at: <https://skl.se/demokratiledningstyrning/politiskstyrningfortroendevalda/kommunaltsjalvstyresastyrskommunenochregionen/sastyrskommunen.735.html> [Accessed 20 May 2019].

Taylor, F.W., 1998. *The principles of scientific management*. Canada, Toronto: General publishing company Ltd.

Van Thiel, S. and Leeuw, F., 2002. The performance paradox in the public sector. *Public Performance and Management Review*, [e-journal] 25(3), pp.267-281. DOI: 10.1080/15309576.2002.11643661.

Watkins, J.M. and Cooperrider, D., 2000. Appreciative inquiry: a transformative paradigm. *OD Practitioner,* 32(1), pp.6-12.

Westher, M., 2018. *Organisationers främsta utmaningar: En studie med fem års horisont*. Göteborg: SIQ – Institutet för Kvalitetsutveckling.

Whitney, D. and Trosten-Bloom A., 2010. *The power of Appreciative Inquiry*. San Francisco: Berrett-Koehler Publishers.

Wu, S., Zhang, D. and Schroeder, R., 2011. Customization of quality practices: the impact of quality culture. *International Journal of Quality & Reliability Management*, [e-journal] 28(3), pp.263-279. DOI: 10.1108/02656711111109883.

ABOUT AUTHORS

Anna Mårtensson – (A.M.) Mid Sweden University, Östersund, Sweden, Department of Quality Technology and Management, Ph.D. Student in Quality Management, e-mail: anna.martensson@miun.se, Author's ORCID: 0000-0001-8609-6290.

Pernilla Ingelsson – (P.I.) Mid Sweden University, Östersund, Sweden, Department of Quality Technology and Management, Assoc. Prof. in Quality Management, e-mail: pernilla.ingelsson@miun.se, Author's ORCID: 0000-0001-5610-2944.

Ingela Bäckström – (I.B.) Mid Sweden University, Östersund, Sweden, Department of Quality Technology and Management, Prof. in Quality Management, e-mail: ingela.backstrom@miun.se, Author's ORCID: 0000-0001-7621-2649.

AUTHOR CONTRIBUTIONS

A.M. – writing—original draft; I.B. – project administration; A.M., P.I. – visualization; I.B., P.I. – funding acquisition; A.M., P.I., I.B. – conceptualization, methodology, validation, formal analysis, investigation, writing—review and editing.

CONFLICTS OF INTEREST

The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Methodological Assessment of Data Suitability for Defect Prediction

DOI: 10.12776/QIP.V24I2.1443

Peter Schlegel, Daniel Buschmann, Max Ellerich, Robert H. Schmitt

Received: 2020-04-03 Accepted: 2020-06-29 Published: 2020-07-31

ABSTRACT

Purpose: This paper provides a domain specific concept to assess data suitability of various data sources along the production chain for defect prediction.

Methodology/Approach: A seven-phase methodology is developed in which the data suitability for defect prediction in interlinked production steps is assessed. For this purpose, the manufacturing process is mapped and potential influencing variables on the origin of defects are identified. The available data is evaluated and quantified with regard to the criteria relevancy, completeness, appropriate amount of data, accessibility and interpretability. The individual assessments are then visualized in an overview, gaps in data acquisition are identified and needs for action are derived.

Findings: The research shows a seven-phase methodology to systematically assess data suitability for defect prediction and identify data gaps in interlinked production steps.

Research Limitation/implication: This research is limited to the analysis of contextual data quality for the use case of defect prediction. Other data analytics applications or processes outside of manufacturing are not included.

Originality/Value of paper: The paper provides a new approach to identify gaps in data acquisition by systematically assessing data suitability for defect prediction and deducting needs for action. The accuracy of predictive defect models is then to be improved by the subsequent optimization of the data basis.

Category: Research paper

Keywords: predictive quality; defect prediction; failure prediction; data suitability; data quality

1 INTRODUCTION

The implementation of Industry 4.0 shapes the competition of manufacturing companies on global markets (Bal and Erkan, 2019). Those who want to stay ahead of their competition for the long term are required not only to record all available information securely and in real time, but also to process it in order to be able to analyze it precisely and continuously (Uhlemann et al., 2017). Having the right information available at the right time is an enormous challenge. Therefore, it is indispensable to recognize patterns in the data stream, to learn from them and to be able to derive the right predictions for the company, the processes and the products (Brecher et al., 2017). This applies in particular to the prediction of product defects in interlinked production steps (Eger et al., 2018). In a data-centric view of the entire production chain, there is great potential for the predictive identification of defect causes, the derivation of suitable measures and thus the reduction of defect costs. Without continuous data acquisition, it is not possible to trace correlations resulting from workpiece handling in different process steps (Ghimire et al., 2015). Classical tools such as Statistical Process Control mostly consider individual process steps and are therefore unsuitable for this application (Škulj et al., 2013).

The vision of the Internet of Production (IoP) describes a real-time, secure availability of information at any time and any place (Brecher et al., 2017; Pennekamp et al., 2019). Precise and continuous data analysis, pattern recognition for prediction and, based on this, reliable decision-making should support production systematically and sustainably.

The IoP infrastructure shown in Figure 1 consists of four underlying layers (Brecher et al., 2017):

- The *Raw Data* layer as well as the raw data access via the respective application software.
- A *Middleware+* for the administration of the data access on different proprietary systems.
- The *Smart Data* layer for the generation of knowledge based on the *Digital Shadow* and the *Smart Expert* layer on which the domain-specific usage of the aggregated knowledge takes place.

The term *Digital Shadow* is defined as the sufficiently precise representation of the processes in production, development and adjoining areas with the purpose of creating a real-time-capable evaluation basis for all relevant data (Bauernhansl et al., 2016). The relevant representation refers specifically to a smaller scope of data than that contained in the raw data, since only data relevant to the application case are passed on (Brecher et al., 2017).

In order to implement models for predicting product defects, a systematic selection of the relevant data in the sense of the *Digital Shadow* must first take place. For this systematic selection, manufacturing companies currently lack

knowledge regarding their own data quality (DQ) and the suitability of collected data for this specific application (Schuh et al., 2019). Previous research work in this context has mainly focused on general description models and metrics for the evaluation of DQ and not on the context-related development of methods for data evaluation (Wang and Strong, 1996; Batini et al., 2009; Zaveri et al., 2015).

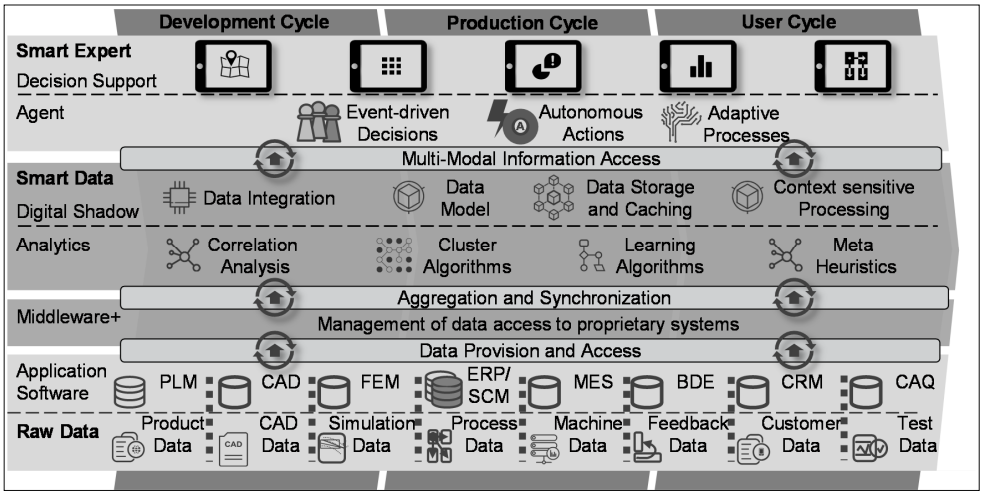

Figure 1 – Infrastructure of the Internet of Production (Source: Based on Brecher et al. (2017))

2 RELATED WORK

Based on the deficits mentioned, a systematic evaluation of the data suitability for the application case of defect prediction is carried out within the framework of the methodology presented. The main objective of the developed methodology is the creation of transparency regarding data suitability of existing data sources, the identification of data gaps and the derivation of need for action.

2.1 Implementation of Defect Prediction

In literature, a variety of data-driven methods and strategies for the implementation of defect prediction in manufacturing can be found. Eger et al. (2018) describe the *ForZDM methodology*, which expands single process boundaries of pre-existing Zero Defect Manufacturing (ZDM) approaches towards a production line perspective. This makes it possible to contrast and counter defects before, during and after their emergence through the integration of multi-level system modelling, big data analysis, Cyber Physical Systems and real-time data management (Eger et al., 2018). Wang (2013) presents a general framework of ZDM and explains how to apply Data Mining approaches to manufacture the products with zero-defect. The developed framework has a modular structure and consists among others of the main components fault prognosis, fault diagnosis as well as the subsequent correction and compensation

(Wang, 2013). Lieber et al. (2013) describe a methodical framework based on data mining for predicting the physical quality of intermediate products in interlinked manufacturing processes in the context of a rolling mill case study. Other approaches to defect prediction also do not include specific modules for they systematic assessment of the available database (Arif, Suryana and Hussin, 2013; Kao et al., 2017; Wuest, Irgens and Thoben, 2013; Schmitt and Deuse, 2018).

From literature, it becomes clear that the implementation of methods for defect prediction is often addressed in the context of ZDM strategies. Implications for the quality of the data basis, which result from the examination of interlinked production steps, are considered little or not at all. The presented frameworks contain modules for data preparation and feature extraction, but no methodology for the context-related evaluation of the data basis.

2.2 Assessment of Data Quality

As mentioned in the introduction, a wide range of general methods and techniques already exist for the assessment of DQ. The earliest work in the area of DQ assessment is published by Wang and Strong (1996). In their work, they present a conceptual framework of DQ including 15 dimensions within four categories. Batini et al. (2009) provide a systematic and comparative description of methodologies that help the selection, customization, and application of DQ assessment. Cai and Zhu (2015) analyze the data characteristics of the big data environment, present quality challenges faced by big data, and formulate a hierarchical DQ framework from the perspective of data users. Zaveri et al. (2015) conducted a comprehensive survey on the assessment of linked DQ and identified 16 quality dimensions that have been studied in the literature. The work unifies and formalizes commonly used terminologies across papers related to DQ and provides a comprehensive list of 18 quality dimensions and 69 metrics (Zaveri et al., 2015). These dimensions were classified into four categories: (i) Accessibility, (ii) Intrinsic, (iii) Contextual, and (iv) Representational. Gürdür, El-khoury and Nyberg (2018) base their analysis on the work of Zaveri et al. (2015) and present a study that explains and applies a DQ assessment methodology as a post-integration phase for linked enterprise data. The authors examine a case study from the automotive industry using the linked enterprise data approach to integrate data from different development tools (Gürdür, El-khoury and Nyberg, 2018). Ardagna et al. (2018) propose a methodology to build a DQ adapter module, which selects the best configuration for a context-aware DQ assessment based on the user main requirements: time minimization, confidence maximization, and budget minimization.

The literature contains a large number of general methods and metrics for evaluating DQ. However, there is no methodology for evaluating the database in the context of defect prediction. Regardless of the formal quality of the data for the application of suitable prediction models, there is uncertainty in many

companies as to whether the information required for defect prediction is available in digital form at all. From this, the need for research is derived to develop a methodology to assess the suitability of the existing database and to identify gaps in data acquisition. The research hypothesis derived from this is:

H: The accuracy of a model for defect prediction can be improved by a systematic assessment of data suitability and the subsequent elimination of the identified data gaps.

3 PREDICTIVE MODELING OF PRODUCT DEFECTS

Predictive modeling is applied to a range of techniques that find relationships between a target variable and the other variables in the data set (Wang, 2013). For example, such functions could be classification, value prediction or association rules (Raudys, 2001; Backhaus et al., 2016). In the given context, the target variable is a specific product defect that is identified by quality control in production. The other influencing variables include the parameters of individual process steps in the production chain as well as recorded sensor and meta data of disturbance and environmental variables. The objective of defect prediction is to predict the occurrence of certain defect types based on process parameters, sensor and meta data of the manufacturing process.

Usually, all available data sources are included in the database and then informative and non-redundant features are extracted (cf. Figure 2). Feature extraction converts raw data into informative features that efficiently represent the information relevant for analysis (Rawat and Khemchandani, 2017). The more accurately the features represent the data set, the faster and more accurately the predictive model will work (Kacprzyk et al., 2006). The features created are iteratively tested and optimized during the modeling process, which allows the dimension of the input data to be reduced (Liu and Motoda, 2008). For a successful feature extraction a suitable pre-processing of the data as well as a high contextual and formal DQ of the raw data is necessary. While the evaluation of formal DQ (e.g. accuracy, consistency etc.) is excluded from this research, the objective is to evaluate the context-related data suitability with a special focus on the degree of information content of the data.

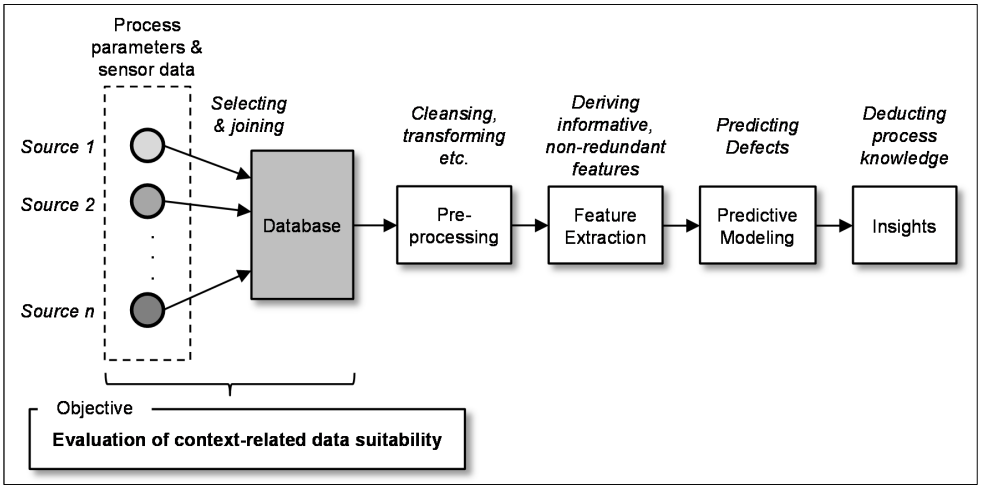

Figure 2 – Localization of the Research Objective in the Predictive Modelling Process

4 METHODOLOGY

DQ is usually understood as a multi-dimensional concept. The dimensions represent the views, criteria, or measurement attributes for DQ problems that can be assessed, interpreted, and possibly improved individually. By assigning scores to these dimensions, the overall DQ can be determined as an aggregated value of individual dimensions relevant in the given application context.

The basis for the developed methodology for evaluating the context-related data suitability for defect prediction is a slightly modified framework for evaluating data quality according to Wang and Strong (1996), which defines 15 DQ dimensions in four categories. Of the 15 dimensions depicted in Figure 3, the following five are addressed within the methodology of this work: *relevancy, completeness, appropriate amount of data, accessibility* and *interpretability*. The two criteria *value-adding* and *timeliness* will not be considered further in the context of the developed methodology for the time being. Whether the available data or the resulting features actually add value can only be determined after modeling. The criterion timeliness and therefore the requirements for the speed with which the data is available depends strongly on the real-time requirements of the respective predictive model and thus on the application case in the production. Therefore, no general evaluation can be made. Since this work is a purely context-related evaluation of DQ, the more accurate term *data suitability* instead of data quality is used from here on.

This methodological approach provides companies with guidance on how to gain insight into the information content of the data collected and its suitability for defect diagnosis and prediction. The need for increased transparency and understanding of the data basis is due to the fact that initiated data analyses in

practical industrial utilization often bring only limited gain in knowledge. The methodical approach describes a step-by-step process that aids stakeholders in discussing, defining and identifying gaps in data collection through systematic uncovering and visualization.

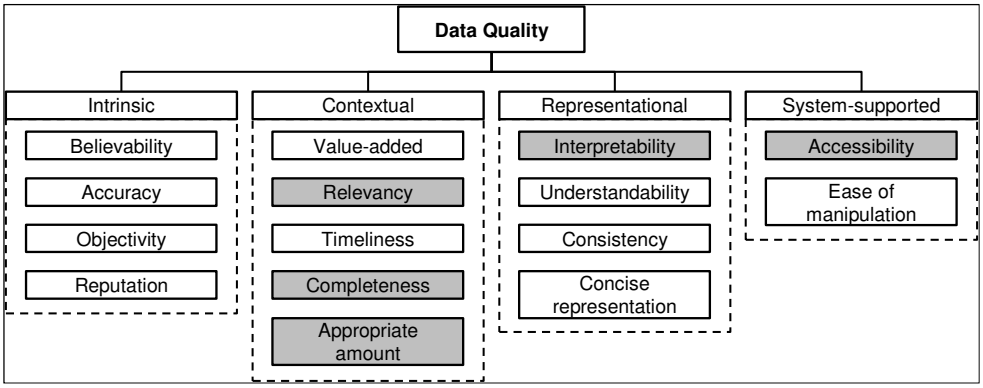

Figure 3 – Framework of 15 Quality Dimensions (Source: Based on Wang and Strong (1996), Hildebrand et al. (2015) and Zaveri et al. (2015))

4.1 Seven-phased Approach

Based on the five identified criteria, a multi-phased methodology for evaluating data suitability is presented in the following. The implementation requires stakeholders from different disciplines to contribute both process and information technology expertise. The methodology consists of the following seven phases:

- Phase 1: Process mapping
- Phase 2: Identification of potential influencing variables (relevancy)
- Phase 3: Evaluation of completeness
- Phase 4: Evaluation of the amount of data
- Phase 5: Evaluation of accessibility
- Phase 6: Evaluation of interpretability
- Phase 7: Identification of data gaps and derivation of need for action

Parallel to the stepwise approach of the methodology, the results are documented in an overview of the framework for visualization purposes (Figure 4).

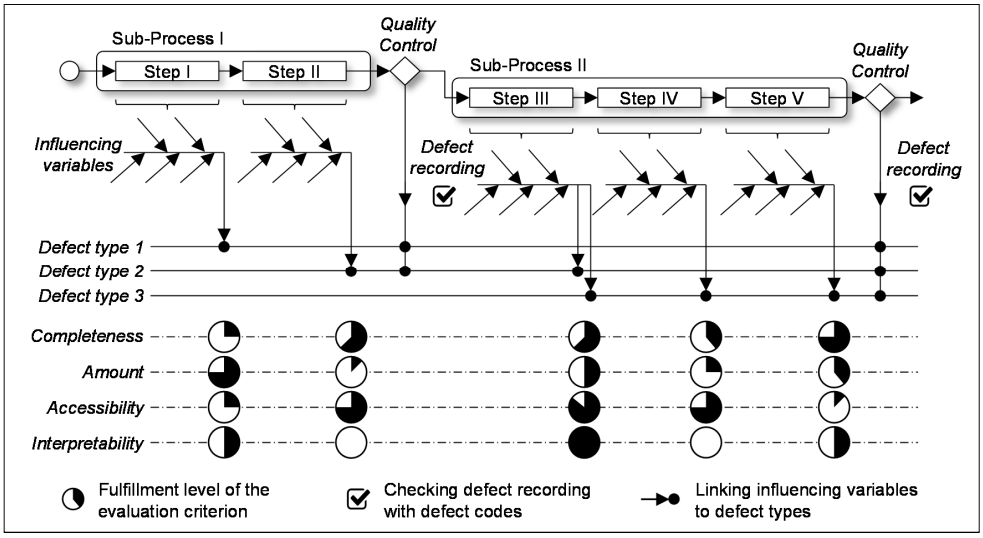

Figure 4 – Framework Visualization of the Methodology

4.1.1 Phase 1: Process Mapping

Previously existing documentation of the manufacturing process is reviewed and a profound understanding of the process is built up within the stakeholder team. In workshops, the framework for the development of a defect prediction model is set and system boundaries as well as target variables are defined. In the context of defect prediction, the target variables are documented defect types with defined defect codes. For subsequent modeling, it can be helpful to restrict the focus to particularly frequent defect types in order to reduce effort. Subsequently, the process is visualized as a functional flow chart in which the essential sub-processes of production are mapped (Figure 4).

4.1.2 Phase 2: Identification of Potential Influencing Variables (Relevancy)

In the next step, the expert team identifies potential influencing variables of the individual process steps, which could affect a specific product characteristic and thus the defect type. This is done completely detached from recorded data and thus represents only the existing process knowledge. A procedural model for the identification of quality-relevant influencing variables has already been elaborated in previous publications (Schmitt et al., 2016; Schmitt et al., 2020). The model in question consists of four steps: the identification of quality characteristics (here: defect types), quality-relevant processes, quality-relevant information and quality-relevant sources of information. The identified potential influencing variables are systematically mapped in a cause-effect diagram and assigned to the specific defect type by linking them to the defect line (Figure 4). Whether the identified influencing variables are reflected in recorded process data is then checked in the further course of the methodology.

4.1.3 Phase 3: Evaluation of the Completeness

The existing database is then evaluated in regard to its degree of coverage of the potential influencing variables identified in phase 2. This evaluation of the criterion *completeness* is carried out separately for each individual process step from a data-centric point of view. The measure of *completeness* for individual influence variables Ii is defined as follows:

$$Q_{complete}(I_i) := \begin{cases} 1 & if\ corresponding\ data\ is\ recorded \\ 0 & otherwise \end{cases}. \quad (1)$$

The evaluation of *completeness* of the data basis for entire process steps Sj results from the arithmetic mean over all potential influencing variables of the process step.

$$Q_{complete}(S_j) := \frac{\sum_{i=1}^{n} Q_{complete}(I_i)}{n} \quad (2)$$

The determined value of $1 \geq Q_{complete}(S_j) \geq 0$ is subsequently visualized as Harvey Ball and included in the methodology's framework (Figure 4) in the process overview. In addition to evaluating the *completeness* of recorded data in relation to individual process steps, it is also checked whether defect data recording has been implemented for individual products with defined defect codes (Figure 4).

4.1.4 Phase 4: Evaluation of the Amount of Data

In phase 4, the amount of available data is evaluated for every individual process step. For this purpose, the storage duration of the data recording is considered, i.e. to what extent historical data is available for the training of a predictive model. Five evaluation levels are introduced for this purpose:

- No storage of historical data
- Storage of historical data for one week
- Storage of historical data for one month
- Storage of historical data for twelve months
- Storage of historical data for > twelve months

In addition, the evaluation of this criterion is affected by the heterogeneity of the parameter variations in the data sets. On top of the parameter variation itself, special focus is placed on the occurrence of the considered defect types in the recording period. This is due to the fact that for a successful training of a predictive model a sufficiently large number of defects must have occurred in the recording period. This ensures that a balanced training data set can be created without the application of over or under sampling techniques. The overall evaluation of the criterion is then visualized in the form of Harvey Balls and placed in the methodology's framework (Figure 4).

4.1.5 Phase 5: Evaluation of Accessibility

In phase 5, the IT systems and their underlying databases in which the identified data is stored are reviewed with regard to data accessibility. The range goes from centrally stored SAP databases to decentrally distributed data in local systems on the shop floor. The focus of the evaluation criterion lies primarily on the possibilities for data export from the individual storage systems. Five evaluation levels are introduced for this purpose:

- Data export is not possible
- Manual screening and transfer of data
- Manual data export via storage device
- Semi-automated export via network
- Automated export via network

Afterwards, the evaluation of data accessibility for single process steps is again visualized in the form of Harvey Balls and placed in the methodology's framework (Figure 4).

4.1.6 Phase 6: Evaluation of Interpretability

The evaluation of the interpretability of the available data takes place in phase 6. For this purpose, the extent to which the collected data can be clearly assigned to a specific product and thus to occurring defects is evaluated. The data integration and therefore the linkability of data from different production steps to a certain product is made possible by the available meta information. If such a linkability via meta data is not given, the information contained in the data can be neither interpreted regarding its influence on the defect cause nor be modeled later. A total of four levels are distinguished in the evaluation:

- The recorded data cannot be assigned to individual products and thus to defects that occur.
- The assignment of recorded data to predefined product types can be implemented.
- The assignment to individual products can be implemented indirectly, e.g. via batch assignment or the time stamp.
- The recorded data can be directly assigned to individual products, e.g. via a product ID.

The results of the evaluation carried out are also visualized and entered in the methodology's framework (Figure 4).

4.1.7 Phase 7: Identification of Data Gaps and Derivation of Need for Action

After the production process has been mapped, potential influencing factors have been derived and the suitability of existing data has been evaluated on the basis of defined criteria, the main fields of action are systematically identified. The degree to which Harvey Balls are filled in serves as an indicator for the greatest gaps in data suitability. The elimination by means of suitable measures is then prioritized. First, measures with low personnel and financial expenses are to be addressed in order to achieve the so-called quick wins. However, in principle the strategy with which the gaps in data acquisition are to be closed depends strongly on the respective target variable. This can, for example, be characterized by a particularly frequent type of defect after a certain process step.

5 OUTLOOK

In future work the presented methodology will be elaborated in more detail as well as extended by further features. For phase 2, the identification of potential influencing variables, a flexible and process-independent meta-model for manufacturing data is currently being created to simplify the identification of influencing variables and their mutual relationships within the manufacturing process. Furthermore, the *understandability* of the data in the form of formal DQ criteria such as data format and data structure is going to be included, which is not particularly addressed in the current draft. The main focus of future work lies on the validation of the methodology on the basis of a real use-case from the field of high volume sink production. Before and after the implementation of the methodology, available data will be pre-processed, features will be extracted and predictive models for defect prediction will be created. The accuracy of these models for defect prediction will be compared before and after applying the methodology in order to verify the formulated research hypothesis: *The accuracy of a model for defect prediction can be improved by a systematic assessment of data suitability and the subsequent elimination of the identified data gaps.*

6 CONCLUSION

The systematic approach, in which the individual phases are completed step by step, makes it possible to evaluate data suitability for defect prediction with regard to various defined criteria. The continuous documentation and visualization of the evaluations allow a quick overview of the data gaps and thus the need for action along the production process. By applying the methodology before the actual modelling, the user receives information about the data suitability at a level of detail, which he does not receive by Machine Learning models with a black box character. The methodology therefore supports a systematic selection of the relevant data in the sense of the *Digital Shadow*. It is postulated that the accuracy of predictive models can be increased by using this systematic evaluation of data suitability and the elimination of identified gaps.

Since the initial effort is rather high, especially for many sub-process steps, it can be useful to limit the focus to certain sub-processes with a high number of occurring defects. In order to train a capable prediction model, formal DQ criteria are relevant as well as the applied content criteria.

ACKNOWLEDGEMENTS

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2023 Internet of Production – 390621612.

REFERENCES

Ardagna, D., Cappiello, C., Samá, W. and Vitali, M., 2018. Context-aware data quality assessment for big data. *Future Generation Computer Systems*, [e-journal] 89, pp.548-562. DOI: 10.1016/j.future.2018.07.014.

Arif, F., Suryana, N. and Hussin, B., 2013. A Data Mining Approach for Developing Quality Prediction Model in Multi-Stage Manufacturing. *International Journal of Computer Applications*, [e-journal] 69(22), pp.35-40. DOI: 10.5120/12106-8375.

Backhaus, K., Erichson, B., Plinke, W. and Weiber, R., 2016. *Multivariate Analysemethoden: Eine anwendungsorientierte Einführung*. Berlin, Heidelberg: Springer. DOI: 10.1007/978-3-662-46076-4.

Bal, H.Ç. and Erkan, Ç., 2019. Industry 4.0 and Competitiveness. *Procedia Computer Science*, [e-journal] 158, pp.625-631. DOI: 10.1016/j.procs.2019.09.096.

Batini, C., Cappiello, C., Francalanci, C. and Maurino, A., 2009. Methodologies for data quality assessment and improvement. *ACM Computing Surveys*, [e-journal] 41(3), pp.1-52. DOI: 10.1145/1541880.1541883.

Bauernhansl, T., Krüger, J., Reinhart, G. and Schuh, G., 2016. *Wgp-Standpunkt Industrie 4.0*. Frankfurt am Main: Wissenschaftliche gesellschaft für produktionstechnik.

Brecher, C., Klocke, F., Schmitt, R. and Schuh, G. eds., 2017. *Internet of Production für agile Unternehmen: AWK Aachener Werkzeugmaschinen-Kolloquium 2017*. Aachen, Germany, 18-19 May 2017. Aachen: Apprimus Verlag.

Cai, L. and Zhu, Y., 2015. The Challenges of Data Quality and Data Quality Assessment in the Big Data Era. *Data Science Journal*, [e-journal] 14(2), pp.1-10. DOI: 10.5334/dsj-2015-002.

Eger, F., Coupek, D., Caputo, D., Colledani, M., Penalva, M., Ortiz, J.A., Freiberger, H. and Kollegger, G., 2018. Zero Defect Manufacturing Strategies for Reduction of Scrap and Inspection Effort in Multi-stage Production Systems. *Procedia CIRP*, [e-journal] 67, pp.368-373. DOI: 10.1016/j.procir.2017.12.228.

Ghimire, S., Melo, R., Ferreira, J., Agostinho, C. and Goncalves, R., 2015. Continuous Data Collection Framework for Manufacturing Industries. In: I. Ciuciu, ed. 2015. *On the move to meaningful internet systems: OTM 2015 workshops, Lecture Notes in Computer Science*. Cham, Heidelberg, New York, Dordrecht, London: Springer. pp.29-40. DOI: 10.1007/978-3-319-26138-6_5.

Gürdür, D., El-khoury, J. and Nyberg, M., 2018. Methodology for linked enterprise data quality assessment through information visualizations. *Journal of Industrial Information Integration*, [e-journal] 15, pp.191-200. DOI: 10.1016/j.jii.2018.11.002.

Hildebrand, K., Gebauer, M., Hinrichs, H. and Mielke, M. eds., 2015. *Daten- und Informationsqualität: Auf dem Weg zur Information Excellence.* Wiesbaden: Springer. DOI: 10.1007/978-3-658-09214-6.

Kacprzyk, J., Gunn, S., Guyon, I., Nikravesh, M. and Zadeh, L.A. eds., 2006. *Feature extraction: Foundations and applications, Studies in Fuzziness and Soft Computing*. Berlin, Heidelberg, Springer. DOI: 10.1007/978-3-540-35488-8.

Kao, H.-A., Hsieh, Y.-S., Chen, C.-H. and Lee, J., 2017. Quality prediction modeling for multistage manufacturing based on classification and association rule mining. *MATEC Web of Conferences*, [e-journal] 123(9), p.00029(2017). DOI: 10.1051/matecconf/201712300029.

Lieber, D., Stolpe, M., Konrad, B., Deuse, J. and Morik, K., 2013, "Quality Prediction in Interlinked Manufacturing Processes based on Supervised & Unsupervised Machine Learning. *Procedia CIRP*, [e-journal] 7, pp.193-198. DOI: 10.1016/j.procir.2013.05.033.

Liu, H. and Motoda, H., 2008. *Computational methods of feature selection, Chapman & Hall / CRC data mining and knowledge discovery series*. Boca Raton: Chapman & Hall/CRC.

Pennekamp, J., Glebke, R., Henze M., Meisen T., Quix, C., Hai, R., Gleim, L., Niemietz, P., Rudack, M., Knape, S., Epple, A., Trauth, D., Vroomen, U., Bergs, T., Brecher, C., Bührig-Polaczek, A., Jarke, M. and Wehrle, K., 2019. Towards an Infrastructure Enabling. In: The Institute of Electrical and Electronics Engineers (IEEE) IEEE Industrial Electronics Society (IES), *Proceedings of the 2nd IEEE International Conference on Industrial Cyber-Physical Systems (ICPS 2019).* Taipei, Taiwan, 6-9 May 2019. IEEE. DOI: 10.1109/ICPHYS.2019.8780276.

Raudys, Š., 2001. *Statistical and Neural Classifiers: An Integrated Approach to Design, Advances in Pattern Recognition.* London: Springer. DOI: 10.1007/978-1-4471-0359-2.

Rawat, T. and Khemchandani, V., 2017. Feature Engineering (FE) Tools and Techniques for Better Classification Performance. *International Journal of Innovations in Engineering and Technology,* [e-journal] 8(2), pp.169-179. DOI: 10.21172/ijiet.82.024.

Schmitt, J. and Deuse, J., 2018. Similarity-search and Prediction Based Process Parameter Adaptation for Quality Improvement in Interlinked Manufacturing Processes. In: The Institute of Electrical and Electronics Engineers (IEEE), *IEEE International Conference on Industrial Engineering and Engineering Management*. Bangkok, Thailand, 16-19 December 2018. IEEE. DOI: 10.1109/IEEM.2018.8607361.

Schmitt, R.H., Ellerich, M., Schlegel, P., Ngo, Q.H., Emonts, D., Montavon, B., Buschmann, D. and Lauther, R., 2020. Datenbasiertes Qualitätsmanagement im Internet of Production. In: W. Frenz, ed. 2020. *Recht und Technik: Handbuch Industrie 4.0*. Berlin, Heidelberg: Springer. DOI: 10.1007/978-3-662-58474-3_25.

Schmitt, R.H., Ngo, Q.H., Groggert, S. and Elser, H., 2016. Datenbasierte Qualitätsregelung. In: R. Refflinghaus, Ch. Kern, and S. Klute-Wenig, eds. 2016. *Qualitätsmanagement 4.0 – Status Quo! Quo vadis? - Bericht zur GQW-Jahrestagung 2016*. Kassel: Kassel University Press. Ch. 6. DOI: 10.19211/KUP9783737600859.

Schuh, G., Rebentisch, E., Riesener, M., Ipers, T., Tönnes, C. and Jank, M.-H., 2019. Data quality program management for digital shadows of products. *Procedia CIRP*, [e-journal] 86, pp.43-48. DOI: 10.1016/j.procir.2020.01.027.

Škulj, G., Vrabič, R., Butala, P. and Sluga, A., 2013. Statistical Process Control as a Service: An Industrial Case Study. *Procedia CIRP*, [e-journal] 7, pp.401-406. DOI: 10.1016/j.procir.2013.06.006.

Uhlemann, T.H.-J., Schock, C., Lehmann, C., Freiberger, S. and Steinhilper, R., 2017. The Digital Twin: Demonstrating the Potential of Real Time Data Acquisition in Production Systems. *Procedia Manufacturing*, [e-journal] 9, pp.113-120. DOI: 10.1016/j.promfg.2017.04.043.

Wang, K.-S., 2013. Towards zero-defect manufacturing (ZDM)—a data mining approach. *Advances in Manufacturing*, [e-journal] 1(1), pp.62-74. DOI: 10.1007/s40436-013-0010-9.

Wang, R.Y. and Strong, D.M., 1996. Beyond Accuracy. What Data Quality Means to Data Consumers. *Journal of Management Information Systems*, [e-journal] 12(4), pp.5-33. DOI: 10.1080/07421222.1996.11518099.

Wuest, T., Irgens, C. and Thoben, K.-D., 2013. Analysis of Manufacturing Process Sequences, Using Machine Learning on Intermediate Product States. In: C. Emmanouilidis, M. Taisch and D. Kiritsis, eds. 2013. *Advances in Production Management Systems, IFIP Advances in Information and Communication Technology*, Berlin, Heidelberg: Springer. Vol. 398. DOI: 10.1007/978-3-642-40361-3_1.

Zaveri, A., Rula, A., Maurino, A., Pietrobon, R., Lehmann, J., Auer, S. and Hitzler, P., 2015. Quality assessment for Linked Data. A Survey. *Semantic Web*, [e-journal] 7(1), pp.63-93. DOI: 10.3233/SW-150175.

ABOUT AUTHORS

Peter Schlegel – (P.S.) Laboratory for Machine Tools and Production Engineering WZL of RWTH Aachen University, Aachen, Germany, Research Associate, e-mail: p.schlegel@wzl.rwth-aachen.de, Author's ORCID: 0000-0003-0906-4644.

Daniel Buschmann – (D.B.) Laboratory for Machine Tools and Production Engineering WZL of RWTH Aachen University, Aachen, Germany, Research Associate, e-mail: d.buschmann@wzl.rwth-aachen.de, Author's ORCID: 0000-0001-8363-1331.

Max Ellerich – (M.E.) Laboratory for Machine Tools and Production Engineering WZL of RWTH Aachen University, Aachen, Germany, Head of Department, e-mail: m.ellerich@wzl.rwth-aachen.de, Author's ORCID: 0000-0001-9113-0203.

Robert Heinrich Schmitt – (R.H.S.) Laboratory for Machine Tools and Production Engineering WZL of RWTH Aachen University, Prof., Aachen, Germany, e-mail: r.schmitt@wzl.rwth-aachen.de, Author's ORCID: 0000-0002-0011-5962.

AUTHOR CONTRIBUTIONS

P.S. – conceptualization, methodology and writing; D.B. – methodology and validation; M.E., R.H.S. – review and editing.

CONFLICTS OF INTEREST

The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.