TQM Framework for Healthcare Sectors: Barriers to Implementation

DOI: 10.12776/QIP.V26I1.1611

Mohammad S.A. Ansari

Received: 2021-09-24 Accepted: 2021-12-10 Published: 2022-03-31

ABSTRACT

Purpose: A qualitative research was carried out with an aim of understanding and addressing the challenges of integrating TQM in the healthcare industry. It examines the existing inadequacies in the service quality, and barriers in implementation of TQM practices,

Methodology/Approach: A conceptual model is developed to explore the issues arising TQM implementation. There are three major components of TQM: such as barriers to implementation, Critical Success Factors and Benefits of TQM. Interviews with health workers and patients were conducted based on pre-structured questionnaires. Seven hypotheses were developed to investigate how TQM can be achieved irrespective of surmounting barriers.

Findings: Findings suggest TQM can be implemented in the right environment with committed leadership and supportive infrastructure, which would drive SQ, improved customers, and employees' satisfaction and loyalty, increase profitability and shareholder values. TQM can deliver high quality medical care for overall performance of the healthcare industry.

Research Limitation/Implication: Interviews were conducted with semi-structured research questionnaires. There may be some inevitable biases present in questionnaires and evaluation of review.

Originality/Value of paper: The study benefits from insights from medical personnel and patients' perspective, in exploring the SQ attributes, i.e., quality circle, continuous improvement, employee empowerment and customer focused approach.

Category: Research paper

Keywords: total service quality; healthcare; organizational excellence; perceived satisfaction; operational efficiency

1 INTRODUCTION

Total Quality Management (TQM) had its origin from the manufacturing sectors to measure the overall performance of a plant or industry. The most important function of TQM is to promote business performance through accomplishment of quality products, service delivery, and improved performance of operation and maintenance. According to Zgodavova and Colesca (2007), quality management systems guarantee the maximum customer satisfaction by constant improvement in processes at minimal spending. The improved service quality (SQ) unquestionably enhances business, and therefore, it is highly adored.

Any product or service without an acceptable SQ will not flourish and is bound to fail; therefore, lack of SQ can be disastrous. Which can be gauged by comparing customers' expectations of quality and what they receive, the difference being called a service gap. The service gap naturally inhibits or discourages customers from being satisfied and loyal (Izogo and Ogba, 2015). Providing poor quality of product or service proved to be costlier than producing high-quality products or services, because poor quality will eventually lead to rework and unsatisfied customers. Therefore, SQ and TQM become an important aspect of any industry or business irrespective of nature or type of business.

The effective implementation of TQM procedures had enhanced functioning of any company (Hassan et al., 2012). It should be developed as part and parcel of organizational culture to embrace TQM with an aim of moving out product, or delivery of service with excellence, at first time and all the time. TQM is an advancement or a progression of appropriate attribute measures within the total range of value addition that enriches accomplishment. It helps in producing improved quality of goods in meeting and exceeding customers' expectations (Yusof and Aspinwall, 2001). Therefore, the organizations that successfully implemented TQM had improved operational efficiency, throughput, sales, revenue, and effectiveness.

The success of TQM mainly depends on committed leadership, strong organizational culture, active participation, and workforce cohesiveness. As per (Dahlgaard, Pettersen and Park, 2011), tangible factors such as leadership, partnership, managing people and service delivery system are crucial for TQM success. This undoubtedly will lead to producing goods and delivery of services at optimal cost. Initially TQM was very much confined to the production line, however, with passage of time it was felt that the service industry too should comply with TQM; therefore, it got similar status in the service industry as well. According to Prajogoa and Hong (2008), TQM is a philosophy that can even promote the right environment other than production or manufacturing. The leadership or top management commitment and orientation towards nurturing quality within and outside of organization are crucial.

1.1 Research Objectives

The aim of this study is to examine the contribution of technology, acquaintance of TQM axioms, practices, and its impact on the bottom-line, particularly aftereffects of pandemic. The debacle has clearly exposed how many gaping loopholes and shocking lacunae exist in the present healthcare set up of many countries. It has forced the healthcare system to seriously relook into the architecture of the whole country and draw up suitable policy and strategy to safeguard society from current and future disasters. According to Kajihara et al. (2016), healthcare services should be able to withstand even during disaster. While doing so it is prudent to integrate the health policy with proper TQM policy, to arrive at a comprehensive system which is efficient and effective. Hence the following research objectives are set for this study:

- To find the contribution of technology in achieving excellence in healthcare services.
- To find to what extent TQM is implemented in the healthcare system.
- To examine the significant correlation between TQM methods and profitability.

1.2 Research Questions

The purpose of the research is to ascertain the present position of TQM and its gaps in the healthcare sector, challenges, barriers, and implementation. The aim is to enhance perception of service providers based on patients' inputs on healthcare competence and quality care endowment. The research attempted to address the following research questions (RQs), which are raised based on literature review and critical success factor (CSF) such as technology, quality compliance, team structure, customer focused approach, employees training and development and to investigate them with an aim of proposing remedial action for accomplishing TQM practices in healthcare sector:

RQ1: Would technology support in implementing TQM?

RQ2: To what extent TQM is understood and implemented?

RQ3: What are the CSFs for healthcare providers from patients' perspectives?

RQ4: What are the critical links between TQM practice, quality care and profitability?

The CSF underlined its effort on attaining SQ, for eventually advancing overall customer fulfilment for improved revenue, and profitability. An extensive evaluation and interpretation of these factors likely to attain success. As a result, the present study is expected to provide guidance to greater extend and suggest in advancing strategies for an efficient and effective TQM implementation in the healthcare sector.

In this connection a TQM framework that is tightly woven with the right healthcare policy framework, will be a formidable shield to protect and save society and make a healthy and happy nation. This study attempted to provide a framework that will be valuable for healthcare providers, hospital administrators, and those concerned with health matters. The following paragraphs discuss literature review, methodologies, analysis, and conclusions.

2 LITERATURE REVIEW

TQM is not a quick-fix solution; therefore, firms must encounter multiple hurdles and challenges in its adaptation and realization. Most of the time TQM benefits could not be derived, due to lack of support from leadership. However, it has received great attention from industries, research and development and academia. In the context of the healthcare industry, TQM can be understood as focused efforts in formulating and transforming for better patient care, system performance, and professional development. According to Ali and Alolayyan (2013), there is an interdependence between TQM practice and hospital achievements. It involves an explicit change in method, support in clinical or administrative systems for eventual improvement of the healthcare system. The following paragraphs discuss SQ improvement, taxonomy in healthcare and how TQM can be practiced.

2.1 Service Quality

It is imperative to examine and comprehend the notion of quality, before discussing the TQM concept. A few questions naturally arise such as, “Are we doing the right thing?” and “Are we doing things right?” These are the twin classic key questions asked in application of quality procedures in any domain of activity. The concept of SERVQUAL, as articulated by Parasuraman, Zeithaml and Berry (1985), helps us to come to terms with the slippery nature of quality. It is a multi-dimensional research instrument of SQ that elegantly and brilliantly captures consumer’s expectations and perceptions of a service. it asserts that a customer’s evaluation of SQ decides the gaps between their anticipations, assessments, and its genuine accomplishment. Parasuraman, Zeithaml and Berry (1985) have further recognized and published the application of SERVQUAL through their research such as Parasuraman, Zeithaml and Berry (1985), Parasuraman and Grewal (2000), and Parasuraman (2010). Originally it was identified with ten components, which later reduced to five Viz Reliability, Assurance, Tangibles, Empathy and Responsiveness (RATER). Among the well-known definitions of quality by various researchers, a few are mentioned below:

- Quality is comparable to end user’s contentment (Kaoru Ishikawa, 1985)
- Quality shall be able to determine by the client (Edwards Deming, 1986)
- Quality is fitness for use (Joseph Juran, 1989)

- All types of behaviour that deemed to put best practices would influence TQM (Hackman and Wageman, 1995)
- Quality is meeting customer requirements (Oakland, 2003)

Although satisfying end users' anticipations and requirements, which is a "common denominator" in all these explanations. Each quality professional describes the quality differently based on their individual insight and experience. According to Parasuraman, Zeithaml and Berry (1985), SQ is essentially a type of mindset, which is formed based on customers' perception and observance, in relation to the expectations. Improved SQ will enhance production and temporarily increase operating cost. Therefore, a manager must decide how much to devote on quality to obtain, the best outcome in terms of return on investment or in other words return on quality. According to Rust, Zahorik and Keiningham (1995), the return of quality methodology empowers executives to ascertain, where to spend, how much, in regard to expected monetary benefits such as profits, and return on investments.

Therefore, service excellence or lack of it can be gauged through evaluating customer's expectation with respect to provider's performance, and the difference representing the perceived service gap. Accordingly, service providers shall benchmark and evaluate the gap between what customers expect, what is delivered and how to close these gaps between perceptions and actual practices. According to Deros, Yusof and Salleh (2006), significant differences between customers' perception of service and actual practices in fact constitute a CSF, such as technology, team structures, customer focus approach, top management commitments towards quality compliance, employees training and quality planning. These can work as a benchmark in calibrating processes and operations. Having seen the importance and necessity of SQ, it is now worth discussing TQM and its ramifications. The following section delves into various aspects of TQM.

2.2 Total Quality Management

The fundamental concepts of TQM were presented in early 1990s by Kaoru Ishikawa (1985), Edwards Deming (1986), Joseph Juran (1989), Hackman and Wageman, (1995) and Oakland (2003). TQM has been defined by different authors in different instances, e.g., "a search for excellence", "perfection, first time and every time", "zero defects", "delighting the customers" and so on. The eventual rationale was for attainment of customer satisfaction, employee satisfaction, and improved products and service through continuous improvement (CI), innovation, reduction of waste, higher productivity, increased sales, revenue, and profitability. According to Hassan et al. (2012), effective implementation and adoption of TQM methods will result in enhancing administrative functioning.

TQM receiving statistical process control is deeply ingrained and embedded. It was coined by "Walter Shewhart" in the early 1920s in the United States of America. The methodology was centered on categorizing specific variables and found inconsistency within manufacturing progression that was not in line with the measurement. It resulted in a logical process centered with the philosophy of PDCA (Plan Do Check and Act) for quality enhancement in the manufacturing sectors (Evans and Lindsay, 2001). TQM grew with passage of time and different researchers based on their findings identified various best practices. According to Dale, Wiele and Iwaarden (2013), the evolutionary phases of TQM are Quality Enhancement, Quality Control (QC), Quality Assurance, and TQM (refer Figure 1).

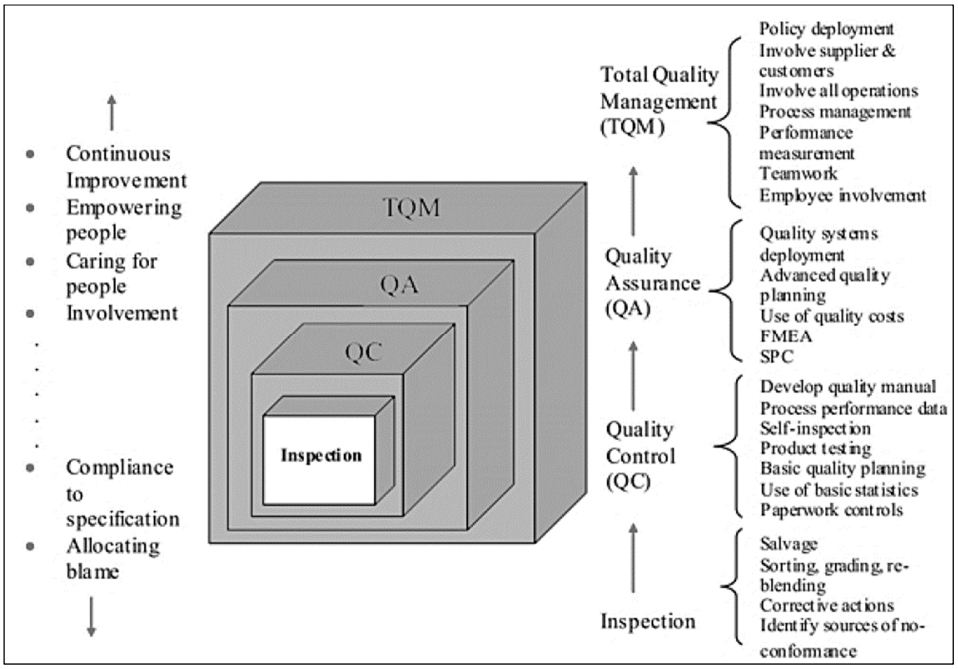


Figure 1 – The Four Stages of Evolution of TQM (Dale, Wiele and Iwaarden, 2013)

Quality improvement and continuous improvement remains a significant business strategy in business surroundings, irrespective of types of business, whether operation, production, healthcare service, hotel, consultancy, or financial services. The victorious organization will never be satisfied with status quo and strive for CI of SQ for attaining TQM and marching toward Six Sigma. According to Antony et al. (2018), SQ implementation in healthcare has abundant prospects for advancement. The successful companies would everlastingly find new ways of improving products and SQ for their definitive success.

Based on literature review (Rahman, 2019) has found seven best practices in TQM implementation, namely:

- Leadership commitment,
- Cooperation and Collaboration,
- Coaching,
- Procedures,
- Progressive growth,
- Skill enhancement, and
- Corporate philosophy.

Rahman also emphasized that these enactments were recognized based on its phenomena and suggested that these applications have great value in healthcare sectors. However, healthcare quality may be defined differently, with multiple healthcare providers, such as number of patients, healthcare stakeholders for ultimately attainment of SERVQUAL. According to Balasubramanian (2016), the SERVQUAL model and its application can support healthcare in accomplishing gratification for service providers, consumers and further it suggests that TQM can have excellent support to healthcare establishments.

2.3 TQM in Healthcare Industries

Healthcare services shall fulfil the likelihood of desired outcomes of organization and should achieve preferred results in line with their expectations and preferences. At the same time, it should follow standards and scientific evidence, clinical and ethical practices, respecting individual preferences, such as customer friendly environment, location, and even ambience could play a greater role in eventually gaining patients' satisfaction. As noted by Fatima, Malik and Shabbir (2018), healthcare SQ aspects like natural ecosystem, consumer-friendly approach, attitude, confidentiality, and security are optimistically associated with patient's loyalty.

The fundamental work of healthcare includes clinical and non-clinical services, procedures and activities carried out in addressing issues like "act virtuously" and "Implementing Things Right". "Doing Right Things" assesses efficacy of clinical service, whereas "Implementing Things Right" contemplates methodological healthcare service. Producing or capable of producing a result, would depend on how procedures and processes are put in place and how implemented in an organization, which certainly would depend on managers' guidance and support. According to Mosadeghrad (2013), TQM enactment and its influence heavily depend on the manager's capability to follow its concepts and principles. Efficiency can be achieved through proper planning and implementation, whereas effectiveness can be achieved by producing a desired result, desired output based on clinical knowledge and skills.

The efficiency can be increased by avoiding waste of materials, man hours, under-utilization of equipment, misuse of materials, energy, amenities, and even non-implementation of right thoughts can lead to inefficiency. According to Dénes et al. (2017), inefficiencies can lead to poor economies of scale and ultimately increased cost. The patients' care and ethical principles involved in those activities should be taken seriously for administrative effectiveness, well-being and company bottom line.

According to Ansari (2020), SQ enhances worker pride and their efforts and motivation develop customers' loyalty, contentment, which promote repeat purchase behaviour. The technical management or clinical performance will improve SQ, end users' satisfaction and loyalty. Contrarily, non-clinical management, such as interpersonal skills can promote patient care and co-production of services. Therefore, clinical, plus nonclinical processes would eventually lead to TQM. According to Ozdal and Oyebamiji (2018), the implementation of TQM methods stretches from top executives to shop floor employees.

3 RESEARCH METHODOLOGIES

Based on the RQs, research methodologies were developed. Accordingly, qualitative research was carried out to investigate the research questions and how to overcome the challenges of the TQM framework in the healthcare industry. It was found that major contributions are from quantitative research, followed by mixed and qualitative research. According to Cameron and Azorin (2011), quantitative techniques take up 76%, which is overwhelmingly the dominant method of choice; mixed methods correspond to 14% and qualitative research 10%. The choice of quantitative research was to explore and extract healthcare professional and public opinions, thoughts, and feelings on how TQM attributes significantly affect the healthcare sector.

Therefore, the present researcher planned to follow qualitative research technique by receiving direct feedback through discussions, from medical professionals such as doctors, paramedics, and other supporting staff such as clinical administrator, hospital administration, medical records administrator, medical secretary, dieticians as well as inpatients and outpatients. A common misguided belief is that sample size in qualitative study is not crucial. Determining adequate sample size is ultimately a matter of judgment, which depends upon individual skills in assessing the quality of data gathered against the sample size. However, researchers have proposed sample sizes depending upon type of research. The sample size shall depend on research objectives and should be sufficient for attainment of saturation. Therefore, beyond the saturation point adding more participants to the sample will not skew the result.

There are various types of qualitative studies such as Ethnography, Grounded theory, and Phenomenological studies. The TQM in the health sector is

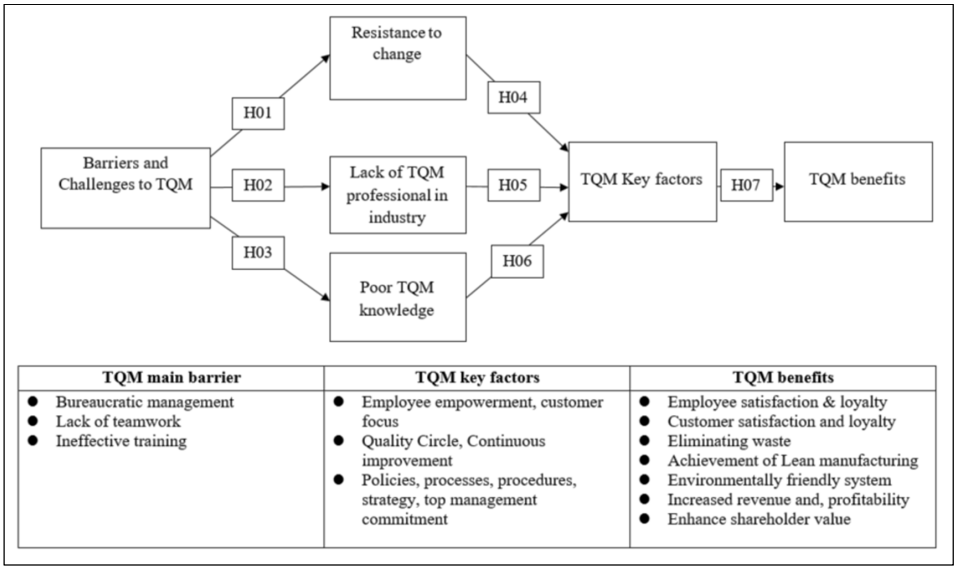
considered to belong to the type of Phenomenology studies, which deals with conscious experience, i.e., feeling from the subjective or first-person point of view. As Creswell (1998) proposes a sample size of 5 to 25 for Phenomenology studies, whereas Morse (1994) advised that at least six samples are essential for such studies. Therefore, based on sample significance, and available recommendations in literature, 5 samples each were collected for doctors, paramedics, supporting staff, inpatients, and outpatients. Sample for earlier research was limited to healthcare staff or in some cases patients' feedback was obtained. In this research an integrated chain of services employees was considered, to overcome limitations of earlier research. Data was collected through direct interviews from physicians, physiotherapists, nurses, technicians, supporting staff, and patients, which formed an integrated chain of service.

A prior email communication was set up for formal interview with professionals explaining the purpose of research and finalization of schedule. Most of the professionals proposed that their interview not be recorded; therefore, discussions are noted on paper. Couple of participants refused to participate either due to unawareness of the subject or unwillingness to participate. Structured and unstructured questions were addressed to the respondents on a case-to-case basis. According to Guetterman (2015), while contemplating sample size, scholars should go ahead of "how many?" to address the queries of "how?" and "why?". Accordingly, pre-drafted questions were addressed to the participants. Patient feedback was taken after seeking their permission, which was done randomly and instantly without any prior appointment.

Interview script was designed by brainstorming and developing open-ended, simple, easy to understand questions that allowed interviewees to do most of the talking. The questionnaire was shared in advance that permitted the narrator to think beforehand. In the beginning simple questions were addressed, whereas in the middle difficult and deeper questions were administered and in the final section, some clarification and word of thanks were exchanged. In most of the cases the interview lasted for 30-40 minutes. Manually coded qualitative data was categorized by figuring out the code frame, identifying which theme came up the most and by subdividing into a hierarchical coding frame for finding relationships embedded with the responding feeling on the subject.

4 SYSTEM MODEL

The research model is developed based on research purpose, research questions, accordingly, the system model was evolved. This system model abstractly captures the essence and spirit of a prototypical healthcare industry with many departmental elements, wherein SQ attributes act as nodes and arrows as "causal" conduits linking them. Particularly, these paths efficiently connect and unify with TQM characteristics and are validated by analysing a cluster of hypotheses. The following sections provide TQM conceptual framework, hypotheses testing, findings, and conclusion (Figure 2).



Notes: H01: Bureaucratic management leads to resistance to change.; H02: Absence of collaboration leads to dearth of TQM professionals in industry.; H03: Ineffective training leads to poor TQM knowledge.; H04: Overcoming resistance to change will improve employee empowerment, and customer focus.; H05: TQM professionals will support Quality Circle and Continuous Improvement.; H06: TQM knowledge will help in setting-up policies, processes, procedures, and strategy.; H07: Key TQM factors will drive TQM benefits.

Figure 2 – TQM Conceptual Framework

4.1 Hypotheses Testing

The conceptual model showing relevant attributes of the present research is shown in Figure 2. It embodies three major segments viz., Barriers, Drivers, CSFs, and Benefits. The model proclaims that there exists a set of barriers that hamper and impede attainment of TQM key factors, which are considered as CSFs. These CSFs later drive benefits and surmounting the barrier to achieve CSFs is a two-step and through three-branch process, as reflected in the model. Three traits connect these two ends/segments viz. proper mindset, professionalism, and awareness. These CSFs are in fact the "enablers" or drives to reap TQM benefits. A Cross-sectional research was accomplished in designing questionnaires and working of hypotheses through literature review (Alzoubi et al., 2019; Antony et al., 2018; Balasubramanian, 2016; Fatima, Malik and Shabbir, 2018; Mosadeghrad, 2013; Ozdal and Oyebamiji, 2018). Consequently, driven by existing conventional insight and engagement in service activities and academia, seven hypotheses (H_{01} to H_{07}) were developed and analysed in subsequent paragraphs.

H_{01}: Bureaucratic management leads to resistance to change.

The TQM program can only succeed, when encouraged by top-management with all-round support. It can be achieved through the right organizational structure and by embracing quality culture. According to Joiner (2007), realization of

TQM methods is accompanied with improved company bottom-line. Hence, organizations shall build a culture or environment of encouragement that would promote and support execution of TQM. There are obstacles in healthcare setup, such as lack of administrative commitment, inexperienced middle level management and resistance to change.

Other crucial obstacles of TQM are lack of employees' curiosity, poor skills, development, workers' devotion, interest, and participation. In addition, there is an overall work pressure of medical professionals hampering acclimatization to TQM culture. As per Sila and Ebrahimpour (2002), the obstacle of achieving TQM is HRM issues, and above all "iron curtain" existing between divisions or functions, high labour turnover and "uncooperative" culture. Sometimes it posing challenge for management to convince and communicate TQM benefits to employees, as a result employees tend to perceive that it is going to overburden or threaten their "identity". Incompetent employees also contribute to this situation, who perhaps feel risk of losing job, due to lack of education or training. Even labour unions can make or break the successful implementation of TQM. *Accordingly, it is hypothesized that bureaucratic management leads resistance to change.*

H_{02}: Absence of collaboration leads to dearth of TQM professionals in industry.

Teamwork can be a magic cure, if these precepts like "working together"; "leaving no-one behind" are ingrained in the system. When a team is cohesive and working for a common goal, employees will feel they are working for each other. On the contrary, poorly run teams can do more harm to groups than having no teams at all. A trusted team leader with leadership skills could build trust and develop mutual support among the members. It can turn out to be supportive and will be able to assist each other in accomplishing the organization's performance. Hence, top executive support will actively lead to better patients' care in the healthcare industry. According to Rahman (2019), TQM implementation can be identified through top-management commitment, leadership; teamwork; training; competency development; and organizational culture. Therefore, teamwork should be institutionalized in the healthcare system through development of the right culture, training and development, and continuous improvement.

The top management affiliation, employee efficiency and commitment in TQM curriculum would achieve success. According to Slimák and Zgodavova (2011), success is only a sustainable fortune for individuals and for organisations. Therefore, teamwork is a significant tool that has potential to create a high-performance, well-organized, and strong healthcare setup. TQM can focus on long, medium-term and short-term goals and can provide a cohesive vision for systemic change to all types of industries including healthcare. According to Al-Dhaafri, Zien and Al-Swidi (2013), the capacity to create a company's ethos that can promote the invention and accept the failure as the way of discovery. As part of senior management commitment, the TQM initiatives should be taken as

managerial solutions towards gaining competitive advantage. *Therefore, it is hypothesized that absence of collaboration leads to dearth of TQM professionals in industry*.

H_{03}: Ineffective training leads to poor TQM knowledge.

The type of training imparted to staff would depend upon individual skills, purpose, needs and hierarchical level. General leadership training can be given to managerial cadres, technical to specialist, TQM training to specialist as well as for leadership cadre. According to job requirements, specialized training such as use of eminence tools or techniques can be given to experts. According to Talib, Rahman and Qureshi (2010), CSFs for successful implementation of TQM in service industries is to achieve business success and perfection in quality service. Every organization needs to decide to whom to impart general or specific training based on expected benefits and CSFs.

Effective and systematic training is required for understanding and implementation of TQM. It can be done by identifying training requirements, planning, and implementation of suitable training programs. Proper training would assure that employees have enough support to develop their skills. It can be personalized in training objectives for gaining efficiency, and by evaluating the effectiveness of training programs. According to Ali et al. (2017), employee training and relationships can positively influence business. The evaluation and effectiveness of training should be carried out by respective line managers. *Accordingly, it is posited that ineffective training will lead to poor TQM knowledge.*

H_{04}: Overcoming resistance to change will improve employee empowerment, and customer focus.

It has been said that “change is always good”, one frequently hears that “nothing is certain except change” or “change is always resisted”. In most of the cases it is observed that employees become the prime opponent of change. Total management support is essential, and without their support, the TQM program would not take-off. The managers shall involve employees in implementation and accomplishment of TQM. According to Sadikoglu and Olcay (2014), firms should enhance employees’ engagement/dedication/understanding towards TQM to overcome its barriers. If a manager instead of being “change agent” takes up a postman role, TQM will suffer enormously. Therefore, the manager shall be responsible and should take a proactive role in TQM accomplishment.

Top management commitment toward employee empowerment is considered as one of the most significant values toward TQM, as it builds strong and everlasting alliance with end users. It will improve customer satisfaction, enhanced reliability, increased sales, improved revenues, and profitability. These can be accomplished only through satisfied and loyal employees, care, and empowerment. Therefore, employee empowerment should be most effective in connecting the silos and bridging the gaps to mitigate employees’ resistance to

change. Once employees are willing to accept change, they will be proactive, satisfied, and loyal, thus will concentrate on customers and become customer focused. *Accordingly, it is posited that overcoming resistance to change will improve employee empowerment and customer focus.*

H_{05}: TQM professional will support Quality circle and uninterrupted progress.

The accomplishment of TQM relies on a simple, yet straightforward organizational structure. However, typical healthcare organizations are systems that are composite by design and complex by operation with multiple levels of bureaucracy, and functional units, posing hurdles for smooth delivery of healthcare services. In the given composite structure, it is challenging to accomplish synchronization. In the complex environment, establishing Quality Circles (QC) is challenging, which is a pre-requisite for TQM for an organization and it will not flourish in a complex and complicated system and that too in the absence of poor leadership commitment and encouragement. According to Albliwi et al. (2014) identified several common reasons for the TQM catastrophe, such as lack of leadership commitment, participation, communication, training, and inadequate resources. The QC is required to be formed by a steering committee, which will require leadership involvement, nomination of QC volunteers and designation of senior staff as facilitators. At the same time, the QC recommendations shall be religiously implemented.

QC is likely to bring change through CI, thereby the organization continually undergoes change to improve overall SQ. Improved quality will fill service gaps to meet and exceed customers' expectations. Accordingly, customers' feedback shall be collected on regular intervals. QC teams typically shall work on a constant basis that will satisfy customers and end users. According to Kassinis and Soteriou (2003) customer loyalty led to increased revenues and reduced future transaction costs. Therefore, QC and CI shall fall within the purview and shall form as foundational "bricks" of TQM. *Accordingly, it is predicted that TQM professionals will improve Quality Circle and Continuous Improvement*.

H_{06}: TQM knowledge will help setting-up policies, processes, procedures, and strategy.

The TQM can be viewed as the convergence of three major streams of thinking about quality that evolved over the past few decades: inspection, quality circle, quality assurance, which led to various certifications such as ISO 9000, ISO 14000 etc. According to Carnerud and Bäckström (2019), quality has centered on SQ & end users' satisfaction; process design & control; ISO certification & standards. ISO was introduced in 1987 by the International Organization for Standardization and it proposed seven basic quality management tools viz. customer focused, leadership, people engagement, process-centric, continuous improvement, and evidence-based decision and rapport management. It mainly focused on leadership in organization and people management. According to Sampaio, Saraiva and Rodrigues (2011), the motivation to implement ISO 9000

is a CSF that speaks about the quality management and financial achievement of many companies.

Lean Management (LM) became an important process of TQM that pleaded for methodical techniques of eradicating waste and incompetency. According to Bhamu and Sangwan (2010), there are an abundance of LM classifications with conflicting objectives and scope. LM looks for things that add value, and those which do not shall be eliminated. It also beseeches for waste reduction in transportation, inventory, action, over- production, over processing, and defects. According to Choomlucksanaa, Ongsaranakorna and Suksabaia (2015), lean manufacturing processing reduced time, cost, and non-value-added activities. LM focuses more on identifying incompetency in the production process, rather than encouraging quality system.

According to Antony, Snee and Hoerl (2017), the integration of Lean and Six Sigma (LSS) is significant as the Lean target on enlightening the exchange of ideas and flow of materials, whereas Six Sigma mechanisms add value in transformations. Similarly, Six Sigma is a business approach for functioning and obtaining service eminence. Hence it is important that healthcare industries take enough trouble to understand the benefits and success of Six Sigma. According to Nakhaiand Neves (2009), the persistent drive of embracing Six Sigma has developed an impractical expectation than what Six Sigma is genuinely capable of accomplishing. It is likely that Six Sigma will enable future work to progress, develop and evolve, which will open new prospects in the field for a sustainable future.

According to Muñoz and Gutierrez (2017), the Six Sigma emphasizes on customer input, improvement in design and control practices, but also exploratory methods, such as invention, originality, or revolution. LM combined with Six Sigma is the futuristic strategy for the operation management and service industry. According to Muraliraj et al. (2018), there is a quandary in combining LSS for those who had preliminary experience with one or the other notion. However, the industry became acquainted with LSS during the initial stage of the new millennium. *Hence, it is predicted that TQM knowledge will set policies, processes, procedures, and strategy.*

H_{07}: TQM key factors will drive TQM benefits.

Healthcare system is gradually moving toward quality improvement; however, many eyebrows were raised whether TQM as a strategy will pay back. According to Kim, Kumar and Murphy (2010), it is important to adjudicate if TQM will thrive or flop, however, the motivation is that TQM has been constantly evolving. TQM will certainly succeed, without fail, when certain sets of practices are put into action. According to Kumar et al. (2018), many companies having the right strategy will be able to accomplish TQM effectively, whereas others will have difficulties. During these efforts while implementing, many companies have come to understand more about the dynamic and a somewhat near “protean” nature of TQM. However, as a concept and philosophy, TQM is universally

accepted to some extent but not widely discussed in the healthcare industry. According to Alzoubi et al. (2019), there is scarcity of TQM research in healthcare perspective.

It was observed that the significance of top management support, employee training and development are extremely important to bring about a transformational change in an organization. It can lead to improved organization overall operational efficiency. According to Taddese and Osada (2010), TQM affects techno-process innovation mediated primarily by human resource and working conditions. The overarching purpose of implementing TQM in the healthcare industry is to effect changes in management style and culture. Improved knowledge of TQM is likely to improve overall working traits in an organization such as job satisfaction, employee loyalty, employee retention and commitment and industry competitiveness. As Ansari, Farooquie and Gattoufi (2018) service organizations should pursue employee loyalty to build improved relationships with customers in delivering better SQ.

The TQM factors will drive employee loyalty. As employees are the extremely valuable resource of any organization for productivity and performance improvement. The successful company considers employees as "prime movers" in achieving results sooner than later. According to Psomas et al. (2014), quality is a considerable component, which promptly impacts employee benefits, customer satisfaction and business accomplishment. The performance improvement should be strongly managed by the company since it directly helps in profitability of the company. According to Yee, Yeung and Cheng (2011) there is persuasive argument among SQ, employee and customers' satisfaction/loyalty and firm profitability. *Accordingly, it is predicted that TQM key factors will drive TQM benefits.*

5 RESEARCH IMPLICATIONS

The research attempted to investigate the contribution of technology on TQM, understanding of TQM practices, its benefits to the patients, society and how it can influence company profitability. TQM can be used as a synergistic approach in implementing systems, overcoming clinical and administrative issues. It can enhance patient satisfaction, loyalty, through continuous improvement and by facilitating healthcare service at a reasonable and competitive cost. TQM has received great attention in all the fields including academia, managerial application, and operation management.

5.1 Implication in Service Industry

There is enough scope for TQM in service industry research. According to Ozdal and Oyebamiji (2018), the implementation of TQM methods stretches from executive management to employees' satisfaction. It can improve overall operational efficiency and improve employees and customers' satisfaction. TQM

needs right planning, financial support, however poor management can put TQM at risk. According to Sampaio, Saraiva and Domingues (2012), numerous necessities should be studied such as leadership dedication, availability of resources, communication, training and development within the company.

5.2 Managerial Applications

The TQM will be a great opportunity from a managerial point of view, in terms of minimized waste, improved product and SQ, improved revenue, and profit. It can be achieved by developing TQM culture, leadership support, employee empowerment, bridging the silos between employees and management that can overcome resistance to change. As per Rahman (2019), the best TQM implementation can be identified through leadership dedication, harmony; and employee coaching. TQM will be appreciated by employees once they are satisfied, and loyal.

5.3 Operation Management Application

The TQM can be vision for change for all types of industries, which can derive benefits to society in terms of increased customer satisfaction, higher productivity, waste reduction, reduced defects, lower cost, increased revenue, and profitability. It will reinforce competitive position due to higher productivity, improved operating cost, improved customer satisfaction, loyalty, and employee's engagement. According to Joiner (2007), realization of TQM methods is associated with enhanced organization performance. It will add value to all stakeholders in terms of improved payback. Industry must be cautious in implementing TQM by keeping labour unions engaged, since any disagreement can make or break the TQM.

6 FINDINGS

The present framework is especially useful, which in turn will drive to achiever TQM benefits. The chief barrier of implementing TQM in healthcare is the bureaucracy, since healthcare setups in many hospitals are departmentalized with numerous units, sub-units, and subsystems, with a web of cross-linking management controls. This entails the necessity of multiple layers of hierarchy with multiple "command and control" structure. Such a bureaucratic style of management is antithesis in adopting TQM.

According to Candido and Santos (2011), though there is no evidence of empirical validation, yet many companies may deter from adopting TQM for the sheer magnitude of this undertaking. The other major barriers of TQM implementation include lack of dedication of executive, or poor teamwork, unproductive training and advancement, absence of employee empowerment, paucity of processes, procedures and even the non-existence of QC and CI

philosophy. As Khan, Malik and Janjua (2008) affirm that TQM substantially defines employee accomplishment due to their contentment and dedication.

7 CONCLUSION

Conclusions suggest that TQM key factors will drive benefits such as, workers' commitment, and loyalty can enhance customers' engagement and patriotism; elimination of waste, by becoming environmentally-friendly; increase in revenue, profitability, and shareholder value. The conclusions also advise that technology supported healthcare TQM services will increase patient satisfaction, loyalty, positive word of mouth, company revenue and profitability. According to Ansari (2020), SQ enhances workers' pride, efforts, develops customers' loyalty, contentment, which promotes repeat purchase behaviour.

To reach that stage, one must bring in suitable change of mindset to accept change and ready to learn more, engage in developing structures, like QC & CI, and acquire professional knowledge, without which one cannot reap TQM benefits. Thus, the present research affirms that TQM enactment in healthcare will provide improved accomplishment when aptly executed in a right condition, with committed leadership, quality circle in place, supportive infrastructure, and with total quality improvement attitude.

From a practical standpoint, the present TQM framework suggests that healthcare providers can figure out realistic factors of TQM implementation through overcoming TQM barriers, putting special emphasis to drive its benefits. Armed with a clear-cut TQM framework, healthcare managers can devise better strategies to keep at bay some of the key challenges and overcome the stubborn barriers to reach a favourable, conducive state, from where one can proceed to get at the "prized" benefits.

A perceived limitation of this study is perhaps due to poor understanding of TQM. However, detailed interviews were conducted with a limited set of respondents that may not represent the view of the complete population. Thus, future research using Structural Equation Modelling with an empirical validation is called for, which can also investigate Lean Manufacturing and Six Sigma combined as LSS to find out how far the healthcare industry can benefit from them.

ACKNOWLEDGEMENTS

Gratitude goes to medical professionals and patients of Oman for expressing their views and sharing their opinions on the benefits and pitfalls of TQM practiced in the healthcare sector in Oman. It is also to declare that there is no assistance or financial support taken from any party and there is conflict of interest.

REFERENCES

Albliwi, S.A., Antony, J., Lim, S.A.H. and Wiele, T.V.D., 2014. Critical failure factors of Lean Six Sigma: a systematic literature review. *International Journal of Quality & Reliability Management*, 31(9), pp.1012-1030. DOI: 10.1108/IJQRM-09-2013-0147.

Al-Dhaafri, H.S., Zien, R. and Al-Swidi, A.K., 2013. The Effect of Total Quality Management, Enterprise Resource Planning and the Entrepreneurial Orientation on the Organizational Performance: The Mediating Role of the Organizational Excellence --- A Proposed Research Framework. *International Journal of Business Administration*, 4(1), pp.66-85. DOI:10.5430/ijba.v4n1p66.

Ali, F., Jain, R., Ali, L. and Munir, K., 2017. The Effects of Quality Practices on the Performance Measurement of Business Management. *Journal of Advanced Management Science*, 5(6), pp.440-444. DOI: 10.18178/joams.5.6.440-444.

Ali, K.A.M. and Alolayyan, M.N., 2013. The impact of total quality management (TQM) on the hospital's performance: an empirical research. *International Journal Services and Operations Management*, 15(4), pp.482-506. DOI: 10.1504/IJSOM.2013.054904.

Alzoubi, M.M., Hayati, K.S., Rosliza, A.M., Ahmad, A.A. and Al-Hamdan, Z.M., 2019. Total quality management in the health-care context: integrating the literature and directing future research. *Risk Management and Healthcare Policy*, 12, pp.167-177. DOI: 10.2147/RMHP.S197038.

Ansari, M.S.A., 2020. Extended Service-Profit Chain for Telecom Service Industry in Oman: An Empirical Validation. *Sustainable Futures*, 2, 10pp. DOI: 10.1016/j.sftr.2020.100032.

Ansari, M.S.A., Farooquie, J.A. and Gattoufi, S.M., 2018. Emotional Intelligence and Extended Service Profit Chain in Telecom Emotional Intelligence and Extended Service Profit Chain in Telecom. *International Business Research*, 11(3), pp.133-148. DOI: 10.5539/ibr.v11n3p133.

Antony, J., Palsuk, P., Gupta, S., Mishra, D. and Barach, P., 2018. Six Sigma in healthcare: a systematic review of the literature. *International Journal of Quality & Reliability Management*, 35(5), pp.1075-1092. DOI: 10.1108/IJQRM-02-2017-0027.

Antony, J., Snee, R. and Hoerl, R., 2017. Lean Six Sigma: yesterday, today and tomorrow. *International Journal of Quality & Reliability Management*, 34(7), pp.1073-1093. DOI: 10.1108/IJQRM-03-2016-0035.

Balasubramanian, M., 2016. Total Quality Management in the Healthcare Industry, Challenges, Barriers and Implementation Developing a Framework for TQM Implementation in a Healthcare Setup. *Science Journal of Public Health*, 4(4), pp.271-278.

Bhamu, J. and Sangwan, K.S., 2010. Lean manufacturing: literature review and research issues. *International Journal of Operations & Production Management*, 34(7), pp.876-940. DOI: 10.1108/IJOPM-08-2012-0315.

Cameron, R. and Azorin, J.F.M., 2011. The acceptance of mixed methods in business and management research. *International Journal of Organizational Analysis*, 19(3), pp.256-271. DOI: 10.1108/19348831111149204.

Candido, C.J.F. and Santos, S.P., 2011. Is TQM more difficult to implement than other transformational strategies. *Total Quality Management journal*, 22(11), pp.1139-1164. DOI: 10.1080/14783363.2011.625185.

Carnerud, D. and Bäckström, I., 2019. Four decades of research on quality: summarizing, trendspotting and looking ahead. *Total Quality Management & Business Excellence*, 32(9-10), pp.1023-1045. DOI: 10.1080/14783363.2019.1655397.

Choomlucksanaa, J., Ongsaranakorna, M. and Suksabaia, P., 2015. Improving the productivity of sheet metal stamping subassembly area using the application of lean manufacturing principles. Procedia Manufacturing, 2, pp.102-107. DOI: 10.1016/j.promfg.2015.07.090.

Creswell, J.W., 1998. *Qualitative inquiry and research design: Choosing among five traditions.* Thousand Oaks, CA: Sage.

Dahlgaard, J.J., Pettersen, J. and Park, S.M.D., 2011. Quality and lean health care: A system for assessing and improving the health of healthcare organisations. *Total Quality Management & Business Excellence*, 22(6), pp.673-689. DOI: 10.1080/14783363.2011.580651.

Dale, B.G., Wiele, T.V.D. and Iwaarden, J.V., 2013. *Managing quality*. 5th ed. Blackwell Publishing.

Deming, E., 1986. *Out of the crisis.* Cambridge, Mass.: Massachusetts Institute of Technology, Centre for Advanced Engineering Study.

Dénes, R.V., Kecskés, J., Koltai, T. and Dénes, Z., 2017. The Application of Data Envelopment Analysis in Healthcare Performance Evaluation of Rehabilitation Departments in Hungary. *Quality Innovation Prosperity*, 21(3), pp.127-142. DOI: 10.12776/qip.v21i3.920.

Deros, B.M., Yusof, S.M. and Salleh, A.M., 2006. Perceptions and practices of critical success factors in benchmarking implementation in Malaysian automotive manufacturing companies. *Journal of Quality Measurement and Analysis*, 2(1), pp.45-61.

Evans, J.R. and Lindsay, W., 2001. *The management and control of quality*. Mason, Ohio: South-Western Cengage Learning.

Fatima, T., Malik, S.A. and Shabbir, A., 2018. Hospital healthcare service quality, patient satisfaction and loyalty: An investigation in context of private healthcare systems. *International Journal of Quality & Reliability Management*, 35(6), pp.1195-1214. DOI: 10.1108/IJQRM-02-2017-0031.

Guetterman, T.C., 2015. Descriptions of sampling practices within five approaches to qualitative research in education and the health sciences. *FORUM: Qualitative Social Research*, 16(2), 25p.

Hackman, J. R., and Wageman, R., 1995. Total quality management: Empirical, conceptual, and practical issues. *Administrative Science Quarterly*, 40,(2), pp.309-342. DOI: 10.2307/2393640.

Hassan, M., Mukhtar, A., Qureshi, S.U. and Sharif, S., 2012. Impact of TQM Practices on Firm's. *Performance of Pakistan's Manufacturing Organizations*, 2(10), pp.232-259.

Ishikawa, K., 1985. *What Is Total Quality Control? The Japanese Way*. Englewood Cliffs, N.J.: Prentice-Hall.

Izogo, E.E. and Ogba, I.E., 2015. Service quality, customer satisfaction and loyalty in automobile repair services sector. *International Journal of Quality & Reliability Management*, 32(3), pp.250-269. DOI: 10.1108/IJQRM-05-2013-0075.

Joiner, T.A., 2007. Total quality management and performance. The role of organization support and co-worker support. *International Journal of Quality & Reliability Management*, 24(6), pp.617-627. DOI: 10.1108/02656710710757808.

Juran, M.J., 1989. *Juran's Quality Handbook*. McGraw-Hill Companies.

Kajihara, C., Munechika, M., Kaneko, M., Sano, M. and Jin, H., 2016. A Matrix of the Functions and Organizations that Ensure Continued Healthcare Services in a Disaster. *Quality Innovation Prosperity*, 20(2), pp.145-156. DOI: 10.12776/qip.v20i2.747.

Kassinis, G.I., and Soteriou, A.C., 2003. Greening the service profit chain. The impact of environmental management practices. *Production and operation management*, 12(3), pp.386-403. DOI: 1059-1478/03/1203/386$1.25.

Khan, M.N., Malik, S.A. and Janjua, S.Y., 2008. Total Quality Management practices and work-related outcomes, A case study of higher education institutions in Pakistan. *International Journal of Quality & Reliability Management*, 36(6), pp.864-874. DOI: 10.1108/IJQRM-04-2018-0097.

Kim, D.Y., Kumar, V. and Murphy, S.A., 2010. European Foundation for Quality Management Business Excellence Model: An integrative review and research agenda. *International Journal of Quality & Reliability Management*, 27(6), pp.684-701. DOI: 10.1108/02656711011054551.

Kumar, V., Sharma, R.R.K., Lai, K.K. and Chang, Y.H., 2018. Mapping the TQM, Implementation, An empirical investigation of the cultural dimensions with different strategic orientation in Indian firms. *Benchmarking: An International Journal*, 25(8), pp.3081-3116. DOI: 10.1108/BIJ-06-2017-0150.

Morse, J.M., 1994. Designing funded qualitative research. In: N.K. Denzin and Y.S. Lincoln, eds. *Handbook of qualitative research.* Thousand Oaks, CA: Sage. pp.220-235.

Mosadeghrad, A.M., 2013. Obstacles to TQM success in health care systems. *International Journal of Health Care Quality Assurance*, 26(2), pp.147-173. DOI: 10.1108/09526861311297352.

Muñoz, C.A. and Gutierrez, L.J., 2017. Six Sigma and organizational ambidexterity: a systematic review and conceptual framework. *International Journal of Lean Six Sigma*, 8(4), pp.436-456. DOI: 10.1108/IJLSS-08-2016-0040.

Muraliraj, J., Zailani, S., Kuppusamy, S. and Santha, C., 2018. Annotated methodological review of Lean Six Sigma. *International Journal of Lean Six Sigma*, 9(1), pp.2-49. DOI: 10.1108/IJLSS-04-2017-0028.

Nakhai, B. and Neves, J.S., 2009. The challenges of six sigma in improving service quality. *International Journal of Quality & Reliability Management*, 26(7), pp.663-684. DOI: 10.1108/02656710910975741.

Oakland, J, S., 2003. *Total Quality Management text with cases.* 3rd ed. Oxford: Burlington, MA: Butterworth-Heinemann.

Ozdal, M.A. and Oyebamiji, B.F., 2018. Implementation of Total Quality Management and its Effect on Employees' Performance in a Teaching Hospital in Oyo State, Nigeria. *Public Health Open Access*, 2(3), pp.1-8. DOI: 10.23880/phoa-16000129.

Parasuraman, A., 2010. Service productivity, quality and innovation. *International journal of quality and service science*, 2(3), pp.277-286. DOI: 10.1108/17566691011090026.

Parasuraman. A. and Grewal, D., 2000. The impact of Technology on the Quality-Value-Loyalty Chain: A research agenda. *Journal of the Academy of Marketing Science*, 28, pp.168-174. DOI: 10.1177/0092070300281015.

Parasuraman. A., Zeithaml, V.A. and Berry, L.L., 1985. A conceptual model of service quality and its implications for future research. *Journal of marketing*, 49(3), pp.41-50.

Prajogoa, D.I. and Hong, S.W., 2008. The effect of TQM performance in R&D environments: A perspective from South Korean firms. *Technovation*, 28(12), pp.855-863. DOI: 10.1016/j.technovation.2008.06.001.

Psomas, E., Vouzas, F. and Kafetzopoulos, D., 2014. Quality management benefits through the "soft" and "hard" aspect of TQM in food companies. *The TQM Journal*, 26(5), pp.431-444. DOI: 10.1108/TQM-02-2013-0017.

Rahman, M., 2019. A Literature Review Based Analysis of Total Quality Management (TQM) Implementation towards Quality Improvement in Bangladeshi Hospitals. *International Journal of Progressive Sciences and Technologies (IJPSAT)*, 17(2), pp.11-16.

Rust, T.R., Zahorik, A.J. and Keiningham, T.L., 1995. Return on Quality (ROC): Making Service Quality Financially Accountable. *Journal of Marketing*, 59(2), pp. 58-70. DOI: 10.2307/1252073.

Sadikoglu, E. and Olcay, H., 2014. The Effects of Total Quality Management Practices on Performance and the Reasons of and the Barriers to TQM Practices in Turkey. *Advances in Decision Sciences*, 2014, 17p. DOI: 10.1155/2014/537605.

Sampaio, P., Saraiva, P. and Domingues, P., 2012. Management systems: integration or addition?. *International Journal of Quality & Reliability Management*, 29(4), pp.402-424. DOI: 10.1108/02656711211224857.

Sampaio, P., Saraiva, P. and Rodrigues, A.G., 2011. The economic impact of quality management systems in Portuguese certified companies' Empirical evidence. *International Journal of Quality & Reliability Management*, 28(9), pp.929-950. DOI: 10.1108/02656711111172522.

Sila, I. and Ebrahimpour, M., 2002. An investigation of the total quality management survey-based research published between 1989 and 2000. *International Journal of Quality & Reliability Management*, 19(7), pp.902-970. DOI: 10.1108/02656710210434801.

Slimák, I. and Zgodavova, K., 2011. Focus on Succes. *Quality Innovation Prosperity*, 15(1), pp.1-4. DOI: 10.12776/qip.v15i1.36.

Taddese, F. and Osada, H., 2010. Process Techno - Innovation Using TQM in Developing Countrie, Empirical Study of Deming Prize Winners. *Journal of Technology and Management Innovation*, 5(2), pp.46-65. DOI: 10.4067/S0718-27242010000200005.

Talib, F., Rahman, Z. and Qureshi, M.N., 2010. Pareto analysis of total quality management factors critical to success for service industries. *International Journal for Quality research*, 4(2), pp.155-168.

Yee, R.W.Y., Yeung, A.C.L. and Cheng, C.E.C., 2011. An empirical Analysis in high contact service industries. *International Journal of Production Economics*, 130(2), pp.236-245. DOI: 10.1016/j.ijpe.2011.01.001.

Yusof, S.M. and Aspinwall, E., 2001. Case studies on the implementation of TQM in the UK automotive SMEs. *International Journal of Quality & Reliability Management*, 18(7), pp.722-743. DOI: 10.1108/02656710110396058.

Zgodavova, K. and Colesca, S.E., 2007. Quality Management Principles – An Approach In Healthcare Institutions. *Management and Marketing Journal*, 5(1), pp.31-38.

ABOUT THE AUTHOR

Mohammad Sultan Ahmad Ansari[0000-0001-7064-3245] – Assist. Prof., Faculty of Business and Economics, Modern College of Business & Science (MCBS), Muscat, Oman, e-mail: mohammad.ansari2040@gmial.com.

CONFLICTS OF INTEREST

The author declares 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 Influencing the Adoption of New Information Technology Supporting Healthy Life Style: The Example of Wearable Self-Tracking Devices

DOI: 10.12776/QIP.V26I1.1612

Stanislava Grosová, Olga Kutnohorská, Marek Botek

Received: 2021-09-26 Accepted: 2022-02-08 Published: 2022-03-31

ABSTRACT

Purpose: The presented study aims to identify key factors affecting the adoption of wearable electronics in Czech women. The results of the study can give insight on how to design an optimized wearable self-tracking device.

Methodology/Approach: The well-established Unified Theory of Acceptance and Use of Technology (UTAUT) framework served as a baseline for research on key determinants of behavioral intention to use wearable self-tracking devices.

Findings: The strongest factor was identified as the habit. The second strongest predictor affecting behavioral intention and use was the construct of performance expectancy. Personal health motivation, as a factor reflecting the nature of the subject examined, was the third strongest factor. The determinants of price value, effort expectancy, and social impact influence the adoption and use of these products. Facilitating conditions, personal inovativeness, personal control over diet and hedonic motivations did not play a significant role.

Research Limitation/Implication: The tested sample included 808 interviewed women, but only from the Czech Republic. Scale already defined Eating control behavior as a measure of healthy lifestyle in terms of eating was the first usage in UTAUT 2 model.

Originality/Value of paper: The study aims primarily to uncover the determinants of the usage of wearable electronics. Secondarily, it extends the theoretical framework of UTAUT2 by testing personal factors such as personal inovativeness, personal motivation to health, and personal control of eating as variables explaining behavioral intention and usage.

Category: Research paper

Keywords: wearable self-tracking devices; UTAUT 2; healthy lifestyle

1 INTRODUCTION

According to the Czech Statistical Office, searches for the word health on the Internet increased by more than 33% between 2007 and 2017 (CZSO, 2018b). At the same time, the Czech Statistical Office states that in April 2018 47% of men and 33% of women suffered from being slightly overweight; obesity then affected almost 20% of men and 18% of women and is still growing (CZSO, 2018a). Generally, people are clearly increasingly interested in their health, but their interest in itself does not bring the desired results, rather the contrary. These two contradictory information points to considerable market potential for companies developing new innovations in the industry. However, it will be important for their success to uncover the determinants that will make potential customers use the new healthy lifestyle-oriented information technology.

In the area of a healthy lifestyle, modern information technologies are increasingly being used to monitor diet, exercise, and personal well-being. They are most often of two types, wearable electronics (wearable self-tracking devices, WSTD) and smart phone applications. There are a number of healthy lifestyle-oriented applications, primarily nutritional applications, weight diary, various fitness applications, or applications that promote relaxation, good sleep, and mental hygiene.

In general, the acceptance of innovative tools has been subjected to intensive research in the last three decades. Several theories have been created to study the reasons for accepting or rejecting innovation. The most recent is "The Unified Theory of Acceptance and Use of Technology" (UTAUT2) introduced by (Venkatesh et al., 2003) can be currently considered as a well-established validated model originally created to study acceptance of new technologies.

Although UTAUT2 was originally designed to study the acceptance of technological innovations in companies, the authors recommend that further research should focus on testing the constructs of UTAUT2 within different contexts, technologies, ages, and geographic locations. A large number of applications and replications of the entire model or its part should contribute to the universal applicability of the model (Venkatesh, Thong and Xu, 2012).

2 METHODOLOGY

The presented research aims to identify what determinants would be considered the most important in relation to the intention to use wearable self-tracking devices. Therefore, an adapted research model based on UTAUT2 was designed to reflect the specific characteristics of these products.

2.1 Research Model

An application of all independent determinants of UTAUT2 in terms of applicability and appropriateness in the context of WSTD has been reviewed.

Following literature review, five in-depth individual interviews with WSTD users were conducted for this work to determine, which variables will affect the WSTD usage and to ensure the questionnaire intelligibility to Czech users. The first two independent variables of the model include expected performance and expected effort. The authors (Venkatesh et al., 2003) identified that the particularly expected performance construct, which is defined as a measure of the belief that the use of technology will help the user achieve improved performance, is the strongest determinant affecting behavioral intent. This is confirmed by studies by Okumus, Bilgihan, and Ozturk (2016), Yuan et al. (2015), which have already explored healthy lifestyle technologies. The construct of the expected effort, described as the degree of simplicity associated with the user's use of the technology, should also be preserved in this work. As a healthy lifestyle is currently a major trend, it is expected that social impact will significantly influence behavior intent. Many WSTDs also allow users to share their activities and progress. All of five in-depth individual respondents confirmed that society, whether friends, family, doctors, or health professionals, has had a major impact on its use. Four of the respondents even share their values with their loved ones or with social networks.

The construct of conditions facilitating the use of technology describes the degree of the user's conviction that the necessary resources and technical support for the use of technology exist. It can be assumed that the better the conditions to facilitate the use of fitness watches, the higher the intention to use them. Under these terms we mean compatibility with another device with which the watch is connected (mobile, tablet, or computer), expertise for their use (orientation in terms that evaluate the watch), and also the possibility to contact someone if they would had any problems.

As already mentioned, the price is expected to have a significant impact on the consumer's intention to use WSTD. On the other hand, there are many models that range from 30 to more than 770 EUR, so it is up to the consumer which features she or he is willing to invest. The study (Gao, Li and Luo, 2015) points out that the price/performance ratio is not the only thing the consumer deals with when buying wearable electronics. Since the user can wear the watch continuously or wear it with increased activity, much more attention is paid to comfort. At the same time, however, it is also an everyday accessory and, therefore, design is an important factor influencing product purchase.

The construct of hedonic motivation, which is described as a measure of pleasure and enjoyment resulting from the use of technology, is also left in this work. In-depth individual interviews have found that users see fitness watches not only as a technology to help them monitor their activities and meals, but also consider them a "modern toy". Respondents confirmed that they like to earn points to achieve their goals and that they like to compete with their loved ones. Subsequently, individual interviews found that the use and monitoring of their activities and diet became a habit. Habit describes the extent to which the user tends to perform the behavior automatically based on what has been learned.

Three constructs were also added to this work, personal innovativeness, motivation for personal health, and personal control over eating. In accordance with the authors (Agarwal and Prasad, 1997) defined as a user's willingness to try and use innovative information technology. In many studies, the construct of personal innovativeness has been added and has often significantly influenced the intention to use the technology (Escobar-Rodriguez, Carvajal-Trujillo and Monge-Lozano, 2014; Shamim et al., 2018). Personal motivation for health describes how an individual perceives health and what exactly is a fitness watch used for. In this case, it is a tool for reducing weight, monitoring sports activities, and monitoring health conditions. The last construct, personal control of eating, focuses more on a healthy lifestyle in terms of eating. It is based on an already defined scale "Eating Control Behavior" (King and Bruner, 2000).

Based on the facts discussed, a research model has been proposed to determine which independent variables will primarily affect the dependent construct of behavioral intention due to the continuous use of WSTD. The framework of this research model can be found in Fig. 1.

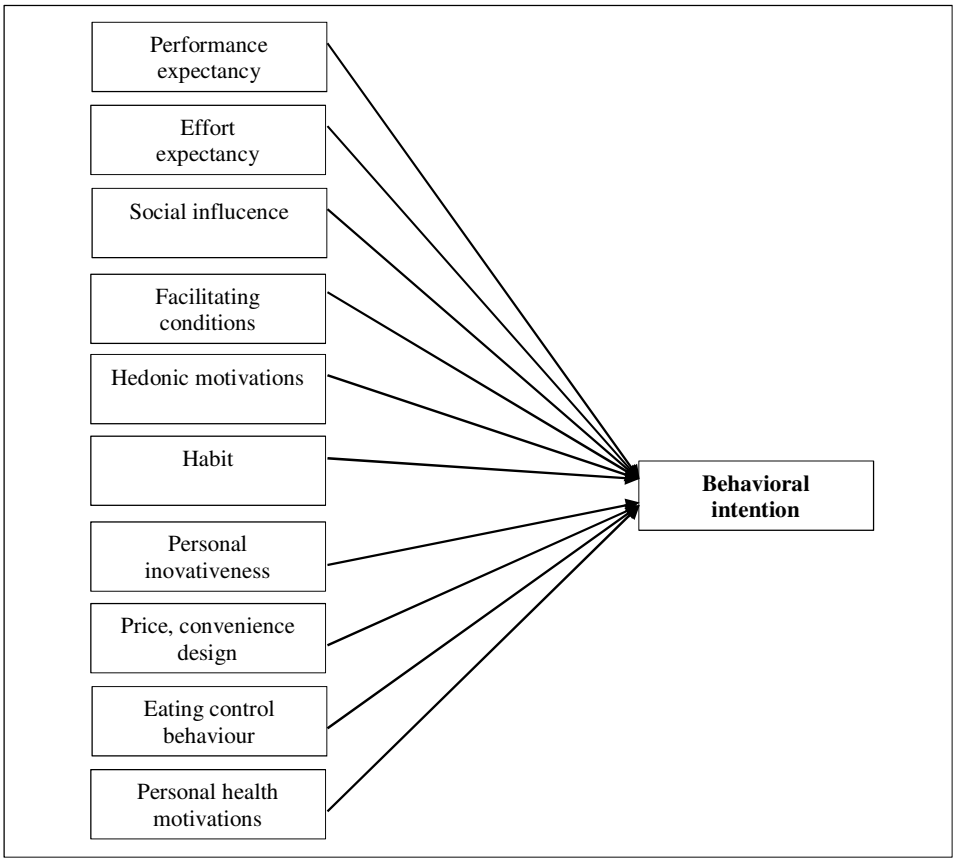


Figure 1 – Research Model

2.2 Research Hypotheses

Based on the facts discussed and on the proposed research model, ten specific hypotheses were determined as follows.

H_{01}: Performance expectation (PE) will positively affect behavioral intentions (BI) to continuous use of WSTD.

H_{02}: Effort expectation (EE) will positively affect behavioural intentions to use WSTD.

H_{03}: Social influence (SI) will positively affect behavioral intentions to use WSTD.

H_{04}: Facilitating conditions (FC) will positively affect the intentions of behavioral use of WSTD.

H_{05}: Hedonic motivations (HM) will positively affect the behavior intentions to use WSTD.

H_{06}: Habit (H) will positively affect the behavioral intentions to use WSTD.

H_{07}: Personal innoveness (PI) will positively affect the behavioral intentions to use WSTD.

H_{08}: The price value ratio (PVL) will positively affect the behavioral intention to use WSTD.

H_{09}: Eating control behavior (ECB) will positively affect the behavior expectation to use WSTD.

H_{10}: Personal motivation to health (PHM) will positively affect the behavioral intentions to use WSTD.

Null hypothesis H_0 was stated for all proposed hypotheses. The null hypothesis assumes that the latent variable does not affect the independent variable (behavioral intention). To decide whether the null hypothesis is supported or not, the relevant null hypothesis needs to be rejected.

2.3 Research Methodology

The methodological approach consisted of several steps following the proposal of the research model, the design of the final questionnaire, the collection of data, and the statistical evaluation. Data for the evaluation of the proposed research model were intended to be collected via an online questionnaire. The final questionnaire included UTAUT2 questions and descriptive questions related to the respondent and its behavior in relation to lifestyle.

Each construct of the proposed research model was represented by questions related to the specific construct in the UTAUT2 part of the questionnaire, see Appendix A. The list of questions was designed mainly from the original questionaire items of the UTAUT2 study (Venkatesh, Thong and Xu, 2012). The

questions were translated from the English original and adapted to the topic of WSTD (fitness watches). New variable personal innovativeness questions were taken from the study "Impact of Personal Innovativeness" (Tor and Øystein, 2005). Questions related to motivation constructs for personal health and personal control of diet were taken from the Marketing Scales Handbook (Bruner, 2013). Translating UTAUT2 statements from English required several consultations to avoid misinterpretation. Prior to launching the questionnaire, a sample of 10 fitness watches users was piloted. It has been confirmed that the statements are formulated in a clear way. Seven-point Likkert scale was used to measure the questionnaire items. The online survey has been conducted since March 2019 until May 2019. People living in the Czech Republic were addressed, especially those who are interested in a healthy lifestyle (whether about diet, exercise, sports activities, or mental hygiene). All respondents owned fitness watches and used some fitness applications. The online survey reached 909 respondents (808 women, 101 men). No one questionnaire had to be excluded due to invalid answers, so 909 questionnaires should be further evaluated. Only data and a model for women are presented. Of the women who responded, 285 (35.4%) were under 20 years of age, 324 (40%) were between 21 and 30 years of age, and 118 (14.6 %) were between 31 and 40 years of age.

3 DATA ANALYSIS AND RESULTS

Structural evaluation modelling was performed to evaluate the relationship of independent variables towards behavioral intention. According to Hair et.al. (2017) partial least squares structural equation modelling (PLS-SEM) is recommended to test the model with the expected high amount of interactions among factors. As all conditions were fulfilled, SmartPLS 3.0 software was employed to assess the gathered (Ringle, Wende and Becker, 2015). A bootstrapping method was used to evaluate the hypothesized relationships. Factor loading and cross-loading were evaluated primarily in line with recent scientific literature. Factor loadings are supposed to be greater than 0.6-0.7 (Chin, Marcolin and Newsted 2003; Hair et al., 2017). This assumption ensures the reliability of the questionnaire items on each construct. Only one construct item (BI3) had not fulfilled the required condition and it was excluded of the experiment. In order to evaluate impact of BI3 on the construct, Cronbach's Alpha of two constructs (with and without BI3) was compared. The results proved BI3 decreases Cronbach's Alpha, therefore, it was dropped from the construct. Factor cross-loadings and Heterotrait-Monotratit Ratios (HTMT) were used in assesing discriminant validity, which implies that variable (construct) is unique and captures phenomena not represented by other constructs in model (Hair et al., 2017). The cross loadings always did not exceed the loadings in any case and HTMT Ratios were lower than treshhold level 0.85, therefore discriminant validity of the variables can be stated as good.

The adjusted coefficient of determination (adj. R2) reached 0.536, so that the proposed research model explains 53.6% of the variability of the dependent variable "behavioral intention (BI)". The statistical significance of the dependency "behavioral intention BI" on latent variables was evaluated using the p-value (see Tab. 1). For six variables, the p-value is lower than the chosen significance level $\alpha = 0.05$ and their null hypothesis H_0 of independence between variables can be rejected. These are the constructs habits, expected performance, price value, motivations for personal health, expected effort, and social impact.

Table 1 – Overall Results of Hypothesis Testing

Relationship between Variables	**Path values**	***p*-value**	**Results**
H_{01} : PE → BI	0.309	0.000	Supported
H_{02} : EE → BI	0.075	0.008	Supported
H_{03} : SI →BI	-0.075	0.008	Supported
H_{04} : FC → BI	0.035	0.274	Non-supported
H_{05} : PVL → BI	0.111	0.001	Supported
H_{06} : HM → BI	0.006	0.871	Non-supported
H_{07} : H → BI	0.368	0.000	Supported
H_{08} : PI → BI	0.020	0.446	Non-supported
H_{09} : PHM →BI	0.105	0.003	Supported
H_{10} : EBC → BI	-0.046	0.08	Non-supported
$R^2 = 0.548$ adj. $R^2 = 0.536$			

Fig. 2 summarizes the structure of the answers to individual items/questions of statistically significant variables.

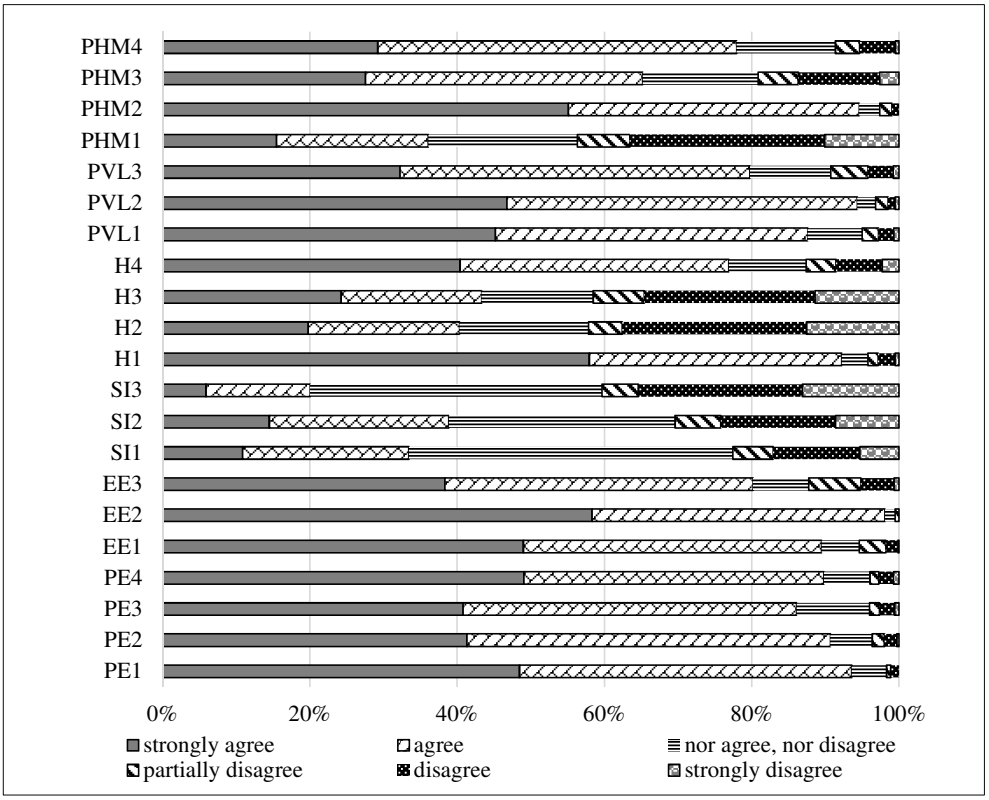


Figure 2 – Structure of Responses of Statistically Significant Variables

4 DISCUSSION

The main aim of our research was the identification of key factors that affect the adoption of wearable electronics among Czech women. The results of our study show that about 55.3% of the variance in the perception of WSTD can be explained by six factors. The most significant factor according to the p-value and the path value is the habit. Respondents strongly agreed with the statement that the use of fitness watches has become a habit and they use it without thinking; see Fig. 2, items H_{01}, H_{04}. The second strongest factor that affects behavioral intention in our study is the construct of performance expectancy. The answers indicate daily use in connection with physical activity. Cost value, convenience, and design factor is the third most influential predictor in our model. From the proportion of agreement with statements, we conclude that users appreciate convenience and design for the price, see Fig. 2, item PVL 2. There is some possibility to modify the price/value ratio. Personal health motivation was newly introduced in the research model. Firstly, the theoretical assumption that the addition of this construct reflecting the nature of the examined subject was confirmed. Second, it was the fourth strongest factor. More than 80% of respondents agree (strongly agree and agree) with statements about the use of

fitness watches in monitoring sport activities and with fitness watches and applications as a part of a healthy lifestyle; see Fig. 2, PHM2 and PHM4 items. This finding emphasizes those characteristics that support sport and fitness usage and can also be used in marketing communication. From the answers to the expected effort control, we deduce that there is a possibility to help the users control more fitness help functions. Some results of this research coincide with the study "An Empirical Study of Wearable Technology Acceptance in Health Care" (Gao, Li and Luo, 2015). In their study, the determinants that most influenced wearable fitness electronics were expected performance, expected effort, social impact, and cost value (hence comfort and de sign) and these are the same determinants that were identified in our research. The construct habit, which was the strongest factor in our research, has not been studied in their study. Similarly, personal health motivations and control of eating behavior were not studied in their study.

Surprisingly, the hedonic motivation construct was not proved according to the p-value. According to the answers, users have no fun nor find fitness watches very interesting and trendy. In contrast to our assumption, personal control over eating did not play a significant role yet. We recommend taking into account the possibility of controlling eating behavior even though it is not a key determinant. There are more than 500 thousand Czech users of the Caloric table application who use it to track daily food nutrients and energy intake. Connection of fitness watches with this application can bring additional value through tracking not only energy output but also intake and support the motivation to use it repeatedly.

The construct of the facilitation conditions proved to be insignificant. Studies that examined the determinants of the use of technology based on healthy lifestyle agree that conditions that facilitate the use of technology do not affect the behavioral intention to use the technology (Yuan et al., 2015). We agree with the same explanation that due to the increased use of smartphones and applications within it, users no longer find the use of fitness applications or wearable electronics complicated.

The construct of personal inovativeness was incorporated in our model. Although in the studies Okumus, Bilgihan and Ozturk (2016) and Pfeiffer et al. (2016), which examine wearable self tracking electronics, this construct was statistically significant, in our work it proved to be statistically insignificant. In the opinion of our respondents, they feel they are experimentators rather than inovators in the area of wearable self-tracking devices.

5 CONCLUSION

Innovations in new information technologies and products that support healthy lifestyles require an understanding of the factors that affect their adoption and continued use. The market for wearable self-tracking devices is subjected to this

research. Classic methodology was used to create a structural model for the identification of factors influencing the adoption and use of products, which has not yet been the focus of applied research in the Czech Republic. Due to the specific properties of the products and their use, it was necessary to modify the models used in other areas. Specifically, there are motives related to personal health, which have been shown to be important in the adoption and use of the aforementioned technology. The results of the presented study show that more than 50% of the variance can be explained by six constructs, habit, performance expectancy, price-value-convenience design, personal health motivation, effort expectancy, and social influence. Practical significance lies in pointing out other directions that the creators of these devices could take and the development of software applications with which they are equipped. Regarding the practical results, we highlight that habit and performance expectancy are essential for the adoption and continued use of the WSTD.

Regarding theoretical contributions, confirmation of the statistical significance of the newly added construct "personal health motivation" supported that the proposed research models should include at least one construct reflecting the nature of the subject examined. Further research in the area of wearable self-tracking devices could focus first on researching models with a new variable "Technical quality" which should uncover technical features and second on the adoption and use of those products and applications in the age group older than forty years. Partial results of our research showed that penetration in this age group is relatively low.

REFERENCES

Agarwal, R. and Prasad, J., 1997. The role of innovation characteristics and perceived voluntariness in the acceptance of information technologies. *Decision Sciences*, 28(3), pp.557-582. DOI: 10.1111/j.1540-5915.1997.tb01322.x.

Bruner, G., 2013. *Marketing Scales Handbook: Multi-Item Measures for Consumer Insight Research (Volume 7).* Fort Worth, Texas USA: GCBII Productions, LLC.

Chin, W.W., Marcolin, B.L. and Newsted, P.R., 2003. A partial least squares latent variable modeling approach for measuring interaction effects: Results from a Monte Carlo simulation study and an electronic-mail emotion/adoption study. *Information Systems Research*, 14(2), pp.189-217. DOI: 10.1287/isre.14.2.189.16018.

CZSO (Czech Statistical Office), 2018a. Household Income and Living Conditions – 2017. *100 Století Statistiky*, [online] Available at: <https://www.czso.cz/csu/stoletistatistiky/jak-jsou-na-tom-cesi-s-chudobou-obezitou-ci-sportovanim> [Accessed 06 February 2020].

CZSO (Czech Statistical Office), 2018b. *Individuals in EU countries searching for health information on the Internet*, *2017.* [Využívání informačních a komunikačních technologií v domácnostech a mezi jednotlivci - 2018 > 8. Používání internetu k činnostem souvisejícím se zdravím] ČSÚ [online] Available at: <https://www.czso.cz/csu/czso/7-vyhledavani-vybranych-informaci-na-internetu-2gm1e9vx8n> [Accessed 06 February 2020].

Escobar-Rodriguez, T., Carvajal-Trujillo, E. and Monge-Lozano, P., 2014. Factors that influence the perceived advantages and relevance of Facebook as a learning tool: An extension of the UTAUT. *Australasian Journal of Educational Technology*, 30(2), pp.136-151. DOI: 10.14742/ajet.585.

Gao, Y., Li, H. and Luo, Y., 2015. An empirical study of wearable technology acceptance in healthcare. *Industrial Management & Data Systems*, 115(9), pp.1704-1723. DOI: 10.1108/imds-03-2015-0087.

Hair, J.F., Hult, G.M.T., Ringle, C.M. and Sarstedt, M., 2017. *A primer on partial least squares structural equation modeling (PLS-SEM).* 2nd ed. Los Angeles: SAGE Publications, Inc.

King, M. F. and Bruner, G.C., 2000. Social desirability bias: A neglected aspect of validity testing. *Psychology & Marketing*, 17(2), pp.79-103. DOI: 10.1002/(sici)1520-6793(200002)17:2<79::Aid-mar2>3.0.Co;2-0.

Okumus, B., Bilgihan, A. and Ozturk, A.B., 2016. Factors Affecting the Acceptance of Smartphone Diet Applications. *Journal of Hospitality Marketing & Management*, 25(6), pp.726-747. DOI: 10.1080/19368623.2016.1082454.

Pfeiffer, J., Entress-Fuersteneck, M.V., Urbach, N. and Buchwald, A., 2016. Quantify-me: Consumer Acceptance of Wearable Self-tracking Devices. In: ECIS, *24th European Conference on Information Systems (ECIS)*. Istanbul, Turkey, 12-15 July 2016. Augsburg: University of Augsburg, Bayreuth: University of Bayreuth.

Ringle, C.M., Wende, S. and Becker, J.M., 2015. *SmartPLS 3.* Boenningstedt: SmartPLS GmbH, [online] Available at: <http://www.smartpls.com> [Accessed 17 January 2020].

Shamim, M., Chiong, R., Bao, Y. and Malik, B., 2018. Acceptance and use predictors of fitness wearable technology and intention to recommend: An empirical study. *Industrial Management & Data Systems*, 119(1), pp.170-188. DOI: 10.1108/IMDS-01-2018-0009.

Tor, J. L. and Øystein, S., 2005. Impact of Personal Innovativeness on the Use of the Internet Among Employees at Work. Jo*urnal of Organizational and End User Computing (JOEUC)*, 17(2), pp.43-63. DOI: 10.4018/joeuc.2005040103.

Venkatesh, V., Morris, M.G., Davis, G.B. and Davis, F.D., 2003. User acceptance of information technology: Toward a unified view. *Mis Quarterly*, 27(3), pp.425-478. DOI: 10.2307/30036540.

Venkatesh, V., Thong, J.Y.L. and Xu, X., 2012. Consumer Acceptance and Use of Information Technology: Extending then Unified Theory of Acceptance and Use of Technology. *Mis Quarterly*, 36(1), pp.157-178. DOI: 10.2307/41410412.

Yuan, S., Ma, W., Kanthawala, S. and Peng, W., 2015. Keep Using My Health Apps: Discover Users' Perception of Health and Fitness Apps with the UTAUT2 Model. *Telemedicine and E-Health*, 21(9), pp.735-741. DOI: 10.1089/tmj.2014.0148.

ABOUT AUTHORS

Stanislava Grosová[0000-0002-9252-800X] (S.G.) – Assoc. Prof., School of Business, University of Chemistry and Technology, Prague, Czech Republic, e-mail: stanislava.grosova@vscht.cz.

Olga Kutnohorská[0000-0002-5255-8001] (O.K.) – Ph.D., School of Business, University of Chemistry and Technology, Prague, Czech Republic, e-mail: olga.kutnohorska@vscht.cz.

Marek Botek[0000-0002-5941-3040] (M.B.) – Ph.D., School of Business, University of Chemistry and Technology, Prague, Czech Republic, e-mail: marek.botek@vscht.cz.

AUTHOR CONTRIBUTIONS

Conceptualization, S.G. and O.K.; Methodology, S.G.; Software, S.G. and O.K.; Validation, S.G. and O.K.; Formal analysis, S.G.; Investigation, O.K. and M.B.; Resources, S.G.; Data curation, O.K.; Original draft preparation, S.G.; Review and editing, O.K. and M.B.; Visualization, O.K.; Supervision, M.B.

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 – Items/ Survey questions/ for variables measurement

Variable: Item/ Question
PE1: I find the fitness watch useful in my daily life.
PE2: I believe that fitness watches are useful for achieving my fitness.
PE3: I believe that fitness watches motivate me to improve my health.
PE4: I think that fitness watches motivate me to move.
EE1: It was easy for me to put a fitness watch into operation and connect it with other technologies.
EE2: I think that using and handling a fitness watch is easy.
EE3: I think I can control all functions that a fitness watch offers.
SI1: People who are close to me recommend me to use fitness watches.
SI2: People I admire recommend me to use fitness watches (eg. blogger, celebrity).
SI3: Healthy lifestyle experts recommend using a fitness watch (eg doctor, trainer, nutritionist).
FC1: I have the resources necessary to use a fitness watch (financial and technological).
FC2: I have enough expertise to use fitness watches (I know the terms REM, kcal, joul, etc.)
FC3: Fitness watches are compatible with other technologies that I use.
FC4: I know who to address when I have trouble using a fitness watch.
HM1: I will have fun using the fitness watch.
HM2: Using a fitness watch makes my fitness activities more enjoyable.
HM3: I feel good with fitness watches because they are trendy.
H1: Using a fitness watch has become a habit for me.
H2: I am addicted to using a fitness watch.
H3: Without a fitness watch, I feel uncomfortable.
H4: I use fitness watches completely without thinking.
PV1: My fitness watch has a nice design for the price.
PV2: My fitness watch is comfortable for the price.
PV3: I think fitness watches generally have a good price / performance ratio.
BI1: In the future, I plan to continue using fitness watches (to achieve my fitness goals, to monitor my state of health).
BI2: I plan to use fitness watches in my daily life.
BI3: In the future, I would like to use my fitness watch even more.
PHM1: I use fitness watches as a tool to reduce weight.
PHM2: I use fitness watches to monitor my sports activities.

Variable: Item/ Question
PHM3: I use fitness watches to monitor my health.
PHM4: I use fitness watches, a healthy lifestyle is important to me.
PI1: As soon as I hear about a technological innovation, I am looking for a way to try it out.
PI2: Among my friends and colleagues, it is usually me who tries new technologies.
PI3: I like to experiment with new technologies.
ECB1: I watch my total caloric intake on my fitness watch and / or in another mobile application.
ECB2: On my fitness watch and / or in another mobile application, I monitor my total intake of nutrients (carbohydrates, fats, proteins).
ECB3: I regularly check how much I eat.
ECB4: I regularly monitor how much I eat on my fitness watch and / or in another mobile application.

More Accurate Knowledge Search in Technological Development for Robust Parameter Design

DOI: 10.12776/QIP.V26I1.1639

Kosuke Oyama, Masato Ohkubo, Yasushi Nagata

Received: 2022-01-05 Accepted: 2021-03-12 Published: 2022-03-31

ABSTRACT

Purpose: The causality search Taguchi (CS-T) method was proposed to support system selection in a robust parameter design. However, the target of the analysis is likely to be quasi-experimental data. This can be difficult to analyse with the CS-T method. Therefore, this study proposes a new analysis approach that can perform a more accurate knowledge search by applying the instrumental variable.

Methodology/Approach: Using the CS-T method, appropriate knowledge search is difficult with quasi-experimental data, including endogeneity. We examined an analytical process that addresses the endogeneity between mechanism and output by utilizing the control and noise factors that constitute the mechanism as instrumental variables.

Findings: The results show that 1) the proposed method has sufficient practical accuracy, even for quasi-experimental data including endogeneity; and 2) the extracted mechanism is less likely to fluctuate depending on the number of experimental conditions used. Moreover, we can clarify the position of the CS-T and proposed methods in system selection.

Research Limitation/Implication: We perform estimation under the assumption that the threshold is known. However, the extracted mechanism may change depending on the threshold; this requires discussing how to determine them.

Originality/Value of paper: Technological development requires a high degree of engineer sophistication. However, this study's analytical process allows conducting more accurate knowledge search in a realistic and systematic way without requiring a high level of engineer input.

Category: Research paper

Keywords: robust parameter design; Taguchi method; technological development; instrumental variable method; statistical modelling

1 INTRODUCTION

Generally, the developmental activities of quality engineering (Taguchi method) are divided into two areas: technological development and product design. In the Taguchi method, system selection is performed during technological development, whereas in product design, robust parameter design (RPD) and tolerance design are performed. The goal is to achieve both quality improvement and cost reduction. The RPD is a system design method based on experimental design, and it decreases the effects of random and systematic errors that can change the function of the system.

However, it is difficult to achieve quality improvement in RPD in product design if the system selection (technological development) in the front-end process is inadequate. Kawada (2013) reports that the same issue applies to its use in research and development. In many technological developmental processes, the best existing system is selected. Thereafter, the system is developed by improving the output/objective characteristics, such as variance and signal-to-noise ratio (SNR), and obtaining ideas for new control factors and generic functions from the mechanisms that achieve the target values. In other words, knowledge search is very important to understand the relationship between mechanisms and output/objective characteristics.

Hosokawa (2020) cites one-factor experiments and the RPD of the objective function as the knowledge search approaches used in this process.

One-factor experiments can focus on a single mechanism and facilitate its comprehensive analysis. Conversely, besides being inefficient, these experiments reduce the ability to detect mechanisms for improvement because the range of change in the output/objective characteristics and mechanisms is small.

The RPD of the objective function can be performed efficiently to a certain extent, notwithstanding that the range of change in the output/objective characteristics expands by moving many factors. Conversely, despite that the mechanism of improvement is unknown, the reliability of the graph of factorial effects may be compromised by the effects of interactions. According to Hosokawa and Miyagi (2019), only a limited amount of studies has been effectively used, specifically Mashhadi et al. (2016), Gamage, Jayamaha and Grigg (2017), and Göhler, Ebro and Howard (2018).

Thus, there is a growing need for technological development of an efficient and accurate method to understand the mechanism of improvement. Therefore, Hosokawa et al. (2015) proposed the causality search - Taguchi (CS-T) method to solve this problem. This method increases the amplitude of the mechanism and output/objective characteristics by moving the factors assigned to the orthogonal array and estimates the relationship between them using the Ta-method (Inoh et al., 2012) contribution ratio and the overall estimated SNR. The Ta-method is a modification of the Taguchi (T) method.

Because the analysis is performed while repeatedly adding experimental conditions, the experiment can be censored, which holds promise as a methodology for efficient knowledge search. For specific examples, see Hosokawa and Miyagi (2019).

However, the data handled using the CS-T method can be interpreted as quasi-experimental data in which observational and experimental data are mixed. Therefore, problems, such as sensor failure (measurement error) or a mixture of unexpected factors (omitted variables), may occur, and bias due to endogeneity may arise when a regression problem is assumed. Meanwhile, the CS-T method, which employs the Ta-method, a regression technique, is at risk of not extracting the mechanisms that improve the output/objective characteristics because the contribution ratio and the overall estimated SNR are affected.

In this study, we propose an analysis process that can remove the bias caused by endogeneity, which hinders the extraction of mechanisms that improve output/objective characteristics by incorporating the concept of instrumental variables (IVs). Further, the proposed method enables the censoring of experiments as well as the CS-T method, which is expected to facilitate efficient knowledge search.

The remainder of this paper proceeds as follows: Section 2 explains the CS-T method for a static system; Section 3 proposes a new knowledge search method based on the description of the assumed data structure and concepts; Section 4 compares the methods through simulations to clarify their respective positions in technological development; and Section 5 summarizes the study and describes future developments.

2 CAUSALITY SEARCH TAGUCHI METHOD

As previously mentioned, system selection must be performed in the RPD front-end process in product design. Taguchi (1993) states that “the system selection is which system (generic function as a technical means) is chosen as the generic function.” Here, the generic function refers to the technical means of achieving the objective function. As an example of a generic function, Taguchi (2004) cites the chemical reaction of an engine. In conjunction with the determination of generic functions, it is also necessary to develop control factors that constitute the system. This is to determine the limits of improvement of the selected system and evaluate the system properly by moving the control factors.

Subsequently, we outline the CS-T method based on Hosokawa (2020) as an approach to support system selection, using a static characteristic system as an example. For the target data, it is assumed that the relationship between the control factor and the mechanism is known to some extent and that the control factor affects the output/objective characteristics only through the mechanism. In the CS-T method, the factors that represent mechanisms are referred to as effective explanation factors (EEFs), including sensor data, physical property

data, and intermediate data in computer-aided engineering, and they cannot be leveled at the time of the experiment neither do they have a target value. In this study, we interpret the EEFs as an intermediate characteristic and proceed with the discussion.

We use the orthogonal array L_8 with four control factors **x** as the inner array and one error factor **z** as the outer array, assigned at two levels for discussion. When the index $i = 1,2, \ldots, I$ $(I = 8)$ denotes the combination of control factors, $j = 1,2$ denotes the combination of noise factors, and $k = 1, \ldots, r$ denotes the types of EEFs, an example of the data used in the CS-T method with the observed EEFs s_{ijk} and output y_{ij} are shown in Table 1.

Table 1 – Example of the Data Format Used in the CS-T Method

No.	x_A	x_B	x_C	x_D	$z = 1$					$z = 2$				
					s_1	s_2	⋯	s_r	y	s_1	s_2	⋯	s_r	y
1	1	1	1	1	s_{111}	s_{112}	⋯	s_{11r}	y_{11}	s_{121}	s_{122}	⋯	s_{12r}	y_{12}
2	1	1	2	2	s_{211}	s_{212}	⋯	s_{21r}	y_{21}	s_{221}	s_{222}	⋯	s_{22r}	y_{22}
3	1	2	1	2										
4	1	2	2	1										
5	2	1	1	2	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
6	2	1	2	1										
7	2	2	1	1										
$I = 8$	2	2	2	2	s_{I11}	s_{I12}	⋯	s_{I1r}	y_{I1}	s_{I21}	s_{I22}	⋯	s_{I2r}	y_{I2}

The method consists of four steps.

In the first step, the data are created. Control factors and noise factors are selected and assigned to the orthogonal array to increase the range of changes in the EEFs and output/objective characteristics. Thereafter, an arbitrary number of experimental conditions is randomly selected from the orthogonal array, and experiments are conducted to measure the output and EEFs. Here, the EEFs and outputs are converted into objective characteristics as necessary.

In the second step, the explanatory rate R^2 of the entire EEF is calculated by conducting the Ta-method with the objective characteristic as the objective variable and the EEF as the explanatory variable. This serve as an indicator of the termination of the experiment. Note, however, that the Ta-method is essentially a method that assumes that the explanatory variables arise from the objective variables.

In the third step, the EEF is assigned to the orthogonal array of the two-level system as the first level if it is included in the estimation equation created in the second step, and the second level if it is not. The overall estimated SNR $\hat{\eta}$ is

calculated after assigning the factors to the orthogonal array. Thereafter, $\hat{\eta}$ is used to perform an analysis of variance to derive the contribution ratio $\hat{\rho}_k$ of each EEF.

In the fourth step, the decision to terminate the experiment is made. If R^2 is greater than 0.6, and the $\hat{\rho}_k$ of any EEF shows a similar trend more than five times, the experiment is terminated. If the conditions are not met, one unselected experimental condition is randomly selected, the output and the EEFs are measured anew, and the two steps are returned.

3 PROPOSAL OF A NEW METHOD

3.1 Data Structures

Hosokawa et al. (2015) and Hosokawa (2020) do not perform simulations, but only validate the results using real data analysis. Therefore, in this section, we discuss and define the data structure while considering the assumptions of the RPD.

The CS-T method assumes that the control factors affect the output/objective characteristics only through the EEFs. Moreover, the Ta-method is used in the CS-T method. With the aforementioned in mind, we refer to the model of static systems in Myers, Khuri and Vining (1992). Thus, the output y_{ij} is:

$$y_{ij} = \delta_0 + \mathbf{s}_{ij}^T \boldsymbol{\delta} + \varepsilon_{ij},$$
$$\mathbf{s}_{ij} = \begin{pmatrix} s_{ij1} \\ \vdots \\ s_{ijr} \end{pmatrix} \tag{1}$$

where δ_0 is the intercept parameter, $\mathbf{s}_{ij}$ denotes the EEFs, ε_{ij} denotes the error terms, and the r-dimensional vector $\boldsymbol{\delta}$ denotes the vector of coefficients for the control factors. The kth EEF is:

$$s_{ijk} = \beta_{0k} + \mathbf{x}_i^T \boldsymbol{\beta}_k + \mathbf{z}_j^T \boldsymbol{\gamma}_k + \mathbf{x}_i^T \boldsymbol{\Omega}_k \mathbf{z}_j + \varepsilon_{ijk},$$
$$\mathbf{x}_i = \begin{pmatrix} x_{i1} \\ \vdots \\ x_{ip} \end{pmatrix}, \mathbf{z}_j = \begin{pmatrix} z_{j1} \\ \vdots \\ z_{jq} \end{pmatrix}, \mathbf{z}_j \sim N\left(0, \sigma_{\mathbf{z}}^2 \mathbf{I}_q\right) \tag{2}$$

where β_{0k} is the intercept parameter, $\mathbf{x}_i$ denotes the control factors, $\mathbf{z}_j$ denotes the noise factors, the p-dimensional vector $\boldsymbol{\beta}_k$ denotes the vector of coefficients for the control factors, the q-dimensional vector $\boldsymbol{\gamma}_k$ denotes that for the noise factors, the $p \times q$ matrix $\boldsymbol{\Omega}_k$ denotes the matrix of control-by-noise interaction coefficients and ε_{ijk} denotes the error terms. Each level of the noise factor follows a normal distribution $N(0, \sigma_{\mathbf{z}}^2)$. The index $i = 1,2, \ldots, I$ denotes the combination of control factors, and $j = 1,2, \ldots, J$ denotes the combination of noise factors.

In this model, the EEFs consist of control and noise factors, and the output consists of EEFs. In other words, the measured output is an aggregate of multiple EEFs, and changes in the output are always the result of changes in some EEFs. Moreover, from the perspective that RPD uses the interaction between control and noise factors, it is more natural for the control and noise factors to have a direct interaction relationship.

Upon applying to the analysis of Hosokawa et al. (2015), it is implied that variations in control factors and paper features affect the distance the paper is folded from the reference point through the paper transport speed. Hence, by choosing a combination of control factors that attenuate the effects of variations in paper features, the paper transport speed and the paper folded distance from the reference point are stabilized.

3.2 Concept of the Proposed Method

Causal inference is a useful method for estimating relationships between variables. Causal inference methods include Gaussian graphical modelling (Lauritzen, 1996), linear non-gaussian acyclic models (Shimizu et al., 2006), and propensity scores (Rosenbaum and Rubin, 1983). However, they are not applicable to this problem.

Therefore, in this study, we apply the IV concept, which is one of the causal inferences. Kuroki, Miyakawa and Kawata (2003) interpret this in terms of causal diagrams as being able to deal with unobserved covariates (omitted variables). It is also tolerant to measurement errors. See Bowden and Turkington (1985) for further details.

Although the CS-T method uses orthogonal arrays, they should be treated as quasi-experimental data. Therefore, it is possible that the missing variables or measurement errors are included in the EEFs. If this happens, endogeneity will occur, which will cause a bias in the regression coefficients when assuming a regression model. This is because as the EEFs change, the error terms and omitted variables also change, thereby making it impossible to determine the effect of the EEFs on the output.

The IV method can be used to solve this problem. Furthermore, because the data covered in this study use orthogonal arrays and assume that the control factors affect the output/objective characteristics only through the EEFs, making the control factors the IVs facilitates the IV selection, which is normally difficult.

Additionally, there are many EEFs that are not involved in the output of the data targeted in this study, and it is necessary to select variables for them. Therefore, in the proposed method, the EEFs are reduced by using stability selection (Meinshausen and Bühlmann, 2010), which is a variable selection method.

3.3 Analysis Process of the Proposed Method

In the following sections, the proposed method is explained in detail by assuming a static system.

The first step entails data creation, similar to the CS-T case. However, the level of each control factor should be extracted such that they are not all the same. The use of a mixed-level orthogonal array is desirable because the combination of rows to be extracted affects the amplitude of each EEF.

The second step involves the selection of EEFs to be used in the estimation of the relationships in the third step. Here, the output/objective characteristics are the objective variables, and all the EEFs are the explanatory variables. Thereafter, we use lasso regression with the regularization parameter λ. However, selection is performed using stability selection.

For stability selection, we first prepare B bootstrap samples. We define $\Psi\{k \in \hat{S}^{\lambda}(\Psi_b)\}$ as the indicator function in which the coefficient of the EEF of variable number k is 1 if it is non-zero and 0 if it is zero in the b-th bootstrap sample Ψ_b. $\hat{S}^{\lambda}$ denotes the selection result of the lasso regression using λ for Ψ_b. Thereafter the choice probability $\widehat{\Pi}_k^{\lambda}(k = 1, \cdots, p)$ is defined as the sum up to $b = 1, \cdots, B$ divided by B. We also treat the set of EEFs larger than an arbitrary threshold $\pi_{thr}(0 < \pi_{thr} < 1)$ as candidate $\hat{S}^{stable}$, as follows:

$$\hat{S}^{stable} = \left\{k : \max_{\lambda \in \mathbf{\Lambda}} \widehat{\Pi}_k^{\lambda} \geq \pi_{thr}\right\} \tag{3}$$

where $\mathbf{\Lambda}$ denotes an arbitrary set of λ.

The third step is the estimation of the relationship between the EEFs and the output/objective characteristics using the conditional IV method (Brito and Pearl, 2002). The use of an IV is meant to exclude bias by examining the effect of the factors assigned to the orthogonal array on the output/objective characteristics of changes in the EEFs.

Here, we consider $\hat{S}^{stable}$ as the treatment variable. Then, the control and noise factors as well as $\hat{S}^{stable}$ which are not the treatment variables, are considered as $n \times l$ matrix Z, $\hat{S}^{stable}$ as $n \times m$ matrix X, and the output/objective characteristics as $n \times 1$ matrix Y. Y corresponds to the outcome variable.

Thereafter, the effect of the treatment variable on the outcome variable (partial regression coefficient) in the conditional IV method can be written as:

$$\hat{\beta}_{IV} = (X^T P_Z X)^{-1} X^T P_Z Y \tag{4}$$

where P_Z is the projection matrix.

From the above-stated, we calculate $\hat{\beta}_{IV}$ for all elements of $\hat{S}^{stable}$.

The conditional IV method is used to deal with the exclusion restriction that the IV affects the outcome variable only through the treatment variable.

In the fourth step, the decision to terminate the experiment is made. When the third step has been performed three or more times, the results up to the time when censoring is considered and the results up to the previous time are considered as two groups. Next, we extract the results of the EEFs that have the strongest relationship with the output/objective characteristics at the times when we consider censoring. The significance level is then set and an F-test is performed. The null hypothesis is that "the variances of the two groups are equal" and the alternative hypothesis is that "the variances of the two groups are not equal". If the null hypothesis can be rejected, one unselected experimental condition is randomly selected, the output/objective characteristics and EEFs are measured anew, and the second step is returned. If the null hypothesis cannot be rejected, the experiment is terminated.

4 VERIFICATION BY SIMULATIONS

4.1 Simulations Outline

In this section, simulations are conducted to compare the methods. Further, the position of the CS-T and the proposed methods in technological development is clarified. Subsequently, the control factor is **x**, the noise factor is **z**, the EEF is **s**, and the output is y.

4.2 Simulations Settings

Verification is performed by analysing the data of the two patterns generated according to the data structure defined in Section 3.

The first pattern is quasi-experimental data, where the EEFs associated with y contain unobserved covariates u_1 and u_2. The model equation is defined as:

$$s_k = \begin{cases} \beta_{Ak}x_A + \beta_{BZk}x_B z + \beta_{Zk}z + \varepsilon_k \ \ (k = 1,2,\cdots,9) \\ 5.0(x_A + x_B z + z + 2.0u_1) + \varepsilon_k \ \ (k = 10) \\ \beta_{Ck}x_C + \beta_{DZk}x_D z + \beta_{Zk}z + \varepsilon_k \ \ (k = 11,12,\cdots,19) \\ 5.0(x_C + x_D z + z + 2.0u_2) + \varepsilon_k \ \ (k = 20) \\ \beta_{Ek}x_E + \beta_{FZk}x_F z + \beta_{Zk}z + \varepsilon_k \ \ (k = 21,22,\cdots,30), \end{cases} \quad (5)$$

$$y = 3.0s_{10} + 2.0s_{20} + 7.0(-u_1 + u_2) + \varepsilon_{31}, \quad (6)$$

$$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_{31}, u_1, u_2 \sim N(0,1), \quad (7)$$

$$\beta_{Ak}, \beta_{BZk}, \beta_{Ck}, \beta_{DZk}, \beta_{Ek}, \beta_{FZk}, \beta_{Zk} \sim unif(1,10) \quad (8)$$

The second pattern is the experimental data, where s_{10} has the most direct effect with y, while s_{20} has the most total effect. The model equation is defined as:

$$s_k = \begin{cases} \beta_{Ak}x_A + \beta_{BZk}x_B z + \beta_{Zk}z + \varepsilon_k \ \ (k = 1,2,\cdots,9) \\ 5.0(x_A + x_B z + z) + \varepsilon_k \ \ (k = 10) \\ \beta_{Ck}x_C + \beta_{DZk}x_D z + \beta_{Zk}z + \varepsilon_k \ \ (k = 11,12,\cdots,19) \\ 50.0(x_C + x_D z + z) + \varepsilon_k \ \ (k = 20) \\ \beta_{Ek}x_E + \beta_{FZk}x_F z + \beta_{Zk}z + \varepsilon_k \ \ (k = 21,22,\cdots,30), \end{cases} \quad (9)$$

$$y = 3.0s_{10} + 2.0s_{20} + \varepsilon_{31}, \quad (10)$$

$$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_{31} \sim N(0,1), \quad (11)$$

$$\beta_{Ak}, \beta_{BZk}, \beta_{Ck}, \beta_{DZk}, \beta_{Ek}, \beta_{FZk}, \beta_{Zk} \sim unif(1,10) \quad (12)$$

4.3 Method of Analysis

The data used for the simulation were generated using the orthogonal array L_{18}. The factors x_A, x_B, x_C, x_D, x_E, and x_F were assigned to the 2nd, 3rd, 4th, 5th, 6th and 7th columns, respectively. The noise factor z was arranged in the outer array. The first level of the noise factor was 1, and the second level was −1. The first level of the control factor was 1, the second level was 0, and the third level was −1. The number of datasets was 5,000. The initial number of experimental conditions was set to nine.

Columns 1 to 30 of the L_{32} orthogonal array were used for the extraction of EEFs in the CS-T method. The hyperparameters of the adaptive lasso were determined by the leave-one-out using $Mean\ Squared\ Error$ (MSE) as an evaluation criterion. Stability selection was set to $B = 100$, $\mathbf{\Lambda} = (1.3^1, 1.3^2, \cdots, 1.3^{30})$. π_{thr} was varied for each pattern, with 0.95 for patterns 1 and 0.80 for pattern 2, respectively. The significance level of the censoring criterion was set to 5%.

We set accuracy evaluation criteria which is the absolute values of the relationship between y and **s** calculated by each method, sorted in descending order, and the probability that the top two are non-zero and correctly aligned with s_{10} and s_{20}. The above-mentioned evaluation criteria are shown in the bar graph. The leftmost bar shows the results when nine experimental conditions are used, while the rightmost bar shows the results when all experimental conditions are used.

4.4 Simulations Results

First, we focus on the first pattern. As can be seen in Figure 1, the accuracy of the CS-T, adaptive lasso, and SS-LM methods is always less than 50%, which is rather impractical. However, the accuracy of KS-IV is always above 90%, which is a practical result.

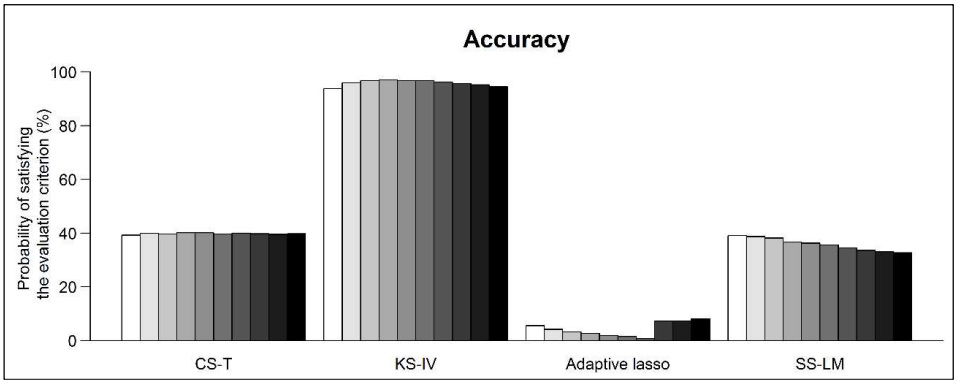


Figure 1 – Pattern 1 Results (Accuracy)

These results show that the CS-T, adaptive lasso, and SS-LM methods are unreliable when omitted variables are mixed. Conversely, the proposed method performs well.

According to the rules of the proposed method and the CS-T method, the probability of the shortest termination is 46.6% at 13 experimental conditions used for the CS-T method and 97.5% at 11 experimental conditions used for KS-IV, indicating that the proposed method can terminate the experiment stably.

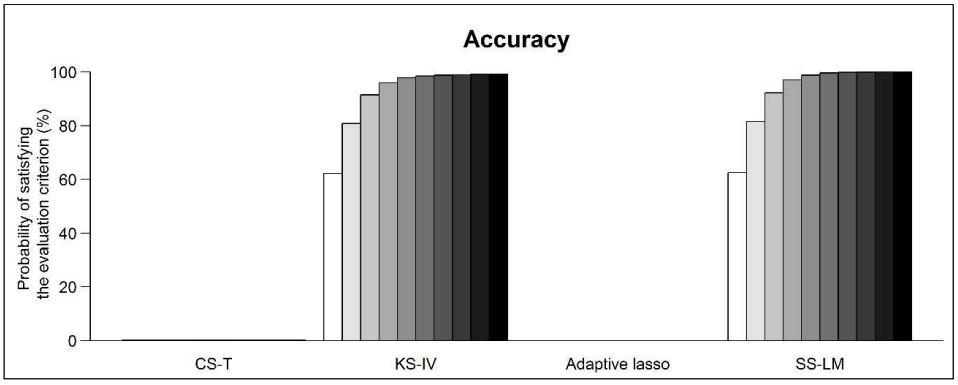


Figure 2 – Pattern 2 Results (Accuracy)

Thereafter, we focus on the second pattern. The rules of the proposed method and the CS-T method, the probability of the shortest termination is 82.8% at 13 experimental conditions used for the CS-T method and 85.5% at 11 experimental

conditions used for KS-IV. Moreover, the accuracy of the proposed method and SS-LM increases with the number of experimental conditions used, as can be seen in Figure 2. However, the CS-T and adaptive lasso methods do not estimate the correct relationship.

These results show that the proposed method performs well even with experimental data where the total and direct effects are different.

Based on the above-stated results, we clarify position of the CS-T and proposed methods in system selection. In models where the total and direct effects differ, the accuracy of the CS-T method is approximately zero. This is because the Ta-method is included, thereby adding the effect from **s** to y to the effect from **x** and **z** to **s**, so that the total effect is estimated.

As a confirmation, the probability that s_{20} has the greatest influence on y is shown in Figure 3. For comparison, the other methods calculate the probability by evaluating the standard partial regression coefficient. Figure 3 shows that the CS-T method has a high probability of determining that s_{20} has the greatest effect on y.

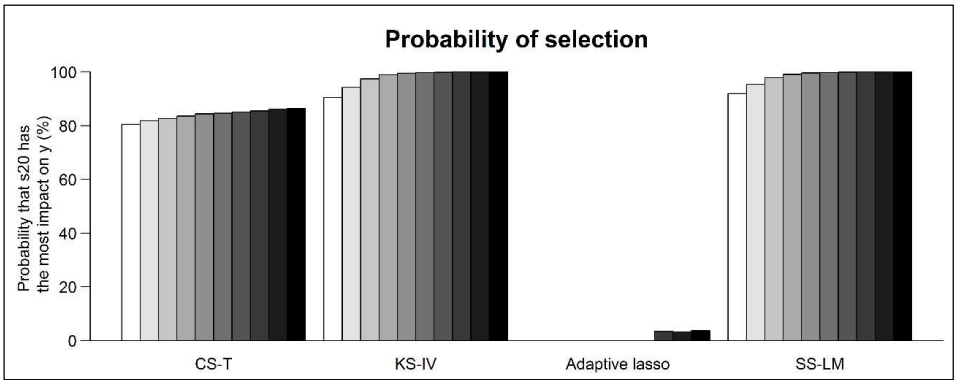


Figure 3 – Pattern 2 Results (Probability that s_{20} has the Most Influence on y)

First, the aim of the proposed and CS-T methods is to convert the extracted **s** into a new control factor or a generic function. For the former, it is always necessary to choose s that has a strong direct effect on y. Ideally, the latter should also consider the effect on **s** when **x** is optimized. In other words, the proposed method is suitable for developing control factors, and the CS-T or the proposed method with standard partial regression coefficients is suitable for developing generic functions.

Nonetheless if there is an omitted variable, it is necessary to choose **s** that has a strong direct effect on y. This is because the effect of **x** and **z** on s is combined with the effect of the omitted variable on **s**.

5 CONCLUSION

The CS-T method proposed by Hosokawa et al. (2015) is a technique for constructing an efficient knowledge search flow for system selection. However, there is room for debate on how to verify the usefulness and deal with the endogenous nature of the data. Therefore, in this study, we attempt to extract mechanisms more accurately and easily by proposing an analysis flow that introduces variable selection by stability selection and the IV concept.

Simulations show that knowledge search using the proposed method has sufficient practical accuracy, even for quasi-experimental data with omitted variables. In addition, the probability of censoring in the shortest time is high. In other words, it can be interpreted that the extracted mechanism is less likely to fluctuate depending on the number of experimental conditions used, and a stable knowledge search can be performed.

Moreover, based on the results of the simulations, we are able to clarify the position of the CS-T and proposed methods in technological development.

Originally, developing control factors and generic functions in technological development requires a high degree of engineer sophistication. However, the analysis process in this study does not require a high level of sense from the engineer, and can be carried out in a realistic and systematic manner. Therefore, it can be expected to establish quality engineering that is both efficient and creative.

Future development entails the determination of the optimal threshold value π_{thr}. Depending on the value of π_{thr}, the EEFs selected in the stability selection will change, which will affect the estimation of the relationship between the EEFs and the output/objective characteristics. In this study, the analysis is conducted with known values, but better results can be expected by changing π_{thr} as the number of experimental conditions used changes.

ACKNOWLEDGEMENTS

We would like to thank the anonymous referees for their valuable comments. This work was partly supported by JSPS Grants-in-Aid for Scientific Research Grant Number 18K11202.

REFERENCES

Bowden, R.J. and Turkington, D.A., 1985. *Instrumental Variables*. Cambridge: Cambridge University Press.

Brito, C. and Pearl, J., 2002. Generalized instrumental variables. *Uncertainty in Artificial Intelligence*, 18, pp.85-93.

Gamage, P., Jayamaha, N.P. and Grigg, N.P., 2017. Acceptance of Taguchi's Quality Philosophy and Practice by Lean practitioners in apparel manufacturing. *Total Quality Management & Business Excellence*, 28(11-12), pp.1322-1338. DOI: 10.1080/14783363.2015.1135729.

Göhler, S.M., Ebro, M. and Howard, T.J., 2018. Mechanism and coherence of robust design methodology: A robust design process proposal. *Total Quality Management & Business Excellence*, 29(3-4), pp.239-259. DOI: 10.1080/14783363.2016.1180952.

Hosokawa, T. and Miyagi, Z., 2019. Quality engineering-based management: a proposal for achieving total optimisation of large systems. *Total Quality Management & Business Excellence*, 30(1), pp.182-194. DOI: 10.1080/14783363.2019.1665843.

Hosokawa, T., 2020. *Taguchi Method for Technology Development CS-T Method for Exploring Generic Functions*. Tokyo, Japan: Nikkagiren.

Hosokawa, T., Okamuro, A., Sasaki, Y. and Tada, K., 2015. A Proposal of Development Methodology Integrating Parameter design and T-Method. *Journal of the Japanese Society for Quality Control*, 45(2), pp.64-72. DOI: 10.20684/quality.45.2_194.

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*, 42(2), pp.265-277. DOI: 10.20684/quality.42.2_265.

Kawada, N., 2013. Application of Quality Engineering in Research and Development. *J-TREC TECHNICAL REVIEW*, 1, pp.38-45.

Kuroki, M., Miyakawa, M. and Kawata, R., 2003. Instrumental Variable(IV)Selection for Estimating Total Effects in Conditional IV Method. *Japanese Journal of Applied Statistics*, 32(2), pp.89-100. DOI: 10.5023/jappstat.32.89.

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

Mashhadi, A.F., Alänge, S., Gustafsson, G. and Roos, L.U., 2016. The Volvo Robust Engineering system: How to make robust design work in an industrial context. *Total Quality Management & Business Excellence*, 27(5-6), pp.647-665. DOI: 10.1080/14783363.2015.1039938.

Meinshausen, N. and Bühlmann, P., 2010. Stability selection. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 72(4), pp.417-473. DOI: 10.1111/j.1467-9868.2010.00740.x.

Myers, R.H., Khuri, A.I. and Vining, G., 1992. Response Surface Alternative to Taguchi Robust Parameter Design Approach. *The American Statistician*, 46(2), pp.131-139. DOI: 10.2307/2684183.

Rosenbaum, P.R. and Rubin, D.B., 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), pp.41-55. DOI: 10.2307/2335942.

Shimizu, S., Hoyer, P.O., Hyvärinen, A. and Kerminen, A., 2006. A linear non-gaussian acyclic model for causal discovery. *Journal of Machine Learning Research: J.M.L.R*, 7, pp.2003-2030. DOI: 10.5555/1248547.1248619#sec-terms.

Taguchi, G., 1993. System Selection. *Journal of Quality Engineering Society*, 1(5), pp.2-8. DOI: 10.18890/qes.1.5_2.

Taguchi, G., 2004. Objective Function and Generic Function (2):—Product Development and Technology Development—. *Journal of Quality Engineering Society*, 12(4), pp.6-13. DOI: 10.18890/qes.12.5_6.

ABOUT AUTHORS

Kosuke Oyama[0000-0002-1195-8172] (K.O.) – Graduate student, Department of Industrial and Management Systems Engineering, Waseda University, Tokyo, Japan, e-mail: whbcbr1000rr@akane.waseda.jp.

Masato Ohkubo[0000-0002-1885-6945] (M.O.) – Lecturer, Department of Business Administration, Toyo University, Tokyo, Japan, e-mail: ohkubo@toyo.jp.

Yasushi Nagata[0000-0002-4947-1204] (Y.N.) – Prof., Department of Industrial and Management Systems Engineering, Waseda University, Tokyo, Japan, e-mail: ynagata@waseda.jp.

AUTHOR CONTRIBUTIONS

Conceptualization, K.O. and M.O.; Methodology, K.O. and M.O.; Software, K.O.; Validation, K.O.; Formal analysis, K.O.; Investigation, K.O.; Resources, K.O.; Data curation, K.O.; Original draft preparation, K.O.; Review and editing, M.O. and Y.N.; Visualization, K.O.; Supervision, M.O. and Y.N.; Project administration, Y.N.; Funding acquisition, Y.N.

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.

Cost-Benefit Analysis of Implementing Circular Economy in a Portuguese Company: From a Case Study to a Model

DOI: 10.12776/QIP.V26I1.1657

Delfina Ramos, Luis Fonseca, Joaquim Gonçalves, Raquel Carvalho, Sandro Carvalho, Gilberto Santos

Received: 2022-02-09 Accepted: 2022-03-19 Published: 2022-03-31

ABSTRACT

Purpose: This work was carried out in a company of the upholstery sector in northern Portugal to demonstrate that Investment in Preventive Measures, within the scope of Environmental Management and supported by a Circular Economy approach, is a significant investment decision, with transversal benefits to the entire organization.

Methodology/Approach: This study focuses on an interview, a financial cost-benefit analysis and a sensitivity analysis. The company is in line with concerns about the Circular Economy.

Findings: The findings support that the cost-benefit analysis is a valuable tool for decision-making and for identifying the potential benefits that can arise from implementing measures from the Circular Economy perspective. An action plan was structured with several recommendations in a logic of action priorities. A case study was carried out in the company and a model was designed.

Research Limitation/Implication: The studied company is still at an embryonic stage. Nevertheless, it was possible to identify all environmental aspects, namely, impacts, consumption. The study in question was carried out only in one company. It can be extended to many more companies.

Originality/Value of paper: The measures to be implemented, in addition to bringing economic and financial benefits to the Company, will contribute to a better and greater environmental sustainability and a better intervention at a social level.

Category: Case study

Keywords: circular economy; environmental management; cost-benefit analysis; sensitivity analysis

1 INTRODUCTION

Circular Economy (CE) can be defined as a regenerative industrial system. It replaces the "end of life" concept of products by regenerating them, in part or in whole. For this purpose, the use of renewable energy is recommended, as well as the elimination of the use of toxic chemical products, which impair the reuse of referred products. This concept aims to eliminate waste through superior design of materials, products and systems, appealing to the use of sustainable materials. It calls for new business models towards sustainability. Hence, the CE economic model aims to conciliate economic and environmental goals and performance. The transition from the traditional economic model "take-make-consume-disposal" to the CE model needs an urgent societal response to tackle environmental problems and promote sustainable (D'Amato, 2021).

Masi, Day and Godsell (2017) performed a systematic review of supply chain configurations regarding CE. Merli, Preziosi and Acampora (2018) identified the following main areas for CE research: social and economic dynamics, firm circular process implementation and related consumption, product design and industrial symbiosis. De Jesus et al. (2018) addressed the role of eco-innovations and their contribution to CE transition. Camacho-Otero, Boks and Pettersen (2018) reviewed consumption related CE research. Sassanelli et al. (2019) focused on the methods for assessing CE. Camón and Celma (2020) under carried a CE review and bibliometric analysis. performed a systematic literature review covering different catalysts, obstacles and ambivalent factors influencing CE implementation in business. Concerning the CE benefits, these authors identified the expected economic and other benefits, the threat for business-as-usual, and the managerial support and existing management systems. To sum up, the transition to CE is a priority theme within the European Union. There is public support for its adoption, and academic research supports the view that it contributes to sustainable development and enduring business performance. Nevertheless, companies consider the economic attractiveness and improvement of profitability and value creation majorly important when addressing CE adoption. Therefore, companies need a support methodology to assess CE cost versus benefit.

This research contributes to CE adoption by performing a cost-benefit analysis of CE implementation in a typical Portuguese Small and Medium Company. The cost-benefit analysis (CBA) compares project implementation's real costs and benefits and updates them to the current moment to make the most beneficial decision (Fonseca et al., 2018; Marques, Guedes and Ferreira, 2017; Ramos, Afonso and Costa, 2020; Ramos, Arezes, P.M. and Afonso, 2017). Thus, CBA is a relevant tool for evaluating, selecting, and analysing CE projects encompassing the financial, economic, social, and environmental dimensions, with the purpose of good management (Bravi et al., 2020; Fonseca, Amaral and Oliveira, 2021; Santos, Rebelo and Santos, 2017; Santos et al., 2021; Rebelo, Santos and Silva, 2015; Carvalho, Santos and Gonçalves, 2020), protecting the environment

(Santos et al., 2014; Teixeira et al., 2021b; Talapatra et al., 2019) and creating value for customers (Santos et al., 2019a; 2019b; Rodrigues et al., 2019).

2 MATERIALS AND METHODS

2.1 Case Study – Interview

In the present study, the interview technique to collect information was used, particularly the semi-structured interview. Regarding the conduct of the interview, an individual and open semi-structured interview was used, based on a previously elaborated script. The interviewee was the managing partner of the organization. The main questions are presented in Table 1.

Table 1 – Interview Script – Main Questions

Main Topics	**Questions**
1. Interviewee Data	Gender: male female; age, Years of experience in the sector;
2. Company Practices in Environmental Issues	• Do you have any specific environmental concerns in your company? What kind of actions do you take to preserve the environment? • What is your waste? What is the destination for your waste? • Is any of the waste likely to re-enter the production of any of your products? • Are you currently using any of the waste that comes from your activity to re-enter the production process? • Does your company recycle any waste?
3.Company Strategy for a CE	• What steps can your company take to achieve a CE and thus reduce or even eliminate waste? • Do you consider the transition to a CE beneficial from an economic and social point of view? Why?

The objective of the interview was to collect qualitative information from the entrepreneur about the Company and the respective sector, with particular emphasis on assessing what potential actions can be implemented to achieve the strategy for the implementation of CE.

2.2 Model Cost-Benefit Analysis

This study was based on and adapted from the model developed by the authors Ramos, Arezes and Afonso (2017) concerning the calculation of the financial cost-benefit ratio in the context of Environmental Management. These authors (Ramos, Arezes and Afonso, 2017) developed a CBA in Occupational Safety and Health (ACBSSO) model, that includes not only financial aspects related to the Company's perspective but also economic aspects (from the perspective of all stakeholders, with particular emphasis on workers and in society).

The B/C ratio is equal to the present value of the project's benefits, divided by the current value of its costs. If the calculated B/C ratio is greater than one, the

project is considered efficient. This equates to the requirement that a Net Present Value (NPV) be positive. A B/C ratio of 1.2, for example, means that the benefits of estimating a project are 20% greater than the costs (Ramos, Arezes and Afonso, 2017). Figure 1 shows the financial B/C ratio in the context of Environmental Management. In this model, the B/C ratio compares avoided costs, measured in terms of reduction of environmental aspects and impacts, and the cost of preventive measures designed and implemented according to the risk analysis.

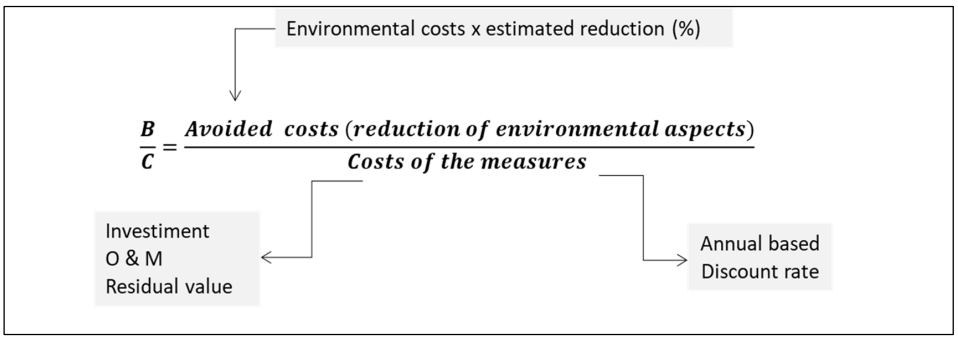


Figure 1 – Schematic representation of the financial Cost-Benefit ratio in the context of Environmental Management, adapted from (Ramos et al., 2017)

CBA is an effective and flexible model for project evaluation based on quantifying inputs and outputs in monetary units. For example, if the result of the ratio is B/C ≥ 1, it means that the benefits are greater than the costs. Therefore, the project should be considered for investment; if B/C < 1, the costs are greater than the benefits. However, the project can be considered if there are unquantified benefits (such as the Company's reputation, employee satisfaction and motivation, among others) and/or positive externalities, such as benefits to society, which arise from quality of products (Costa et al., 2019; Araújo et al., 2019; Sá et al., 2020), new production development methods (Zgodavova et al., 2020; Zgodavova and Slimak, 2008) as well CE (Ramos, Arezes and Afonso, 2017). According to Guide to CBA of Investment Projects 2014-2020 (European Commission, 2014) the standard CBA is structured in seven steps:

1. Description of the context
2. Definition of objectives
3. Identification of the project
4. Technical feasibility & Environmental sustainability (e.g., technical design, cost estimates and implementation schedule)
5. Financial analysis
6. Economic analysis
7. Risk assessment (e.g., sensitivity analysis)

2.3 Sensitivity Analysis of the B/C Ratio

The B/C ratio proposed allows quantifying the benefits of preventive measures on environmental management and as such, environmental aspects for all the stakeholders. It can be an essential support tool to validate, understand and analyse the implementation of preventive measures. The analysis of the B/C ratio will permit to show which preventive measures defined in the risk assessment process are cost-effective and should be implemented. Nevertheless, the quality and availability of environmental aspects data is still a significant limitation to link environmental aspects and business performance (Ramos, Arezes and Afonso, 2017). The B/C ratio is a valuable tool for decision-making related to environmental management. However, it must be emphasized that the manager must understand that the behaviour of the B/C ratio depends on the assumptions made and the contribution of several variables to the result obtained by Silva (2020). The object of the sensitivity analysis is the selection of the critical variables and parameters of the CBA model, that is, those whose variations, positive or negative concerning the value used as the best estimate in the reference case, have a more pronounced effect on the financial parameters and determining economics. Thus, parameters whose variation of, for example, 1% translates into a variation of 5% of the base value of the NPV should be considered.

Assessing a project is not without risks or uncertainties as it is a forecasting exercise. In a sensitivity analysis, the model's impact parameters are determined and subjected to elasticity tests, analysing the reflection in the VAL values. Although the values of this indicator are defined based on forecasts, it may happen that they don't match to the values detected throughout the project. It is essential to carry out tests on the response of forecasts, to possible changes in various parameters (Teixeira et al., 2021a; Carvalho, 2015).

2.4 Company Presentation

The study Company is a Portuguese family micro-enterprise in the upholstery sector, founded in 1984, headquartered in the north of Portugal, and has 9 employees. Its market segment is medium-high quality, operating in 3 different market niches: Upholstery for Decoration, Orthopaedic Chairs for Geriatrics/Health and Articulated Chairs for Private Movie Theatres. Its distribution network in the decoration segment is divided into resale and direct sales to the public; in the remaining segments, distribution is made only through the retail channel.

The Company uses solid wood structure as raw materials, all types of coatings (cotton, PU synthetics, PVC, fabrics with various compositions, natural skins, etc.), foams, threads, fasteners, Velcros, folders, conforel, screws, nuts, metallic mechanisms, motors, among others. As an accessory material, it uses chemical products such as glues, thinners, paints, among others. The Company is not ISO 14001:2015 certified.

3 RESULTS AND DISCUSSION

At this point, the results, and the respective discussion of the interview with the Company's managing partner, financial CBA and sensitivity analysis are presented.

3.1 Interview

The CEO said that its Company has always been concerned with environmental issues about the Company's actions to preserve the environment. They forward all waste to licensed operators, and they internally separate common (urban) garbage. About any of the raw materials they use harmful to the environment, which one(s)? and what do they do to minimize its impact(s)? the CEO reported that both skins and synthetic materials in PVC and foams are harmful to the environment, as they are not biodegradable. The glues are also toxic and are harmful to the environment. However, the Company has a glue cabin with a filter that retains most toxic particles and is not polluting. One possibility would be the replacement with water glues. However, there are still not many options on the market, and the one that exists turns out to be more expensive. When asked if the Company recycles any waste, the CEO said that it is not in its Company. About if there is any concern on the part of the Company in the acquisition of cleaner technology, the CEO related that whenever they purchase new equipment, they are concerned about its energy consumption. The CEO reported that all their products use raw materials of high quality and durability. As a quality product, customers end up having the products for more years, and when they want to change it because they get tired of the model, they try to resell it or give it to family members because the product is still sufficient quality to be used.

In the second group of questions the four questions had the following answers:

1. If they think the transition to a CE is important for the future of companies and the planet, the CEO agrees. However, he thinks that at the level of the Portuguese business fabric, namely micro and small companies, there is still little information on what companies can do to achieve a CE that is economically and environmentally sustainable. Many measures are not economically viable for micro and small businesses. In their case, as an example, they have the replacement of the usual raw material for another that is not harmful to the environment since its cost is much higher, and the consumer is not yet sensitive to this price variation.
2. When asked about what steps their Company can take to achieve a CE and to reduce or even eliminate waste, the CEO related that they never thought about that situation very much. They will have to see what possible measures they can take to minimize the impacts of their activity on the environment and thus achieve a CE.

3. When asked if the transition to a CE is beneficial from an economic and social point of view, the CEO said that it might not be advantageous from an economic point of view and in their activity due to the costs it entails. In addition, the population is still not very open to buying more expensive products just because they are environmentally friendly. In general, the CE will bring economic and social benefits in the long run and producing environmentally friendly articles can bring significant positive impacts to the society. Nevertheless, currently, they work more in specific sectors of activity, such as clothing and food, among others. They will consider possible measures that could benefit their sector, the environment and society.
4. When asked if there are many barriers to implementing a CE, the CEO related that the barriers, in a way, are in line with what he mentioned earlier, raw materials that are less harmful to the environment still have a high cost that is not valued by the population. They believe that incentives should also come from governing bodies at the national and European level, promoting incentives for companies that adopt policies that promote a CE. They are not aware of government incentives to produce environmentally friendly items and even apply for eco-certification. They have been betting on producing high-quality items that will have a longer lifespan, that is, a longer shelf life, extending the period for deposition. In addition, they have a restoration section, and since the product is of good quality, they encourage customers to reuse it, replacing only the covering.

3.2 Analyse of Financial Cost-Benefit

CBA is an analytical tool for judging the economic ad-vantages or disadvantages of an investment decision by assessing its costs and benefits in order to assess the welfare change attributable to it. Environmental impacts can result in significant costs for companies. However, it is not always easy to demonstrate the advantages of investing in prevention. The best way to do this will be through financial assessments and analysis (Ramos, 2013).

This study aims to demonstrate that Investment in Preventive Measures within the scope of Environmental Management is an essential investment for the Company with benefits that are transversal to the entire organization. The most significant costs were selected and realized a financial CBA for Energy Consumption.

Table 2 – Financial Cost-Benefit Ratio of the Acquisition of Photovoltaic Panels

Years	Annual energy consumption (kWh)	Energy cost (EUR)	Benefit (EUR)	Depreciation cost (EUR)	Net benefit (EUR)
1	9,767.00	2,350.83	1,207.15	1,777.50	-570.35
2	10,255.35	2,468.37	1,267.51	1,777.50	-509.99
3	10,768.12	2,591.79	1,330.88	1,777.50	-446.62
4	11,306.52	2,721.38	1,397.43	1,777.50	-380.07
5	11,871.85	2,857.45	1,467.30	0.00	1,467.30
6	12,465.44	3,000.32	1,540.66	0.00	1,540.66
7	13,088.71	3,150.34	1,617.70	0.00	1,617.70
8	13,743.15	3,307.85	1,698.58	0.00	1,698.58
9	14,430.31	3,473.25	1,783.51	0.00	1,783.51
10	15,151.82	3,646.91	1,872.69	0.00	1,872.69

The year 2018 was considered as a reference for presented data related to actual costs and consumption used as a basis for calculation. Table 2 shows the calculations of the annual net financial cost and main benefits obtained by implementing this measure, that was calculated according to the methodology indicated above.

The analysis was carried out on an annual basis and a forecast for 10 years (data presented in columns 2 to 10). The first line presents the annual energy consumption in kWh, with year 1 being the reference year 2018, whose consumption value was 9,767.00 EUR, considering an annual consumption/production growth of 5% for calculation of consumption for subsequent years. Line 2 shows the energy cost using the same logic as consumption, with year 1 being the reference year 2018 and the others, what was calculated based on the same annual growth. Finally, line 3 calculates the benefit obtained by reducing the target of 65% to the energy cost, deducted from the tax 21%. For example, in year 1, the 2,350.83 EUR energy cost was multiplied by 65%, which was also multiplied by (1-21%) to deduct taxes. In the following years the calculation followed the same reasoning.

The acquisition of photovoltaic panels including a 6 kW UPAC central with 6.555 Wp installed, 23 modules of 285 W, 46 fiscol structures, a three-phase suntrio, an inverter, 250 6 MM cables, a circuit breaker, and the installation will have an initial cost of 9,000 EUR. Line 4 shows the cost of equipment depreciation. According to the regulatory decree 25/2009 of 14 September 2009, alternative energy equipment is depreciated at a rate of 25%, having a useful fiscal life of 4 years. Therefore, 25% of the value of the investment, about 9,767.00 EUR, was calculated, and the taxes were deducted (ex: year 1; 9,767.00

EUR x 0.25 x (1-0.21). Therefore. this amount was depreciated from years 1 to 4 only. In line 5 the annual net benefit was calculated by reducing the cost of depreciation (line 4), to the benefit obtained (line 3). As a result, the sum of the net benefit obtained in year 1 to year 10 is 8,073.42 EUR, equal to the value of cash flows without applying the update factor.

The calculation of NPV is presented below, considering the data obtained in the previous table. The calculation of NPV is shown in Table 3 about acquisition of photovoltaic panels.

As stated before, the initial investment is 9,000.00 EUR presented in year zero, the year of acquisition. In line 2, the benefit before tax is calculated (e.g. year 1; 1,528.04 EUR obtained by multiplying the energy cost. 2,350.83 EUR by 65%, which is the desired reduction with the acquisition of the panels). This calculation formula was applied for all years. In line 3, the calculation of the additional taxes is presented in which the benefit (e.g. year 1; 1,528.04 EUR) is deducted from the depreciation amount (9,000.00 EUR x 0.25) multiplied by the tax rate (21%). From the 5th year onwards given that there is no longer any value to depreciate. The tax is calculated directly on the benefit.

Table 3 – NPV of the Acquisition of Photovoltaic Panels

Years	Initial Investment (EUR)	Benefit (EUR)	Additional IRC (EUR)	Cash-flows (EUR)	Actualization factor	Updated CF (EUR)	VAL	Time Years
0	9,000.00			-9,000.00	1.00	9,000.00	5,189.63	6.00
1		1,528.04	-151.61	1,679.65	0.97	1,622.85		-7,377.15
2		1,604.44	-135.57	1,740.01	0.93	1,624.32		-5,752.83
3		1,684.66	-118.72	1,803.38	0.90	1,626.55		-4,126.28
4		1,768.90	-101.03	1,869.93	0.87	1,629.53		-2,496.75
5		1,857.34	390.04	1,467.30	0.84	1,235.43		-1,261.32
6		1,950.21	409.54	1,540.66	0.81	1,253.33		-7.99
7		2,047.72	430.02	1,617.70	0.79	1,271.50		1,263.51
8		2,150.11	451.52	1,698.58	0.76	1,289.91		2,553.43
9		2,257.61	474.10	1,783.51	0.73	1,308.62		3,862.05
10		2,370.49	497.80	1,872.69	0.71	1,327.58		5,189.63

In line 4, cash flows are obtained by reducing the additional tax to the benefit. Thus, the sum of these cash flows will be equal to the sum of the annual net benefit calculated previously. To be able to calculate the NPV according to the methodology specified above, it is necessary to update the value of cash flows annually during the 10 years. In line 5, the update factor was calculated considering the respective year. The discount rate considered was 3.5%. For

example, for year 1. 1/ (1+3.5%) ^ 1 was calculated, where 3.5% is the discount rate and ^1 is the year to which it relates.

In line 6, the updated cash flows are presented considering the updating factor calculated in line 5. By the sum of all cash flows considering year zero, we obtained an NPV of 5,189.63 EUR. In line 8, it was calculated the period when the NPV assumes a value of zero. as a measure of moving to the investment recovery situation. According to this calculation, after 6 years more precisely in the passage from the 6th to the 7th year the NPV changes from negative to positive. The implementation of this measure is evident since the sum of the net benefit from years 1 to 10 is positive, and the NPV is also positive.

Concerning this measure, in addition to the financial benefit, the Company will produce cleaner and more renewable energy and at the same time will contribute to reducing the ecological footprint.

It is important to mention a recent study in this area using the same methodology. The authors Bošnjakovic, Cikic and Zlatunic (2021) highlight in their research "Cost-Benefit Analysis of Small-Scale Rooftop PV Systems: The Case of Dragotin. Croatia" the total investment equipment costs vary significantly depending on the type of installation, the capacity of the PV (photovoltaic) system, and the country in which the system is installed. These prices in 2020 range from 555 EUR/kWh in India up to 2,125 EUR/kWh in Switzerland. The investment costs of large PV systems are significantly lower. In Europe. the cost ranges from 600 EUR/kWh in the Czech Republic to 1,040 EUR/kWh in Ireland.

To determine the payback period, it is necessary to determine the profile of the average daily load for the analysed facility and the profile of the average daily production of electricity for the selected PV system. For the analysed facility, with the existing electricity prices in Croatia, which are among the lowest in the EU and with the existing legislation related to net metering, the roof PV system can be paid for in 10.5 years without government incentives. In addition to reducing electricity bills, installing PV roofing systems provides benefits, such as energy autonomy, reduced carbon emissions, and the creation of new local jobs.

According to the analysis above, it should be stated that financial (NPV methodology) and economic (CBA methodology) analyses have similar characteristics. Both estimate the net benefits of a project investment based on the difference between the situation with and without the project. However, financial analysis is largely confined to organizations or their units. It involves a detailed approach that compares the expenses and revenues of an enterprise to determine its profitability and thus its sustainability. For example, the financial analysis based on the NPV methodology measures cash flows and considers an opportunity cost of capital. In turn, economic analysis takes a broader view and assesses the analysis of the project's impact on society (e.g. Bošnjakovic, Cikic and Zlatunic, 2021). To do this. it should consider the viewpoints of all stakeholders and how the results of a project align with economic and social policies. Typically, costs in an economy are a measure of the resources that society collectively invests in

the realization of the project. The benefits, however, need not only be limited and often intangible benefits. Hence, the economic analysis based on the benefit cost analysis goes beyond cash flows and analyses externalities as well.

3.3 Sensitivity Analysis of the B/C Ratio

This study focused on the financial CBA, having analysed the possible investment from the perspective of its profitability, and calculated the NPV.

According to Guide to CBA of Investment Projects 2014-2020 (European Comission, 2014) the Sensitivity Analysis enables the identification of the 'critical' variables of the project. Such variables are those whose variations, be they positive or negative, have the largest impact on the project's financial and/or economic performance. The analysis is carried out by varying one variable at a time and determining the effect of that change on the NPV. As a guiding criterion, the recommendation is to consider 'critical' those variables for which a variation of ±1 % of the value adopted in the base case gives rise to a variation of more than 1 % in the value of the NPV.

The result of the B/C ratio and the NPV depends on the assumptions made and the variables identified. Considering that a change in a variable could cause the result to vary, a sensitivity analysis was carried out to study the acquisition of photovoltaic panels. The manager should not limit himself to the calculation of the B/C ratio but should also carry out a sensitivity analysis, considering the impact of the variation of the variables involved in the indicator. A sensitivity analysis will allow it to compare different scenarios that can better guide the manager in decision-making. For higher production growth rates, scenarios can become more favourable for decision making. Table 4 presents the sensitivity analysis according to the variation of some variables.

Note that the increase of 1% of the production tax rate leads to increase of NPV much higher than 1%, which assures that the production tax rate is a critical variable (European Commission, 2014).

In the beginning stages, to calculate the NPV of the investment in photovoltaic panels, the production growth rate variable was considered. Production was assumed to grow 5% annually over the period under review. If we increase the growth rate, we verify from the data that the NPV also increases, concluding that the greater the production growth, the greater the return on investment.

Decision making is simple because the return is always guaranteed, the manager knowing that the higher the level of sales and production, the greater the return on investments will be. As measures to be implemented. the acquisition of 23 modules of 285 W photo-voltaic panels is recommended. with a view to reducing the annual energy cost by 65%, as well as promoting the use of renewable and clean energy.

The measures must be integrated into a sustainability strategy based on technological innovation models aimed at valorisation, reduction, reuse, and recycling.

Table 4 – Sensitivity Analysis Decimal Points

	Energy	Photovoltaic panels
Critical variables		Production tax rate
Initial Value		5%
NPV		5,189.63 €
Variation 1		6%
NPV 1		5,751.79 €
Variation 2		7%
NPV 2		6,343.59 €
Variation 3		8%
NPV 3		6,966.57 €
Variation 4		9%
NPV 4		7,622.31 €
Variation 5		-
NPV5		-

3.4 The model - from Linear Economy to Circular Economy

Bennett, Pearce and Turner (1991) explain that the economic paradigm must change from a traditional economic system to one characterized by the CE. It is based on three economic functions: supply of resources, life support and source of absorption of waste and emissions.

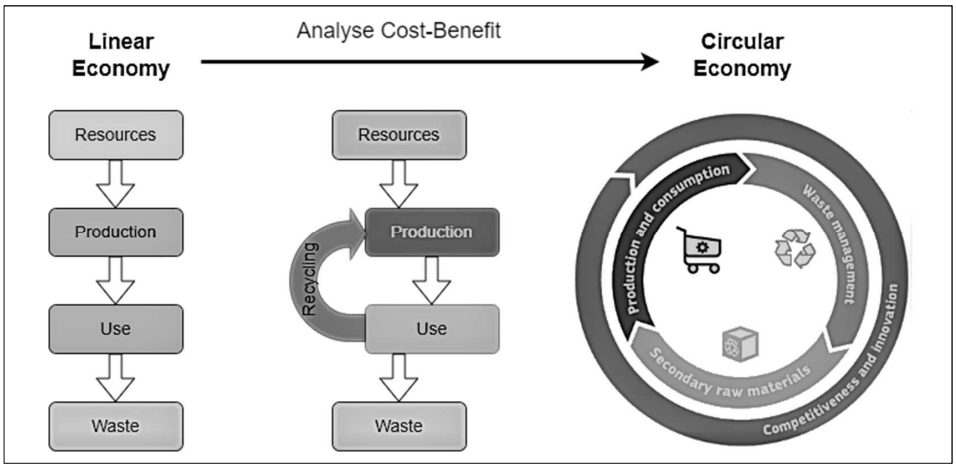


Figure 2 – Transition from Linear Economy to Circular Economy (Adapted from Bennett, Pearce and Turner, 1991; Portugal, 2019)

The CE is not only characterized by product recycling, but it is a new paradigm in which each material is reused as much as possible, reducing waste. It requires a new way of thinking, as well as how to live and consume sustainably. According to Figure 2, the model elaborated according to Eurostat (Portugal, 2019; European Commission, 2015; 2019a; 2019b; 2020). it is necessary to make a CBA and understand that the CE pays off from a financial point of view, and it is good for environment. Hence, it is necessary to adjust the productive system of companies. in the choice of raw materials, in the development of products, use of waste as by-products and recycling, among others.

4 CONCLUSIONS

The studied Company is still at an embryonic stage concerning CE. Nevertheless, it was possible to identify all environmental aspects and impacts and consumption and associated costs, perspectives, and strategies to be implemented. A survey of possible improvement measures to implement in this sector was carried. A relevant CBA was carried out in financial terms and in terms of costs and returns. It was verified the possibility of the Company reaching a more advanced state and approaching that of a CE. From the interview directed at the Company's CEO, from the documentary and empirical survey carried out, and from the application of the financial CBA method, it can be concluded that the implementation of a series of measures already identified will be advantageous.

It was also concluded that the Company should invest in the installation of photovoltaic panels to reduce the cost of energy produced and replace electricity with alternative energy. The Company intends to increase its production in the medium to long term. Hence, the electricity consumption also will increase, making it even more economically advantageous to be a producer of alternative energies. The conclusion supports that the CBA is a valuable tool for decision-making and for identifying the potential benefits that can arise from implementing measures from the CE perspective.

REFERENCES

Araújo, R., Santos, G., 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 journal*, 23(1), pp.1-24. DOI: 10.12776/QIP.V23I1.1132.

Bennett, J.W., Pearce, D.W. and Turner, R.K., 1991. Economics of Natural Resources and the Environment. *American Journal of Agricultural Economics*, 73(1), pp.227-231. DOI:10.2307/1242904.

Bošnjakovic, M., Cikic, A. and Zlatunic, B., 2021. Cost-Benefit Analysis of Small-Scale Rooftop PV Systems: The Case of Dragotin, Croatia. *Appl. Sci*. 11(19), pp.9318. DOI: 10.3390/app11199318.

Bravi, L., Santos, G., Pagano, A. and Murmura, F., 2020. Environmental management system according to ISO 14001:2015 as a driver to sustainable development. *Corporate Social Responsibility and Environmental Management*, 27(6), pp.2599-2614.

Camacho-Otero, J., Boks, C. and Pettersen, I.N., 2018. Consumption in the circular economy: a literature review. *Sustainability*, 10(8), pp.2758. DOI: 10.3390/su10082758.

Camón, L.E. and Celma, D., 2020. Circular Economy. A Review and Bibliometric Analysis. *Sustainability* 12(16), pp.6381. DOI: 10.3390/su12166381.

Carvalho, F., Santos, G. and Gonçalves, J., 2020. Critical analysis of information about integrated management systems and environmental policy on the Portuguese firms' website, towards sustainable development. *Corporate Social Responsibility and Environmental Management*, 27(2), pp.1069-1088.

Carvalho, H., 2015. Metodologia de análise de custo benefício para investimentos em infraestruturas energéticas. [pdf] U. Porto: FEUP. Available at: <https://paginas.fe.up.pt/~ee08333/wp-content/uploads/2013/10/HCV9.pdf> [Accessed 20 July 2021].

Costa, A.R., Barbosa, C., Santos, G. and Alves, M.R., 2019. Six Sigma: Main Metrics and R Based Software for Training Purposes and Practical Industrial Quality Control. *Quality Innovation Prosperity journal*, 23(2), pp.83-100. DOI: 10.12776/qip.v23i2.1278.

D'Amato, D., 2021. Sustainability Narratives as Transformative Solution Pathways: Zooming in on the Circular Economy. *Circ. Econ. Sust*., 1, pp.231-242. DOI: 10.1007/s43615-021-00008-1.

De Jesus, A., Antunes, P., Santos, R. and Mendonça, S., 2018. Eco-innovation in the transition to a circular economy: an analytical literature review. *J. Clean. Prod.*, 172, pp.2999-3018. DOI: 10.1016/j.jclepro.2017.11.111.

European Commission, 2014. *Guide to Cost Benefit Analysis of Investment Projects - Economic Appraisal Tool for Cohesion Policy 2014-2020*. Luxembourg: Publications Office of the European Union.

European Commission, 2015. *Public Consultation on the Circular Economy*. [online] EU Database. Available at: <https://ec.europa.eu/environment/consultations/closing_the_loop_en.htm> [Accessed 20 July 2021].

European Commission, 2019a. *Circular economy indicators.* [Indicators > Main tables > Circular economy indicators] Eurostat. Available at: <https://ec.europa.eu/eurostat/web/circular-economy/indicators/main-tables> [Accessed 20 July 2021].

European Commission, 2019b. *Circular Economy: Implementation of the Circular Economy Action Plan*. [online] European Union. Available at: <https://ec.europa.eu/environment/circular-economy/index_en.htm> [Accessed 20 July 2021].

European Commission, 2020. *Circular economy action plan.* [pdf] Luxembourg: Publications Office of the European Union. Available at: <https://ec.europa.eu/environment/circular-economy/pdf/new_circular_ economy_action_plan.pdf> [Accessed 20 July 2021].

Fonseca, L.M., Amaral, A. and Oliveira, J., 2021. Quality 4.0: The EFQM 2020 Model and Industry 4.0 Relationships and Implications. *Sustainability*, 13(6), pp.3107. DOI: 10.3390/su13063107.

Fonseca, L.M., Domingues, J.P., Pereira, M.T., Martins, F.F. and Zimon, D., 2018. Assessment of circular economy within Portuguese organizations. *Sustainability,* 10(7), pp.2521. DOI: 10.3390/su10072521.

Marques, A., Guedes, G. and Ferreira, F., 2017. Leather wastes in the Portuguese footwear industry: New framework according design principles and circular economy. *Procedia Engineering*, 200, pp.303-308. DOI: 10.1016/j.proeng.2017.07.043.

Masi, D., Day, S. and Godsell, J., 2017. Supply chain configurations in the circular economy: a systematic literature review. *Sustainability*, 9(9), pp.1602. DOI: 10.3390/su9091602.

Merli, R., Preziosi, M. and Acampora, A., 2018. How do scholars approach the circular economy? A systematic literature review. *J. Clean. Prod.* 178, pp.703-722. DOI: 10.1016/j.jclepro.2017.12.112.

Portugal, C.E., 2019. *Circular Economy Portugal.* [online] Retrieved from: <https://www.circulareconomy.pt> [Accessed 20 July 2021].

Ramos, D., 2013. *Cost - benefit analysis in occupational risk assessment*. Thesis. Tese de doutoramento - Programa Doutoral em Engenharia Industrial e de Sistemas. Universidade do Minho. Available at: <http://repositorium.sdum.uminho.pt/handle/1822/28766> [Accessed 20 July 2021].

Ramos, D.G., Afonso, P. and Costa, R., 2020. Cost-Benefit Analysis of Occupational Health and Safety: A Case Study. In: P.M. Arezes, J. Santos, Baptista, M.P. Barroso, P. Carneiro, P. Cordeiro, N. Costa, R.B. Melo, A.S. Miguel and G. Perestrelo, eds. 2020. *Occupational and Environmental Safety and Health II.* Springer Nature Switzerland. pp.689-695. Available at: <https://link.springer.com/chapter/10.1007%2F978-3-030-41486-3_73> [Accessed 20 July 2021].

Ramos, D.G., Arezes, P.M. and Afonso, P., 2017. Analysis of the return on preventive measures in musculoskeletal disorders through the benefit–cost ratio: A case study in a hospital. *International Journal of Industrial Ergonomics*, 60, pp.14-25.

Rebelo, M., Santos, G. and Silva, R., 2015. Integration of Standardized Management Systems: A Dilemma? *Systems*, 3, pp.45-59.

Rodrigues, J., Sá, J.C., Ferreira, L.P., Silva, F.J.G. and Santos, G., 2019. Lean management "quick-wins": Results of implementation. A case study. *Quality Innovation Prosperity Journal,* 23(3), pp.3-21. DOI: 10.12776/QIP.V23I3.1291.

Sá, J.C., Vaz, S., Carvalho, O., Morgado, L., Lima, V., Fonseca, L., Doiro, M. and Santos, G., 2020. A model of integration ISO 9001 with Lean six sigma and main benefits achieved. *Total Quality Management & Business Excellence*, 33(1-2), pp.218-242. DOI: 10.1080/14783363.2020.1829969.

Santos, D., Rebelo, M. and Santos, G., 2017. The Integration of certified Management Systems. Case Study Organizations located at the district of Braga. Portugal. *Procedia Manufacturing*, 13, pp.964-971.

Santos, G., Doiro, M., Mandado, E. and Silva, R., 2019a. Engineering learning objectives and computer assisted tools. *European Journal of Engineering Education*, 44(4), pp.616-628.

Santos, G., Gomes, S., Braga, V., Braga, A., Lima, V. and Teixeira, P., 2019b. Value creation through quality and innovation - a case study on Portugal. *TQM Journal*, 31(6), pp.928-947.

Santos, G., Sá, J.C., Félix, M.J., Barreto, L., Carvalho, F., Doiro, M., Zgodavová, K. and Stefanović, M., 2021. New Needed Quality Management Skills for Quality Managers 4.0. *Sustainability,* 13(11), pp.6149. DOI: 10.3390/su13116149.

Santos. G., Rebelo, M., Barros, S., Silva, R., Pereira, M., Ramos, G. and Nuno Lopes, N., 2014. Developments regarding the integration of the Occupational Safety and Health with Quality and Environment Management Systems (2014). In: I. Kavouras and M.-C.G. Chalbot, eds., 2014. *Occupational Safety and Health – Public Health in the 21st Century*. New York: The Nova Science Publishers. Chapter 6.

Sassanelli, C., Rosa, P., Rocca, R. and Terzi, S., 2019. Circular economy performance assessment methods: a systematic literature review. *J. Clean. Prod.,* 229, pp.440-453. DOI: 10.1016/j.jclepro.2019.05.019.

Silva, S., 2020. *Cost-Benefit analysis of implementing measures for a circular economy: a case study in an upholstery company.* [pdf] Instituto politécnico do Cávado e do Ave. Available at: <http://winlib.biblioteca.ipca.pt/Nyron/Library/Catalog/winlibsrch.aspx?skey=00C7E81F0BF64A15A9FCBB90C7FE5BBC&cap=&pesq=5&thes1=23966&doc=11659> [Accessed 20 July 2021].

Talapatra, S., Santos, G., Uddin, K. and Carvalho, F., 2019. Main benefits of Integrated Management Systems Through Literature Review. *International Journal of Quality Research*, 13(4), pp.1037-1054. DOI: 10.24874/IJQR13.04-19.

Teixeira, P., Araújo, M.J.F., Araújo, M.V.F., Araújo. A.H., Barros, J., Almeida, M., Fonseca, B., Reis, J.M., Barbosa, L., Santos, G. and Sampaio, N., 2021a. Pollution Credit Certificates Theory: An Analysis on the Quality of Solid Waste Management in Brazil. *Quality Innovation Prosperity Journal*, 25(3), pp.1-17. DOI: 10.12776/QIP.V25I3.1574.

Teixeira, P., Sá, J.C., Silva, F.J.G., Santos, G., Ferreira, L.P. and Fontoura, P., 2021b. Connecting lean and green with sustainability towards a conceptual model. *Journal Cleaner production*, 322, pp.129047. DOI: 10.1016/j.jclepro.2021.129047.

Zgodavová, K. and Slimák, I. 2008. Advanced Improvement of Quality. In: B. Katalinic, ed., *19th DAAAM International Symposium on Intelligent Manufacturing and Automation (DAAAM 2008)*. Trnava, Slovakia, 22-25 October 2008. Red Hook, NY: DAAAM International Vienna. pp.1551-1553.

Zgodavova, K., Bober, P., Majstorovic, V., Monkova, K., Santos, G. and Juhaszova, D., 2020. Innovative methods for small mixed batches production system improvement: The case of a bakery machine manufacturer. *Sustainability* 12(15), pp.6266. DOI: 10.3390/su12156266.

ABOUT AUTHORS

Delfina Ramos[0000-0003-2862-2883] (D.R.) – PhD., Researcher at Institute of Science and Innovation in Mechanical and Industrial Engineering (INEGI), Porto, Portugal and Algoritmi Centre, University of Minho, Guimaraes, Portugal; e-mail: dgr@isep.ipp.pt

Luis Fonseca[0000-0001-5668-0656] (L.F.) – PhD., Prof. at Engineering School, Polytechnic of Porto, and Researcher at Institute of Science and Innovation in Mechanical and Industrial Engineering (INEGI), Porto, Portugal, e-mail: lmf@isep.ipp.pt.

Joaquim Gonçalves[0000-0003-2219-1816] (J.G.) – PhD., Prof. at Technological School. Polytechnic Institute Cávado Ave, Barcelos, Portugal, e-mail: jgoncalves@ipca.pt.

Raquel Carvalho[0000-0003-3428-2221] (R.C.) – MSc., Prof. at School of Business Sciences, Polytechnic Institute of Viana do Castelo, Valença, Portugal, e-mail: raquelmcarvalhos@gmail.com

Sandro Carvalho[0000-0003-4470-4993](S.C.) – PhD., Prof. at Technological School, Polytechnic Institute Cávado Ave, Barcelos, Portugal, e-mail: scarvalho@ipca.pt

Gilberto Santos[0000-0001-9268-3272] (G.S.) – Dr.S., Prof. at Design School at the Polytechnic Institute of Cávado Ave (IPCA) and Researcher at ID+, Campus do IPCA, Barcelos, Portugal, e-mail: gsantos@ipca.pt.

AUTHOR CONTRIBUTIONS

Conceptualization, D.R. and L.F.; Methodology, R.C. and S.C.; Software, J.G.; Validation, J.G., D.R. and G.S.; Formal analysis, L.F.; Investigation, D.R.; Resources, D.R.; Data curation, J.G.; Original draft preparation, D.R. and L.F.; Review and editing, S.C. and R.C.; Visualization, L.F. and S.C.; Supervision, G.S.; Project administration, D.R.

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.

The Impact of Employees' Training on Their Performance Improvements

DOI: 10.12776/QIP.V26I1.1665

Martin Flegl, Lucie Depoo, Mayra Alcázar

Received: 2022-02-25 Accepted: 2022-03-19 Published: 2022-03-31

ABSTRACT

Purpose: Performance management is an increasingly important part of strategic human resources management. Employee performance is related to many factors, but education and training are two of the most significant. So far, literature has confirmed that a strategic focus on the employee training leads to the successful achievement of an organization's goals, development, and ultimately to performance. However, there is no consensus about the right level of the training.

Methodology/Approach: The objective of this article is to analyze the impact of training programs on employees' performance in case of 839 analysts in a Mexican public financial institution. The research was carried out over a two-year period and the data were evaluated using descriptive statistics and ANOVA.

Findings: The results reveal that low number of training hours, together with excessive training of more than 166 hours per year has limited or no impact on the performance. In fact, when employees had more than 166 hours of training, the training was negatively related to their performance.

Research Limitation/Implication: Management of organizations should carefully plan the amount of provided training hours per each employee. The impact of training varies based on seniority and number of hours spent on training, but there are no significant differences between employees' gender.

Originality/Value of paper: The originality of this article lies in the identification and evaluation of impact of employee training hours on performance in organization, as well as in the identification of breaking point of number of training hours which may lead to actual decrease of performance.

Category: Research paper

Keywords: employee assessment; performance; training adaption; seniority; workplace

1 INTRODUCTION

Current growing importance of information and innovation in today's competitive environment, especially now when dealing with impact by COVID-19 has started faster process of interrelated changes with impacts on each employee of an organization, as well as on an organization as a whole (Cerne, Jaklic and Skerlavaj, 2013). To develop skills, competences and potential of employees, each organization needs to improve and train their knowledge, skills and abilities to support development of productivity across different conditions and innovation potential of each individual (Na-Nan and Sanamthong, 2019). Every employee is a bearer of knowledge and has potential that an organization needs to use to its own development, but research shows that organizational development is usually limited by lack of educated and qualified human resources (Diesel and Scheepers 2019; Acebo and Viltard, 2018).

Employee knowledge, skills and abilities must be continuously developed through continuous training, increasing their qualifications, and thus increasing their productivity (Cerne, Jaklic and Skerlavaj, 2013; Avolio, Walumbwa and Weber, 2009). According to current studies, impact of training assessment significantly affects performance of the employees and plays a positive role in improving employees' performance (Mahmud, Wahid and Arif, 2019). Also, Guan and Frenkel (2019) found that work engagement mediates the relationship between training and in-role task performance. Furthermore, relationships between task performance and organizational citizenship behavior are moderated by type of human resource management in an organization.

Adequate focus on to human resources and their remuneration, benefits, motivation and stimulation are considered as a key area of management of every organization worldwide (Fajčíková, Urbancová and Kučírková, 2018; Turner and Kalman, 2015). Currently, we can see significant change in social responsibility of organizations towards to an active approach to employees and their remuneration management (Jain and Bhatt, 2015). It is important to realize that efficient remuneration management contains tangible and intangible rewards that leads to better employer branding and talent attraction (Jain and Bhatt, 2015; Vnoučková, Urbancová and Smolová, 2018). Festing and Schäfer (2014) proved that it is crucial to pay attention to employees' motivation through adequate remuneration and benefit system to reach positive psychological contract with employees and to increase their retention. Wilden, Gudergan and Lings (2010) state that today's high competition in the labor market, requires strategic investments in employee satisfaction. Talent management can be understood as a group of processes to attract, develop, motivate and retain employees to make them perform better (Swailes and Blackburn, 2016).

Farndale et al. (2014) pointed out that talent management is a strategic resource integration including proactive identification, development and strategic deployment of key employees with a high potential. As a result, talent management is crucial for organizations to remain competitive (Collings and

Mellahi, 2009; Goldstein and Ford, 2002) although it is sometimes unclear how the talent is translated to performance levels in organizations (McCracken, Currie and Harrison, 2016). As it is increasingly hard for companies to attract and retain talented workers, particularly those who are young and highly skilled, it is vital to offer benefits that capture employees' preferences no matter of the size of an organization, or its industry (Salgado, Flegl and Fejfarová, 2020). In this sense, talent management is a process through which it is possible to ensure that organizations have employees of required quantity and quality in accordance with current and future priorities of the organization (Collings and Mellahi, 2009).

Training activities have a positive impact on the performance of individual, teams and organizations (Aguinis and Kraige, 2009; Mpofu and Hlatywayo, 2015). For this reason, training programs have become an enormous business in terms of both the amount of effort expended and the money spent (Goldstein and Ford, 2002). The problem is that the benefit of the training is not evident on a daily basis and, thus, the expenses are hard to justify. Due to high costs of the training programs, organizations need to control and demonstrate their added value (Goldstein and Ford, 2002). According to Aguinis and Kraige (2009), U.S. organizations alone spend more than $126 billion annually on employee training and development. Most of the money in training is spent on developing technical skills (due to its importance to get specialized employees) and management-supervisory skills, which helps to develop leaders with the capacity to impact the business outcomes through producing extraordinary bottom-line results (Goldstein and Ford, 2002). Nevertheless, talent training must not only build new capabilities, but also continue to support and strengthen already gained capabilities (Joyce and Slocum, 2012). Therefore, our first hypothesis is developed as follows:

H_{01}: Training has a positive effect on employees' performance.

In this case, the more training an employee takes, the better the measurable improvement.

As Farjad (2012, pp.2838) pointed out, "training evaluation is a critical component of analyzing, designing, developing, and implementing an effective training programme". However, many organizations do not collect the information to determine the usefulness of their own training programs. Usually, companies evaluate the impact of training using employees' reactions at the end of each course (Alcázar and Flegl, 2019; Goldstein and Ford, 2002). As a result, organizations fail to capitalize on the opportunity that talent management can bring them (Joyce and Slocum, 2012). So, how can we know that we have well-designed training programs? Aguinis and Kraige (2009) suggest that the effects of training on performance, although subtle, can be verified to the extent that a program created new knowledge, skills or it led to a positive change in the work environment. Furthermore, organizations must ensure that training programs meet the expectations of both organizations and individual learners, facilitate

collection and interpretation of information, in order to make responsible decisions (Abdala, 2009; Goldstein and Ford, 2002). In this sense, the main challenge for any organization is to create an objective, reliable and hard data-based training evaluation. Qualitative evaluation might not be reliable since the impressions are not enough to know the effectiveness of the educational process.

To evaluate the training process, it is necessary to first check the degree of practical application that the acquired knowledge could have had in a certain period of time, reflected in the improvements in the skills and competences of the employees who took it. This leads us to think of employee performance evaluation as a fundamental element to evaluate the training, since it works as one of the main sources of information about the effects on the business or environment resulting from the improved performance of the trainee measured by the business or organizational key performance indicators (Gökhan, 2015). According to Berke (2001), performance evaluation is present at three main levels:

1. Organizational, which is linked to strategic planning, both in its component areas and their interrelation, and in the processes they will carry out;
2. Processes, which is focused on measuring the effectiveness and coherence of the processes established at the organizational level;
3. Individual, whose purpose is the establishment of goals and indicators at the level of employees, in order to plan and evaluate individual's performance and to establish strategic contribution that each employee has regarding his/her position.

Also, Musriha (2019) and Su (2020) state that strategy improving employees training, compensation, motivation and organizational commitment positively affect their performance.

The performance evaluation is a continuous process for all members of the organization, in which all are evaluated by their results (quantitative objectives) and by their attitudes (qualitative objectives or competencies), in order to discover how they could be more productive in the future (Aswathappa, 2005; Gan and Trigné, 2006). All this taking into account, the evaluation must be based on objective criteria to identify what has been done well to recognize it and what has been done wrong to correct it (Gan and Trigné, 2006). In this sense, it must be clear that it is necessary that both the employee and the supervisor understand the requirements of the position, as well as the expectations to be covered, since without this, it would be impossible to perform the evaluation (Daley, 1992). If the training is not evaluated, then we have no arguments to know whether the given training is adequate, or whether the contents of the training are precise and updated. The evaluation of the training is a critical component to analyze, design, develop and implement an effective program (Farjad, 2012), which can be differentiated by age, gender and hierarchical level allowing each employee to

receive specific training necessary to improve his/her performance, avoiding frustration and unnecessary stress of employees (Gursoy, Chi and Karadag, 2013). Based on above mentioned, the second hypothesis is the following:

H_{02}: Training has a positive effect on employees' performance regardless employees' gender (H_{02a}), age (H_{02b}) and seniority (H_{02c}).

Therefore, the objective of this article is to analyze the impact of training programs on employees' performance, considering employees' personal characteristics.

2 MATERIALS AND METHODS

2.1 Sample

The analysis is based on results from 839 employees from analysts' position in a Mexican public financial institution. More precisely, the sample only includes employees that worked in the institution during 2018 and 2019 and were evaluated regarding their performance in both years. The sample consists of 344 (41%) female and 495 (59%) male employees. The average employees' age was 36.25 years (Standard Deviation 8.86 years), in case of females the average age was 35.11 years (SD 8.85) and males 37.04 years (SD 8.86). The average employees' seniority was 11.47 years (SD 9.44), in case of females the average age was 9.97 years (SD 9.44) and males 12.50 years (SD 9.45). In order to be able to apply the statistical analysis, the sample was divided into ten categories regarding employees' age and seniority (Fig. 1 and Tab. 1).

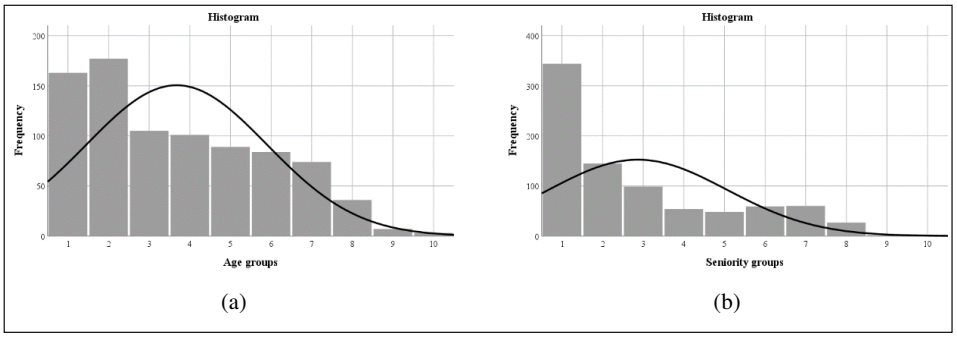


Figure 1 – Sample Distribution and Its Categories Regarding Employees' Age (a) and Seniority (b)

Table 1 – Frequency Distribution of Employees' Age and Seniority

Age			Seniority		
Category	**N**	**%**	**Category**	**N**	**%**
<23-27>	163	19.43%	<1.6-5.2>	304	36.23%
(27-31>	177	21.10%	(5.2-8.8>	152	18.12%
(31-35>	105	12.51%	(8.8-12.4>	87	10.37%
(35-39>	101	12.04%	(12.4-15.9>	69	8.22%
(39-43>	89	10.61%	(15.9-19.5>	40	4.77%
(43-47>	84	10.01%	(19.5-23.1>	38	4.53%
(47-51>	74	8.82%	(23.1-26.7>	58	6.91%
(51-55>	36	4.29%	(26.7-30.3>	45	5.36%
(55-59>	7	0.83%	(30.3-33.9>	42	5.01%
(59-63>	3	0.36%	(33.9-44.0>	4	0.48%

2.2 Employees' Performance Evaluation

To evaluate the effect of training on employees' performance, we used information related to number of hours each employee took of any kind of training during the year 2019. This information was taken from the internal information system of the institution. The training is divided into five basic instruments: Diploma and Specialty courses, Specific courses, Language courses, Study commissions and Internal courses. For the purpose of our analysis, the first instrument was split into two instruments. Similarly, the Internal courses were divided into three separate instruments: Leadership, Culture and Other courses. Thus, at the end, the analysis included information related to eight specific training instruments:

1. Diploma courses: Framed within formal education, these studies are organized by universities and allow employees to update their knowledge for the efficient performance of the tasks according to their job profile. Each employee receives a diploma after the satisfactory completion of his/her studies. The duration of these studies is less than one year.
2. Specialty courses: Framed within formal education, these studies are organized by universities and allow employees to update their knowledge for their efficient performance of job tasks according to their job profile. Each employee receives a diploma after the satisfactory completion of his/her studies. The duration of these studies is more than one year.
3. Specific courses: Framed within non-formal education, these are events such as courses, conferences, congresses, seminars, symposia and meetings, among others, organized by associations, institutes, universities,

specialized centers and other public or private entities. These courses aim at specific employees' needs. After the completion, employee does not achieve any academic degree.

4. Language courses: These courses are offered by specialized institutes and educational centers (accredited by the institution). The objective is to satisfy employees' needs to improve level of foreign languages according to their job profile.
5. Study commissions: Courses, conferences, congresses, seminars, symposiums and meetings, among others, organized by associations, institutes, universities, specialized centers or other entities outside the position of assignment of workers, both in national territory and abroad, aimed at covering specific needs that do not confer any academic degree.
6. Leadership courses: Courses belonging to the internal offer of the Institution. These courses can be, at the consideration of each area, granted to those workers with an analyst level who, due to their performance, are considered with high potential in order to prepare them for a possible promotion.
7. Culture courses: Open access courses that belong to the internal offer of the institution of a transversal nature aimed at reinforcing the identity of the institution in topics such as values, ethics, risk prevention, or diversity and inclusion.
8. Other courses: Open access courses that belong to the internal offer of the institution of a transversal nature aimed at the development of soft skills established in an institutional way for all hierarchical levels.

During the year 2019, employees took between one and 777 hours of training, where the average was 66.56 hours (SD 93.68 hours). The most common level of training was of 3 hours (23.0% of employees), followed by 7 hours (6.1%) and 4 hours (2.4%). For the purpose of the analysis, the amount of training was also divided into ten categories (Fig. 2 and Tab. 2).

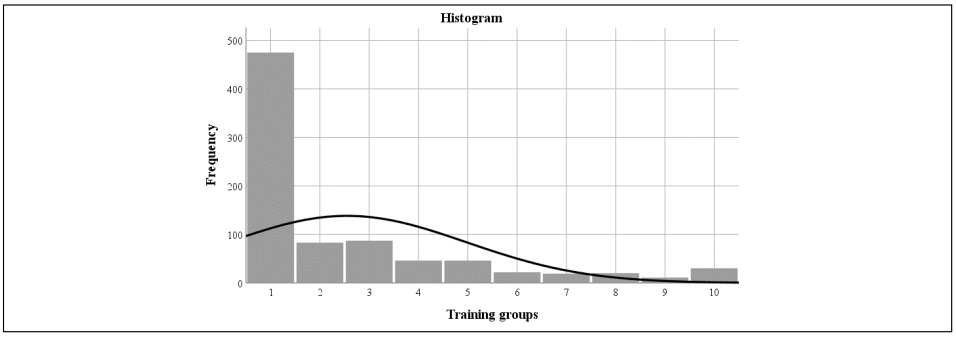


Figure 2 – Sample Distribution and Its Categories Regarding the Level of Employees' Training

Table 2 – Frequency Distribution of the Level of Employees' Training

Level of training		
Category	**N**	**%**
<1-33.2>	475	56.62%
(33.2-66.5>	83	9.89%
(66.5-99.7>	87	10.37%
(99.7-132.9>	46	5.48%
(132.9-166.1>	46	5.48%
(166.1-199.4>	22	2.62%
(199.4-232.6>	19	2.26%
(232.6-265.8>	20	2.38%
(265.8-299.1>	11	1.31%
(299.1-777>	30	3.58%

The employees' performance was measured on a 5-point scale according to the internal annual evaluation process in which managers are asked to assign a score to each person on their team, based on their performance and achievement of their goals. The final grade is obtained by weighing hard skills by 90% and soft skills by 10%. The average evaluation in 2018 was 3.54 pts (SD 0.390) and in 2019 the average evaluation was 3.574 pts (SD 0.402).

2.3 Analysis of Variances (ANOVA)

Analysis of Variance (ANOVA) is based on the comparison of the variance due to the between-groups variability with that of the within-group variability. Under the null hypothesis (stating that there are no significant differences between groups in the population), the variance estimated from the within-group variability should be about the same as the variance estimated from the between-groups variability (Hinton, 2014). For our purpose, we used Multivariate ANOVA (MANOVA), concretely two-way ANOVA. This method compares the mean differences between groups that have been split into two independent factors (Hinton, 2014). The two-way ANOVA in this paper test if there is an interaction between the four independent factors (employees' gender, age and seniority and amount of training hours) on the dependent variables (performance evaluation in 2018 and 2019, and difference between the evaluation from 2018 to 2019). All the calculations were made using IBM SPSS Statistics 25.

3 RESULTS

3.1 Performance Evaluation in 2018

First, we analyze the characteristics of the performance evaluation in 2018 to obtain initial perspective about respondents' evaluation. The average evaluation of all analysts was 3.536 pts. Regarding the employees' gender, there is no statistically significant difference in the evaluation (p = 0.27), although females had an average evaluation of 3.489 pts compared to males 3.434 pts, i.e. the difference is +0.055. What is interesting, that females are much better evaluated in the youngest age group (between 23 and 27 years), where females have an average evaluation of 4.073 pts (+0.537 pts above the company average and 0.474 pts above males, p = 0.001). The youngest group of employees includes 19.4% of all employees in the company. Then, the older an employee is, the lower the evaluation. However, the evaluation is decreasing in a faster way in case of females (in average by -15.584% every age group compare to -4.035% in case of males).

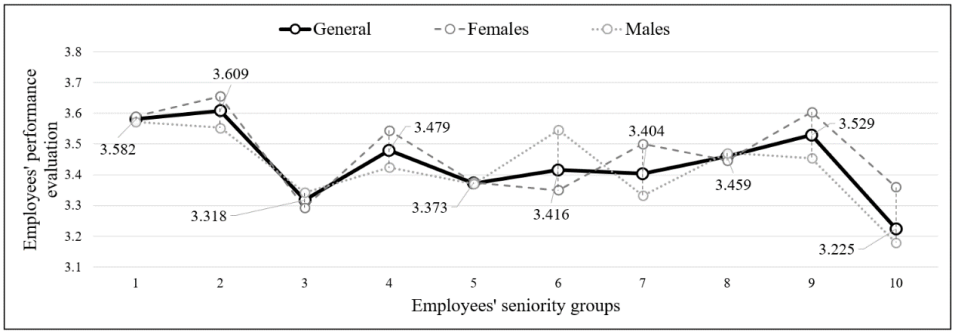


Figure 3 – The Effect of Employees' Seniority on Their Evaluation

Between 27 and 55 years of age, the evaluation of both genders is more less the same without any statistically significant difference. Males are much higher evaluated in the age group of 55 to 59 years (+0.362 pts), however, due to the few observations (only 7 employees representing 0.8% of the company), this difference is not statistically significant (p = 0.312). This result corresponds to the overall effect of the employees' age. Younger employees have higher evaluation. The average evaluation of the youngest generation is 3.837 pts, which is significantly higher than the rest of the groups.

When we consider the employees' seniority (Fig. 3), the highest evaluation can be observed in case of employees between 1.5 and 8.8 years of seniority (3.582 pts, respectively 3.609 pts) and between 30.3 and 33.9 years of seniority (3.529 pts). The effect of gender in the seniority does not have statistically significant effect on the evaluation.

3.2 Performance Evaluation in 2019

Second, we analyze the characteristics of the performance evaluation in 2019 in order to discover possible differences in both years. Employees' gender has no statistically significant effect on the evaluation again (p = 0.119), although the difference between females and males increased to +0.062 pts (+0.007 pts) compared to the evaluation in 2018. The average evaluation in case of females is 3.609 pts, whereas the average evaluation of males is 3.546 pts. The growth in the difference corresponds to the overall evaluation growth. The pattern of better evaluation in case of younger female employees is not confirmed in this case. Contrary, the youngest male employees are better evaluated than females of the same age (3.745 pts compared to 3.69 pts), but this difference is not statistically significant (p = 0.559). What is confirmed is the negative effect of the employees' age on the evaluation. The older the employee is, the lower the evaluation is. Males lose in average -6.659% every age group and females lose -4.482% in the same way. Regardless employees' gender, younger employees have statistically higher evaluation. The average evaluation of the youngest employees is in average 3.716, which is higher by +0.519 pts (+13.97%).

Considering the employees' seniority, the effect is surprisingly negative. The more seniority each employee has, the lower the evaluation. The highest evaluation can be observed in case of employees with seniority between 12.4 and 15.9 years (3.647 pts). However, excluding this seniority group, then the effect is "linearly" negative. The employees with the least seniority have the average evaluation of 3.623 pts, where the following seniority groups lose in average 2.42% (-0.024 pts). This effect has no statistically significant differences when the employees' gender is considered. In both cases the effect is negative, male employees lose in average 3.20% (-0.116 pts) with the growth of seniority group, whereas the effect is slightly lower in case of female employees (-1.75%; -0.063 pts). The effect would be similar, but females report growth in their evaluation during seniority of 5.2 to 8.8 years and 12.4 to 15.9 years. However, these two blips do not eliminate the negative effect of seniority.

3.3 Effect of Employees' Training on the Evaluation in 2019

As the analyzed sample includes the information about the employees' training during 2019, we are able to analyze the effect of training on the evaluation. The results indicate that the level of training has a positive effect on the evaluation. The highest effect is reported for the level of training between 265 and 299 hours (3.712 pts) and 166 and 199 hours (3.708 pts), representing 1.3% and 2.6% of employees. The majority of employees (56.6%) took between 1-33 hours of training during the year, resulting in average evaluation of 3.483 pts (-0.091 pts below the overall average), followed by the group with 66 and 99 hours (10.4%, 3.577 pts). In majority of the cases, employees' gender and level of training has no effect on the evaluation. The only exception is in level of 166 to 199 hours of training, where females have statistically higher evaluation by +0.368 pts

(p = 0.086). Male employees have higher evaluation for excessive level of training between 199 and 265 hours of training (without statistical significance).

Considering the employees' seniority, we can observe several differences. The most hours of training courses were taken by the employees with few years in the company (in average 98.050 hours of training). Then, the amount of training hours has negative trend (Fig. 4). On average, each seniority level takes -10.06% less training. We can observe significant differences between females and males and their seniority. In general, males take more hours of training: especially between seniority of 8.7 and 12.3 years (+19.91 hours more of training), 15.9 and 19.5 years (+29.11 hours more) and 19.5 to 23.1 years (35.63 hours more). However, as observed before, more training hours did not lead to statistically higher evaluation of male employees.

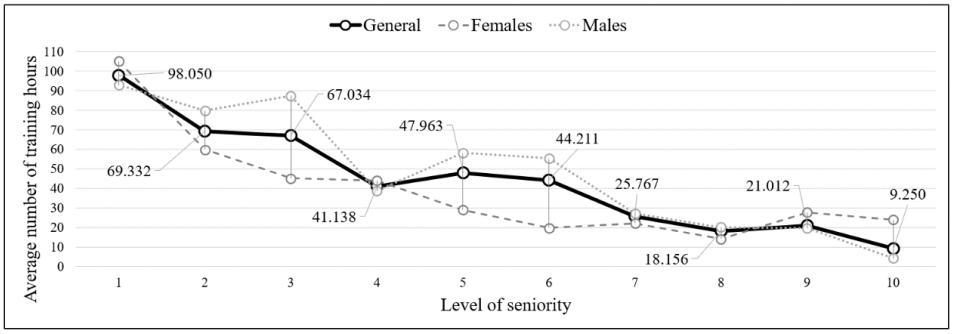


Figure 4 – The Effect of Employees' Seniority on Level of Training

3.4 Effect of Employees' Training on the Evaluation Improvements

In the previous section, we explored the characteristics of employees' evaluation and the effect of the training on the evaluation in 2019. As the results reveal, the effect of the training is positive. The more training, the better the evaluation. However, so far, we have not considered whether this effect leads to improvements over time. Therefore, in this section, we consider as a dependent variable the difference between the evaluations from 2018 to 2019. The working hypotheses states that training improves the evaluation and that the more training an employee takes, the better the improvement should be (H_{01}).

The average change in the evaluation is +0.0379 pts, and 95% of employees have the difference in the evaluation between +0.0116 and +0.0643 pts. In general, training has positive effect on the employees' improvements. For example, employees who took between 1 and 33 hours of training improve their evaluation by +0.047 pts. This is important as 56.6% of the employees belong to this group. As Fig. 5 indicates, the effect increases by +0.028 (p = 0.648) in the group of 33 and 66 hours of training (9.9% of employees), but then the effect is lighter and lighter. For employees between 133 and 166 hours of the training (5.5% of employees) the effect is almost zero (+0.01 pts) and the biggest negative effect is reported by employees with level of training between 166 and 199 hours (-0.188

pts). This negative effect is statistically significant compare to most of the other groups. The peak of the improvement by +0.252 pts in case of the group between 265 and 299 hours is only statistically significant to group 6 and 10 due to the low number of employees in this group (11 employees representing 1.3% of the sample). As a summary, training had positive impact on 88.3% of employees, but no strong correlation is observed (r = 0.0263) between the hours of training and evaluation improvements.

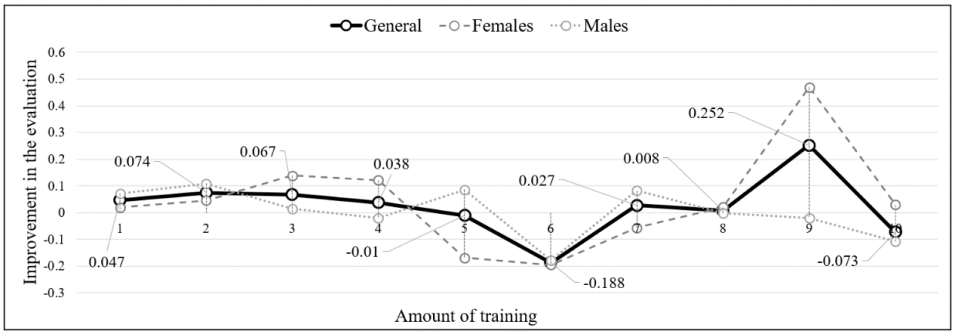


Figure 5 – The Effect of Training on Employees' Evaluation Improvement

Considering the employees' gender (H_{02a}), there is no statistically significant difference in the improvements (p = 0.838), as females have the difference higher by only +0.008 pts (0.043 pts versus 0.035 pts). However, if we consider the amount of training hours, differences can be observed. Female employees have peak in the improvements between 66 to 133 hours of training (+0.139 pts and +0.121 pts), which is followed by a significant drop resulting in negative change in the evaluation. Male employees report the peak evaluation between 33 and 66 hours of training (+0.109 pts) and 133 and 166 hours (+0.086). For both genders, there is the absolute bottom in for 166 and 199 hours of training (Fig. 5). Although females and males have different effect of the training, the only statistically significant difference occurs in the level of 133 to 166 hours of training (p = 0.067), where males have higher effect by +0.254 pts, and in the level of 266 to 299 hours (p = 0.048; +0.489 pts for females).

Regardless the gender, the effect of age (H_{02b}) on the improvements is rather positive, which is in contradiction to the basic characteristics of the 2018 and 2019 evaluation. In this case, the older the employee is, the bigger the effect of training is (although the effect is not statistically significant). The youngest group of employees took the highest amount of training (Fig. 6), but the improvement is negative (-0.029 pts). After, the level of improvements grows until the age group of 51-55 years (+0.124 pts), but no statistically significant differences can be found. Further, if we consider employees' gender, significant differences occur. Negative effect (-0.148 pts) is observed in case of youngest females (age 23 to 27 years), resulting in males' higher improvement by +0.248 pts (p = 0.005). Males have negative improvements in age group of 35 to 39 years (-0.067 pts), but the difference compared to females -0.146 pts is not statistically significant, as the

rest of the groups are not statistically significant. Moreover, there is no correlation at all (r = -0.00007) between the employees' age and the evaluation improvements.

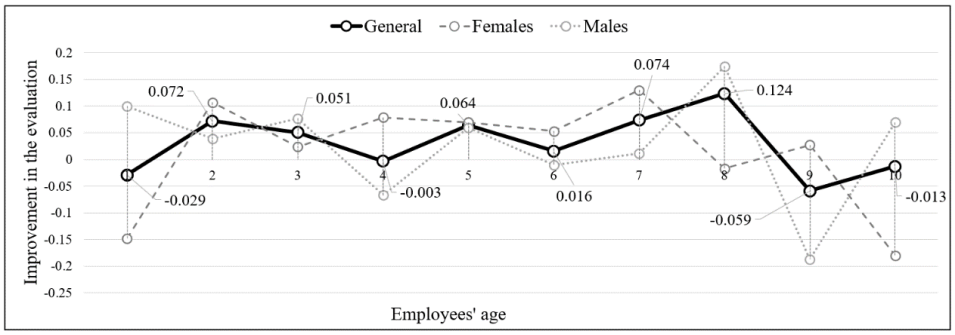


Figure 6 – Relation between Employees' Age and Evaluation Improvements

If we consider the employees' seniority (H_{02c}), we can observe significant drop in the case of 5.2 to 8.8 years of seniority (-0.082 pts), followed by a significant growth with more seniority (Fig. 7). The highest effect of the training is during the seniority of 8.8 to 12.4 years (+0.154 pts), which is statistically significant to most of the other groups of seniority. The effect of training losses its impact between 26.7 and 30.3 years of seniority (-0.001 pts). Different pattern can be observed if we consider employees' gender. Male employees follow the general patter, except the huge drop in seniority of 19.5 to 23.1 years (-0.064 pts), lower by -0.276 pts (p = 0.075) compared to females of the same seniority (in their case is the peak of the improvements). Females report lower growth of the improvements between 8.8 and 12.4 years of seniority (-0.160 pts compared to males, p = 0.125) and decrease in 15.9 to 19.5 years of seniority (-0.141 pts; p = 0.364). No correlation (r = 0.0056) between employees' seniority and evaluation improvements is observed.

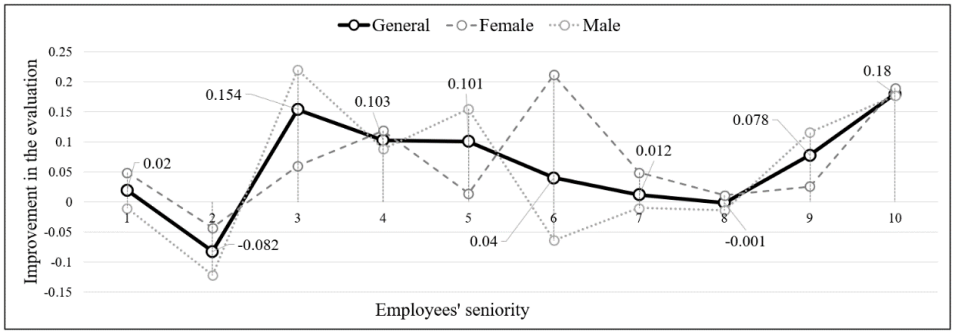


Figure 7 – Relation between Employees' Seniority and Evaluation Improvements

4 DISCUSSION

Nowadays, it is crucial for organizations to search areas for possible improvements in order to enhance their performance and remain competitive (Collings and Mellahi, 2009). Employees' training ensures an employee new skills and abilities, which is a valid and useful way to increase their performance and flexibility leading to a competitive advantage, which is beneficial for their organization (Pinzone et al., 2019). Employee's performance means that the employee has the capabilities and efficiency for their present job. What is more, competitive performance likely enables employees to feel satisfied or successful in their careers (De Vos, Van der Heijden and Akkerman, 2018), which is vital for long-term organization's goals. That is why, the employees' training and development has become one of the key aspects in improving organizational performance and growth (Mpofu and Hlatywayo, 2015) as investigated, correct training programs can have a positive impact not only on the individual performance, but also on teams and the whole organization (Aguinis and Kraige, 2009). Our results confirmed positive relationship between employee training and performance.

However, to secure the functionality of the training it is necessary to differentiate training programs regarding employees' characteristics, such as gender, seniority or their hierarchical level, as confirmed by our study. Ignoring such characteristics can lead to employees' frustration, greater tensions among employees and malfunctioned training programs (Gursoy, Chi and Karadag, 2013). The results detect that organizations must carefully evaluate the amount of provided training. Excessive training of more than 166 hours per year is the breaking point between positive and negative impact on the performance (Fig. 6). Thus, employee's requests for additional training hours should be critically evaluated as the added value can be negative.

Therefore, it must be clear what competences are required in each department of an organization (as well as in the organization as a whole) in order to capture current and future priorities of the organization (Collings and Mellahi, 2009) to maximize their impact. Similarly, organizations must understand their employees' characteristics to achieve maximum from the training. Our study shows the differences in gender, age and seniority as determinants of variability of employees and the necessity to treat them according to their characteristics to achieve desired output of education and training. Other researches, for example by Álcazar and Flegl (2019) observed that female training should be linked more to soft skills training (teamwork, communication, problem-solving and flexibility), whereas males' training should be linked more to hard skills (such as leadership, empathy and also communication, but in the way of typing, writing and software). Thus, managers should identify exact employees' needs and combine them with their personal characteristics to maximize the benefits from training. Moreover, employees must be ready and motivated for training. Organization should demonstrate the value of training before training begins, make sure employees are highly involved and engaged with their job (Aguinis

and Kraige, 2009; Hanaysha, 2016), which was confirmed by our research. Employees with high motivation express their performance through the effort they use to develop their activities. Nye et al. (2017) demonstrated that employees' interest is positively related to their training performance and on-the-job task performance. As the results of this paper show, organizations should not assign any number of training hours to all employees as it can have negative effect on their performance improvements. Instead, organizations should precisely identify employees' interest in specific areas, and target the training in this direction. Understanding employees' interests, particularly the degree of fit between individuals' interests and work environments, is useful for enhancing our understanding about how people do work, who excels, and who falls short at work (Su, 2020).

The key question is how to demonstrate the added value of such training. One of the ways how to demonstrate this is to analyze improvements in employees' skills/competences. Organizations are usually more interested in training transferable skills (communication, problem-solving and teamwork) and employee's personality over job-oriented skills and knowledge (Huq and Gilbert, 2013; McCracken, Currie and Harrison, 2016). This orientation creates mainly problems in soft-skills training evaluation as these skills are hard to measure. Thus, one of the possibilities is to use historical data and link the training effect to improvements in employees' performance evaluations. As the results demonstrated, the effect of training is different when one-year evaluation is considered, and when data across periods are considered. Historical comparison improves the interpretative capability. In a similar way, to be able to make the right conclusion about the added value of the training, the evaluation should consider employees with similar characteristics, such as employees from the same position, seniority, etc. In this case, these employees take very similar (if not the same) training courses and thus the improvements in performance should be very similar.

Moreover, it is important to remember that employee training is an investment into the organizational workforce and its results may not be visible immediately. In some cases, trained competence can have visible impact after a combination of several training programs, or after a period of expertise adjustments (experience). Furthermore, there is a risk of subjectivity when evaluating competences in a performance evaluation, as consequence of how managers respond in terms of their cognition, affectivity, and behavior (Fischer, 2010). Subjectivity is naturally included in any decision-making process and its effect can be both positive and negative. The accuracy of the performance evaluation then depends on decision-maker's experience. Still, although the added value of training includes many direct and indirect variables, all studies confirmed that training has positive effect on employees' performance, if carefully handled.

5 CONCLUSION

The results shown that focus on employee training brings improvements in performance over years and it forms prospective area for organizational development (H_{01} was confirmed). It was found that inadequate number of training hours, together with excessive training of more than 166 hours per year has limited or no impact on performance. Employees who completed more than 166 hours of training had worsened their performance. Thus, number of training hours should be critically evaluated as high number of training limits actual job performance of defined tasks and its value is disputable. On the other hand, the most efficient training was proven in new cohorts of employees (shortly after they joined studied organization). Thus, it is highly recommended to invest in training of new employees as training hours of freshmen and newcomers has the highest impact on performance (H_{02c} was conformed). Organizational management or human resource department should carefully plan the amount and type of training provided to each employee. Overall, the impact of training varies based on seniority and number of hours spent on training. On the other hand, there are no significant differences among gender of employees (H_{02a} was rejected). Some slight differences in training perceptions were found among age, but those were not statistically significant (H_{02b} was not rejected).

A practical recommendation for organizations is the application of employee training with an emphasis on organizational development and ensuring knowledge continuity. Organizations can use training to increase the level of their internal employee performance. It has also impact on employer branding, retention of employees, continuous education of key employees, and maintaining knowledge within an organization. Furthermore, employee training improves reactions on crises, staff shortages, knowledge continuity problems, and stabilization of staff and harmonization of internal environment of an organization, organizational culture and climate.

The follow-up research may focus on investigation of types of training and their effect on employees based on their gender, age, and seniority. In terms of limitation of this study, the first one is focus on only one organization and type of employees. The second limitation is the process of interpretation of results by researchers, when mediating effect of other factors leading to performance were not considered. But due to the fact that outputs are based on real employee performance and training hours loaded in case organization it shall be reliable.

REFERENCES

Abdala, E., 2009. La evaluación de los programas de capacitación laboral para jóvenes en Sudamérica. *Papeles de población*, 15(59), pp.11-82.

Acebo, M.N. and Viltard, L.A., 2018. Corporate Culture: A Key to Stimulate Innovation. *Independent Journal of Management & Production*, 9(3), pp.869-888. DOI: 10.14807/ijmp.v9i3.735.

Aguinis, H. and Kraiger, K., 2009. Benefits of Training and Development for Individuals and Teams, Organizations, and Society. *Annual Review of Psychology*, 60, pp.451-474.

Alcázar, M. and Flegl, M., 2019. Is popularity of training programs reflected in performance improvements? A case of Mexican public financial institution. *Revista Latinoamericana de Investigación Social*, 2(2), pp.24-41.

Aswathappa, K., 2005. *Human Resource and personnel management*. Nueva Delhi: Tata McGrawHill Education.

Avolio, B.J., Walumbwa, F.O. and Weber, T.J., 2009. Leadership: Current Theories, Research, and Future Directions. *Annual Review of Psychology Book Series: Annual Review of Psychology*, 60, pp.421-449. DOI: 10.1146/annurev.psych.60.110707.163621.

Berke, G., 2001. *How to conduct a performance appraisal*. Alexandría: ASTD.

Cerne, M., Jaklic, M. and Skerlavaj, M., 2013. Authentic leadership, creativity, and innovation: A multilevel perspective. *Leadership*, 9(1), pp.63-85. DOI: 10.1177/1742715012455130.

Collings, D.G. and Mellahi, K., 2009. Strategic talent management: A review and research agenda. *Human Resource Management Review*, 19(4), pp.304-313. DOI: 10.1016/j.hrmr.2009.04.001.

Daley, D., 1992. *Performance appraisal in the public sector: techniques applications*. Connecticut: Quorum Books.

De Vos, A., Van der Heijden, B. and Akkermans, J., 2020. Sustainable careers: Towards a conceptual model. *Journal of Vocational Behavior*, 117, pp.103196. DOI: 10.1016/j.jvb.2018.06.011.

Diesel, R. and Scheepers, C.B., 2019. Innovation climate mediating complexity leadership and ambidexterity. *Personnel Review*, 48(7), pp.1782-1808. DOI: 10.1108/PR-11-2018.

Fajčíková, A., Urbancová, H. and Kučírková, L., 2018. Decisive factors of talent management implementation in Czech Organisations. *Journal on Efficiency and Responsibility in Education and Science*, 11(1), pp.9-15. DOI: 10.7160/eriesj.2018.110102.

Farjad, S., 2012. The Evaluation Effectiveness of training courses in University by Kirkpatrick Model (case study: Islamshahr university). *Procedia - Social and Behavioral Sciences*, 46, pp. 2837-2841. DOI: 10.1016/j.sbspro.2012.05.573.

Farndale, E., Pai, A., Sparrow, P. and Scullion, H., 2014. Balancing individual and organizational goals in global talent management: A mutual-benefits perspective. *Journal of World Business*, 49(2), pp.204-214. DOI: 10.1016/j.jwb.2013.11.004.

Festing, M. and Schäfer, L., 2014. Generational challenges to talent management: A framework for talent retention based on the psychological-contract perspective. *Journal of World Business*, 49(2), pp.262-271. DOI: 10.1016/j.jwb.2013.11.010.

Fischer, F.M., 2010. *The Application of the Controllability Principle and Managers' Responses: A Role Theory Perspective*. Heidelberg: Gabler Verlag.

Gan, F. and Trigné, J., 2006. *Manual de instrumentos de gestión y desarrollo de las personas en las organizaciones*. Madrid: Díaz de Santos.

Gökhan, Ö., 2015. Program Evaluation through Kirkpatrick's Framework. *Pacific Business Review International*, 8(1), pp.106-111.

Goldstein, I.L. and Ford, J.K., 2002. *Training in Organizations*, Belmont, CA: Wadsworth.

Guan, X.Y. and Frenkel, S., 2019. How perceptions of training impact employee performance: Evidence from two Chinese manufacturing firms. *Personnel review*, 48(1), pp.163-183. DOI: 10.1108/PR-05-2017-0141.

Gursoy, D., Chi, C.G. and Karadag, E., 2013. Generational differences in work values and attitudes among frontline and service contact employees. *International Journal of Hospitality Management*, 32, pp.40-48. DOI: 10.1016/j.ijhm.2012.04.002.

Hanaysha, J., 2016. Examining the Effects of Employee Empowerment, Teamwork, and Employee Training on Organizational Commitment. *Procedia - Social and Behavioral Sciences*, 229, pp.298-306. DOI: 10.1016/j.sbspro.2016.07.140.

Hinton, P.R., 2014. *Statistics Explained*. Routledge.

Huq, A. and Gilbert, D.H., 2013. Enhancing graduate employability through work-based learning in social entrepreneurship: A case study. *Education + Training*, 55(6), pp.550-572. DOI: 10.1108/ET-04-2012-0047.

Jain, N. and Bhatt, P., 2015. Employment preferences of job applicants: Unfolding employer branding determinants. *Journal of Management Development*, 34(6), pp.634-652. DOI: 10.1108/JMD-09-2013-0106.

Joyce, W.F. and Slocum, J.W., 2012. Top management talent, strategic capabilities, and firm performance. *Organizational Dynamics*, 41(3), pp.183-193. DOI: 10.1016/j.orgdyn.2012.03.001.

Mahmud, K.T., Wahid, I.S. and Arif, I., 2019. Impact of training needs assessment on the performance of employees: Evidence from Bangladesh. *Cogent Social Sciences*, 5(1), pp.1705627. DOI: 10.1080/23311886.2019.1705627.

McCracken, M., Currie, D. and Harrison, J., 2016. Understanding graduate recruitment, development and retention for the enhancement of talent management: sharpening 'the edge' of graduate talent. *The International Journal of Human Resource Management*, 27(22), pp.2727-2752. DOI: 10.1080/09585192.2015.1102159.

Mpofu, M. and Hlatywayo, C., 2015. Training and development as a tool for improving basic service delivery; the case of a selected municipality. *Journal of Economics, Finance and Administrative Science*, 20(39), pp.133-136. DOI: 10.1016/j.jefas.2015.10.004.

Musriha, M., 2019. The implication of strategy improving employees training, compensation, motivation and organisational commitment as predictors of work performance in private commercial banks Indonesia. *International Journal of Business Performance Management*, 20(1), pp.1-15.

Na-Nan, K. and Sanamthong, E., 2019. Self-efficacy and employee job performance: Mediating effects of perceived workplace support, motivation to transfer and transfer of training. *International Journal of Quality & Reliability Management*, 37(1), pp.1-17. DOI: 10.1108/IJQRM-01-2019-0013.

Nye, C.D., Su, R., Rounds, J. and Drasgow, F., 2017. Interest congruence and performance: Revisiting recent meta-analytic findings. *Journal of Vocational Behavior*, 98, pp.138-151. DOI: 10.1016/j.jvb.2016.11.002.

Pinzone, M., Guerci, M., Lettieri, E. and Huisingh, D., 2019. Effects of 'green' training on pro-environmental behaviors and job satisfaction: Evidence from the Italian healthcare sector. *Journal of Cleaner Production*, 226, pp.221-232. DOI: 10.1016/j.jclepro.2019.04.048.

Salgado, K., Flegl, M. and Fejfarová, M., 2020. Factors affecting talent retention in tech start-ups. *Scientific Papers of the University of Pardubice – Series D*, 28(1), pp.138-149.

Su, R., 2020. The three faces of interests: An integrative review of interest research in vocational, organizational, and educational psychology. *Journal of Vocational Behavior*, 116(Part B), pp.103240. DOI: 10.1016/j.jvb.2018.10.016.

Swailes, S. and Blackburn, M., 2016. Employee reactions to talent pool membership. *Employee Relations*, 38(1), pp.112-128. DOI: 10.1108/ER-02-2015-0030.

Turner, P. and Kalman, D., 2015. Make your people before making your products: How to maximize the potential of an inclusive workforce. *Human Resource Management International Digest*, 23(1), pp. 28-31.

Vnoučková, L., Urbancová, H. and Smolová, H., 2018. Building Employer Image Thanks to Talent Programs in Czech Organizations. *Engineering Economics*, 29(3), pp.319-331. DOI: 10.5755/j01.ee.29.3.13975.

Wilden, R., Gudergan, S. and Lings, I., 2010. Employer branding: strategic implications for staff recruitment. *Journal of Marketing Management*, 26(1-2), pp.56-73. DOI: 10.1080/02672570903577091.

ABOUT AUTHORS

Martin Flegl[0000-0002-9944-8475] (M.F.) – Ph.D., Prof., School of Engineering and Sciences, Tecnologico de Monterrey, Mexico, e-mail: martin.flegl@tec.mx.

Lucie Depoo[0000-0003-4296-875X] (L.D.) – Ph.D., Assist. Professor, Management at the University of Economics and Management, Department of Management, Czech Republic, e-mail: lucie.vnouckova@vsem.cz.

Mayra Alcázar (M.A.) – posgraduate student, Business School, Universidad La Salle México, Mexico, e-mail: mayralcazar@gmail.com.

AUTHOR CONTRIBUTIONS

Conceptualization, M.F. and L.D.; Methodology, M.F. and M.A.; Software, M.F.; Validation, M.F., L.D. and M.A.; Formal analysis, M.F.; Investigation, M.F. and L.D.; Data curation, M.A.; Original draft preparation, M.F., L.D. and M.A.; Review and editing, M.F. and L.D.; Visualization, M.F.; Supervision, M.F.; Project administration, M.F.

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.

Innovative Expert Methods in Strategic Decision Making

DOI: 10.12776/QIP.V26I1.1643

Dagmar Gombitová, Dušana Dokupilová

Received: 2021-01-14 Accepted: 2022-03-28 Published: 2022-03-31

ABSTRACT

Purpose: This study introduces managerial techniques applied for the first time in the high-level strategic public policy decision-making process in Slovakia with an aim to assess the strategic decision-making of groups of experts in a methodologically supported environment. It compares groups of internal analysts and external specialists and should demonstrate the extent to which these two groups are able to process problems analytically and suppress intuition.

Methodology/Approach: Multi-criteria decision methods are used when deciding on complex problems. One of the most popular and most frequently used is the Analytic Hierarchy Process. Application of this method enables measurement of preference consistency, and its relationship with cognitive reflection.

Findings: Consistency of judgement was very similar in both groups. The prioritisation of measures resulted in a similar set of priorities determined by both groups. The assumed relationship of consistency and cognitive reflection score and/or overconfidence was not detected, and decision makers proved to be well calibrated.

Research Limitation/Implication: The main limitation of our research was the small sample size of decision makers, which complied with the requirements of the decision method, but was not sufficient to confirm the statistical validity.

Originality/Value of paper: The introduction of the multi-criteria decision method into decision-making for public policy strategies combines practical policy exercises with scientific research on high-stakes decisions and enables to carry out participatory decision-making process with relevant stakeholders.

Category: Research paper

Keywords: managerial innovation; decision making; consistency; analytic hierarchy process

1 INTRODUCTION

Decision-making in public policy is complex and often costly, with a high degree of uncertainty and risk, as well as often having long-term consequences for a large number of people. Current public policies should ensure that this process is participatory, based on evidence and applies the latest scientific knowledge on decision-making. Strategic decision-making involves the selection of the best among the possible alternatives, although this process is largely associated with uncertainty resulting from inadequate knowledge and excessive complexity. The framework for high-stakes decision-making should therefore involve a methodology for choosing the best among the alternatives; this should take into consideration future situations, use techniques to assess alternatives, understand the behaviour of the system and incorporate expert opinions into the process (Bhushan and Rai, 2007). A review of the literature and common practice in Slovak high-stakes decision-making suggests that there is a research gap. As mentioned by Bačová (2010), there is no unified approach or guidelines, and not much is known about the strategic decision-making in Slovakia.

Our research introduced new managerial techniques – that is, a new way of organising the participatory decision-making process and assessing the strategic decision-making of groups of experts in a methodologically supported environment. Our study demonstrates the extent to which groups of experts/decision makers are able to process problems analytically and suppress intuition and/or to confirm to what extent this process can be considered rational in terms of economic theory. This was examined through consistency and its relationship with cognitive reflection and overconfidence. The contribution of our research lies in the examination of the real-life strategic decision-making of experts through a well-established scientific method, Analytic Hierarchy Process (AHP). AHP applications have grown exponentially in last decades (Emrouznejad and Marra, 2017; Zyoud and Fuchs-Hanusch, 2017), and it has been used in many policy areas, such as strategic socio-economic, technology or environmental decisions in numerous countries (Subramanian and Ramanathan, 2012).

There are a few documented cases of the use of AHP in the Slovak environment, but these cases did not concern public policy decisions (Peregrin and Karahuta, 2014; Šoltés and Gavurová, 2014). To the best knowledge of the authors, our real-life experiment was the first attempt in this geographic area to examine and compare the consistency of two expert groups in high-stakes decision-making. In the next part we briefly introduce the background of the experiment and a literature review of the main concepts applied in the research. The following methodological section introduces our approach focused on the goal of the experiment. The results section presents the calculations, and the concluding chapter summarises the main findings and states the limitations of the research.

The authors of the paper closely cooperated with the Ministry of Investment, Regional Development and Informatisation in Slovakia (MIRDI) and were

assisting with the setup of priorities for future support from European Union sources. The new programming period for the use of the European Structural and Investment Funds in Slovakia has been planned since 2019. The Partnership Agreement as an umbrella document setting future strategic priorities for support from the European sources in the new programming period 2021-2027 was being prepared. During its preparation, the pilot testing introduced the process of high-stakes decision-making utilising AHP for the first time (Baláž, Dokupilová and Filčák, 2021; Dokupilová et al., 2020). Two years later, this novel instrument for decision-making at a high strategic level was repeated. The implementation of one Operational Programme Slovakia 2021-2027 was agreed, which largely followed the structure set out in the Partnership Agreement; however, new measures were introduced to the political objective dealing with social issues, so new prioritisation measures was necessary.

1.1 Rational Decision-Making

New decision-making theories respect the fact that decisions are not strictly governed by the rules of mathematical logic, but that various heuristics/shortcuts are applied in the decision-making process, which manifest as deviations from rational thinking. Heuristics are commonly defined as cognitive shortcuts or general rules that simplify decisions, especially under conditions of uncertainty. Contemporary behavioural science knows several dozen heuristics (Kahneman, 2011), and they can lead to cognitive bias. Such distortions are sometimes appropriate because they are part of people's adaptive responses to situations (Lockton, 2012). Cognitive biases or deviations are repeated systematic deviations from the norm or errors that affect our judgment and decision-making. The terms heuristics and cognitive biases generally refer to the strategies used in decision-making; they differ from the rational model of decision-making. Rational decision-making is not seen as synonymous with intelligence, so it cannot be measured as IQ. Rational decision-making is inherently very complex and cannot be assessed through a single characteristic or variable. To assess cognitive abilities, the relationships of these abilities with heuristics and biases are correlated (Čavojová and Hanák, 2014; Toplak, West and Stanovich, 2014). Given the presumed multifaceted nature of rational decision-making, it is important to determine which measures are most appropriate for assessing rationality. One of the most frequently used measure is the Cognitive Reflection Test (CRT), which is a tool to assess cognitive abilities.

1.2 Decision-Making Methods

There are numerous methods applied in decision-making. The introduction of a participatory approach reduces the selection scope. As the most suitable approach for analysing major development challenges, the Delphi method was applied in the Slovak pilot project mentioned above. 'Classical Delphi' generates consensus on major development challenges, and 'political Delphi' encourages structured public dialogue to generate policy alternatives for further prioritisation

(Dokupilová et al., 2020). When decisions are made for complex problems, a multi-criterion decision-making (MCDM) can be applied. MCDM methods are designed to suppress intuition and 'objectify' the decision-making process as much as possible. MCDM aims to select the best alternatives from a set of many potential candidates or alternatives. The order of the alternatives is based on the evaluation of several conflicting criteria, which MCDM organises. The final number of alternatives is considered and evaluated according to the different criteria. MCDM has been applied in various fields (Wallenius et al., 2008), and about 100 MCDM methods have been used in business and the public domain (Danesh, Ryan and Abbasi, 2017; Velasquez and Hester, 2013). A meta-analysis of expert decision-making methods (Danesh, Ryan and Abbasi, 2017) indicated that one of the most popular and the most frequently used methods is AHP (Saaty, 2008), which was developed as a reaction to the lack of common, easily understood and easy-to-implement methodologies for complex decisions. AHP has found use in business, government, social studies and many other domains involving important and complex decisions in which choice and/or prioritisation is needed (Bhushan and Rai, 2007). AHP has been proved (Danesh, Ryan and Abbasi, 2017; Velasquez and Hester, 2013) to be a theoretically sound, tested and accepted tool.

This type of decision-making involves the participation of experts. Expert decision-making consists of the identification of alternatives and the selection of the most suitable ones based on the criteria and preferences of the decision makers. The expert decision-making method has several properties (Vidal, Marle and Bocquet, 2011) and should have several criteria for decision-making, including qualitative ones, and determine the order of priorities for alternatives based on the selected criteria. At the same time, the method should be reliable, accessible and understandable even to those decision makers who are experts in their field but do not have extensive experience in applying expert decision-making methods. Given the importance of the decision-making problem and the structure of the group of decision makers, when choosing a decision-making method, we consider the following factors: reliability, accessibility and the complexity of the decision-making alternatives. Performing pairwise comparisons of individual alternatives is to be preferred for all criteria, as it is easier and more accurate to express one's opinion on only two alternatives than simultaneously on all the alternatives (Ishizaka and Labib, 2011). For pairwise comparisons, AHP provides the possibility of creating one's own set of criteria that best capture the goal. This decision-making technique helps to alleviate any subjectivity or intuition involved in decision-making. In general, it is completely natural for people to follow intuition or to be subjective, and in some respects, it is a remarkable survival technique that can lead to quick decisions based on personal experience, but biased.

1.3 Expert's Decision-Making

In the pilot programme, decision-making had to be participatory (EC Directive request), and due to its nature, it required involvement of experts. The term 'expert' (also specialist and/or researcher) is very broadly defined as an individual/person who has some special knowledge and skills most often assessed based on criteria such as length of practice or education/training. Several research studies have concluded that experts can make the same mistakes in decision-making as lay people, they have the same limited cognitive capacity and use heuristics (Baláž, 2009; Bazerman and Moore, 2013; Kahneman, 2011; Larrick and Feiler, 2015; Tetlock, 2005). Further research has, however, presented a more positive view of expert decision-making and emphasised that it has room for improvement (Burgman et al., 2011; Hutton and Klein, 1990; Shanteau and Stewart, 1992; Wagenaar and Keren, 1985). Experts are usually unaware of subjective influences and often overestimate their abilities. Nevertheless, the use of experts in decision-making is important and useful for several reasons: they possess more appropriate decision-making skills, are faster and more efficient (Gilmour and Corner, 1998). Appropriately selected methods increase the accuracy and calibration of experts' judgement. Studies have confirmed that the selection of experts based on their expertise or experience is crucial, as their expertise declines dramatically outside their field. Some professions (e.g. chess players, meteorologists) are much more consistent and accurate, but they mostly study physical phenomena (as opposed to human behaviour, society) or recurring events where decisions are made intuitively and have a decision-making apparatus or tool (Shanteau and Stewart, 1992). Group decision-making is also beneficial, because these estimates are better than the estimates of individuals (Saaty and Peniwati, 2013). Our selection of experts was in line with the natural selection of experts when a participatory approach is required in Slovakia. The terms 'expert' and 'specialist' are often used as synonyms, but they should be distinguished. A specialist is a person who has the capacity to substantiate many suggestions in a subdomain or area. Specialists are therefore a subset of experts. Experts do not have to be specialists; an obvious example of this is a general practitioner in medicine who is an expert but not a specialist (Weinstein, 1993).

Two groups of experts participated in the experiment: permanent experts (analysts) and ad-hoc experts (specialists). The first group consisted of analysts from the analytical units of the relevant ministries and other state institutions. Analytical units are part of these institutions, and their role is to provide analytical services for the needs of the ministry or other central state administration body to which they belong. The addressed analytical units were from the Ministry of Labour, Social Affairs and Family; Ministry of Education, Science, Research, and Sport; and the Ministry of Health of the Slovak Republic. Staff of the Institute of Financial Policy and the Value for Money Department from the Ministry of Finance, and Analytical Unit of the Office of Government, who also deal with the sectoral strategies were also included. The employees of

the analytical units should thus have a very good overall overview of what is happening in the respective sectors. The second group of experts (specialists) was composed of knowledgeable and experienced people from the professional community. These specialists are more engaged in specific issues from the relevant sector. Unlike the first group of analysts, the second group has a much narrower professional focus.

2 METHODOLOGY

2.1 Process

The groups of decision makers and/or potential respondents were invited by the MIRDI to take part in the experiment. The aim of the Ministry was to set out the priority measures for the political objective Social Europe in the Operational Programme Slovakia 2021-2027. To enable the use of AHP to elicit policy priorities, an online assessment tool was developed. The outputs were recorded in Qualtrics XM. The initial part of the tool consisted of the decision-making exercise. The political objective covered three main topics of social sector: employment and labour market; education; and social inclusion and social-health services. The respondent could choose to prioritise measures in one or more topics. Each topic was divided into three or four measures (alternatives), which were prioritised. A detailed description of each measure comprising the rationale, target group, expected impact and type of supported activities was made available during the assessment process. The respondents had to read and process information and carry out pairwise comparisons for all alternatives. This process usually took from 30 to 90 minutes and generated a significant cognitive burden for the respondents. First, the standard weights of the criteria were set. The respondents started with mutual comparison of the three determined criteria. A pairwise comparison for each pair of criteria defined which is preferred, as well as by how much on a scale from 9 (criteria A is 9 times more important than criteria B) to 9 (criteria B is 9 times more important than criteria A). Subsequently, participants compared all alternatives according to the individual criteria. The maximum number of AHP criteria was limited to three and the number of alternatives to four as the overload (3+18 comparisons per topic with 4 alternatives, 3+30 comparisons with 5 alternatives) can reduce the quality of policy evaluation. For each criterion, a matrix of preferences was compiled for each alternative, which consisted of standardised weights of alternatives according to the individual criteria. Priorities were set by a rigorous mathematical process the priority vector of weights w. (Mu and Pereyra-Rojas, 2017). The vector method (Saaty, 2008) was used to calculate the priority vector in the AHP method.

When the decision-making process was completed, the respondents were asked to resolve an additional task. This consisted of 10 verbal cognitive reflection tasks (Sirota et al., 2021). Subsequently, the last task provided an estimated

number of correct answers in a cognitive reflection test. The online tool was accessible for one month and data collection took place in November 2021.

2.2 AHP Criteria

AHP breaks down the decision-making problem into elements and levels according to common characteristics. It creates a logical hierarchy that systematically evaluates pairs of alternatives according to specific criteria (Berrittella et al., 2011). To reflect the implementation structure of the Operational Programme, three criteria were defined for what the individual decision makers are assessing according to a given criterion:

1. Relevance determines the alternatives that could significantly affect the social, economic and environmental development of the Slovak Republic. The alternative (measure) with the highest priority should significantly help the country to face major societal challenges in the next decade. Failure to implement this alternative would lead to serious economic, social and environmental consequences and irreversible changes.
2. Urgency sets the alternatives that must be implemented immediately or at least in the shortest possible time, as this alternative may be a necessary condition for the introduction of other alternatives. Postponing this alternative to a later date would lead to serious economic, social and environmental consequences and irreversible changes.
3. Feasibility concerns the alternatives the Government of the Slovak Republic, individual ministries and their agencies can put into practice. No alternative alone is likely to change people's thinking, stop demographic change or solve the global climate crisis; however, some alternatives can have a significant impact on our society, economy and environment, although in the past, we have failed to implement them. We therefore need to know if it is possible to implement a given alternative with the current understanding of the issue.

2.3 Consistency

Many decision makers are inconsistent in their decision-making, especially when it comes to making comprehensive assessments of complex information. When asked to evaluate the same information twice, they often respond differently, and they are not consistent when choosing preferences (Tversky, 1969). It is evident and common that people systematically violate the principle of consistency even in relatively simple decisions (Dawes, 1979; Kahneman, 2011), but a certain level of inconsistency in decision-making is acceptable (Ishizaka and Labib, 2011). As the consistency of preferences is considered one of the basic elements of a rational decision-making process, the rational actor should be an actor with consistent preferences (Sičáková-Beblavá, 2015). As people violate this principle, we examined the consistency of experts in the process of comparing

alternatives. Consistent decisions meant the preference of the most advantageous alternative while applying transitivity, which is the basic principle of decision models based on pairwise comparison (Wu and Tu, 2020). The method transforms mostly empirical data, into numerical values, which are further compared (Vargas, 2010). More precisely, if one considers alternative A to be three times more important than B and B twice as important as C, then alternative A must be (3*2) six times more important than C. According to the principle of transitivity, individuals should have a well-defined preferential structure – that is, they should be consistent. Along with consistency, it is also possible to test other psychological characteristics.

2.4 Cognitive Reflection

The Cognitive Reflection Test (CRT) is a simple test and functions effectively as a demonstration of cognitive ability (Frederick, 2005). CRT is a good measure of incorrect information processing and measures the tendency to replace fast, intuitive, but incorrect choices (Oppenheimer and Thomson, 2016; Sirota et al., 2018; Sleboda and Sokolowska, 2017; Toplak, West and Stanovich, 2014). Toplak, West and Stanovich (2011) found that CRT estimates rational behaviour better than measures of cognitive ability, thinking dispositions, and executive functioning. The CRT test is also related to academic results (Frederick, 2005; Welsh, Burns and Delfabbro, 2013) and is therefore used to measure analytical cognitive style (Čavojová and Hanák, 2014). The original test consists of three simple numeric tasks that evoke intuitive, albeit incorrect, answers. Toplak, West and Stanovich (2014) presented its extended version with seven tasks. Sinayev and Peters (2015) provided evidence that numeracy skills are strongly associated with reflection ability, and this type of test may also cause a gender performance gap (Juanchich, Sirota and Bonnefon, 2020). New versions of the CRT have been developed (Primi et al., 2016; Sirota and Juanchich, 2018) that assume not all individuals are mathematically proficient. To avoid the problems with numeracy and the fact that the original test is well known, the new tested version of verbal CRT was used (Sirota et al., 2021).

2.5 Overconfidence

Cognitive reflection predicts several heuristics, and people with lower cognitive reflection are significantly more likely to be subject to overconfidence, the illusion of control and conservatism (Noori, 2016). Overconfidence is one of the most common heuristics and exists in several forms. The most frequently studied form is overestimation, which is a systematic overestimation of one's abilities compared to reality. Previous research has confirmed that less competent people tended to overestimate their abilities, while more competent people tended to underestimate them (Dunning et al., 2003; Kruger and Dunning, 1999; Pennycook et al., 2017). It can therefore be assumed that more competent experts would not be too confident.

2.6 Sample Size and Structure

There is a wide-ranging debate in AHP groups about group size, as a high number of experts may increase the inconsistency of judgements in the AHP. Saaty and Özdemir (2014) point to the consistency of experts and the validity of their experiences and recommend a maximum of seven experts to a group. The experts' opinions differ from popular opinion polls. Less competent experts can, however, adversely influence the final decision, even if the majority consists of high-quality experts. They argue that a larger number of specialists (experts) can be useful in complex environments where different specialisations are required to assess specific aspects of a decision.

When selecting experts, the most common methods use accreditation (if available), or a method based on experience and/or identification by collaborators. The method of experience, which considers the number of years worked or the number of specific experiences, is one of the oldest assuming that if someone performs in the area for a certain number of years, s/he is considered an expert (Shanteau et al., 2002).

The system of advisory bodies varies considerably from country to country and sector to sector. In most countries, there are permanent and ad-hoc advisory bodies. Permanent bodies make the system more stable but also less flexible in terms of its ability to respond to new problems. They often have a wide scope, long duration and expertise in a certain area; they monitor, estimate trends and collect data. These bodies are usually a part of the public administration. Our first group of experts consisted of permanent advisors/analysts. The selection process for experts was determined by the position of the staff at the ministries and included staff from the analytical units.

Ad-hoc bodies may lack a sufficiently broad basis for structural and broader analyses and are used to answer questions quickly. They consist of external specialists, who are part of a so-called partnership ensuring the participatory approach when designing EU-supported interventions. This partnership includes regional and local administration, the academy, non-governmental organisations and social partners. Their staffing is more diverse, inclusive and external (OECD, 2017). The second group, comprised of specialists, represents the ad-hoc advisory body. Our selection was based on two criteria: at least 10 years of experience and recommendation by several members of the partnership. This approach reflects the common selection procedure when external experts are involved in the decision-making process in Slovakia.

3 RESULTS

An important indication of judgement quality is consistency. Consistency in the AHP is measured via the Consistency Index (CI) and Consistency Ratio (CR). Research has confirmed that $CR \leq 0.1$ is acceptable for the AHP exercise (Chu and Kuang-Han, 2002; Franek and Kresta, 2014). Consistency was measured for

all respondents in both groups. The respondents with a high level of inconsistency were excluded from the sample. Over 100 experts were addressed. The overall number of completed replies was 45 (20 analysts and 25 specialists), which can be considered a very good response rate, but only 64% of those were consistent (14 analysts and 15 specialists) and could be further processed (Table 1). Therefore, the size of individual decision-making groups is not always optimal. This does not allow generalisation of our findings but provides sufficiently good indications to identify differences. It limits the risk of unwanted influence caused by inconsistent decision makers on the final decision.

Table 1 – Final Number of Replies in Both Groups

Topic	**Analysts**		**Specialists**	
	Consistent	Inconsistent	Consistent	Inconsistent
1. Employment and Labour Market	5	3	3	4
2. Education	5	2	4	3
3. Social Inclusion and Soc./Health Serv.	4	1	7	3
Total	14	6	15	10

Out of 29 consistent replies, six respondents provided replies for two topics (three from each group) and two assessed measures for all three topics (one from each group). However, consistent replies were provided by only one person for all three topics and by three respondents dealing with two topics (Table 2). This may suggest that a deeper knowledge of a certain area could be crucial for consistent decision-making. This assumption has not been found in the studied literature, but it offers the opportunity for future research.

Table 2 – Number of Replies and Consistent Replies Provided by the Groups

Decision makers	**1 Topic**		**2 Topics**		**3 Topics**	
	Replied	Consistent	Replied	Consistent	Replied	Consistent
Analysts	12	8	3	3	1	0
Specialists	16	11	3	0	1	1
Overall	28	19	6	3	2	1

Consistency of judgement (especially inconsistency on criteria and total inconsistency) was very similar in both the analyst and specialist groups (Table 3).

Table 3 – Consistency of Groups (AIJ) on Individual Criteria

Topic	Analysts	Specialists	Overall
Inconsistency on criteria	0.030	0.027	0.029
Inconsistency on relevance	0.081	0.032	0.056
Inconsistency on urgency	0.066	0.092	0.080
Inconsistency on feasibility	0.020	0.037	0.029
Total inconsistency	0.040	0.038	0.039

To process all collected inputs, aggregated individual judgement (AIJ) and aggregated individual priority (AIP) procedures can be applied. Both AIP and AIJ procedures conduct aggregation through the weighted arithmetic mean (WAM) or weighted geometric mean (WGM) methods, depending on the group size and structure (Bernasconi, Choirat and Seri, 2014). The AIJ procedure cannot guarantee the Pareto optimality axiom, which states that if all group members prefer alternative A1 to alternative A2, then the group should prefer A1 to A2 as well (Ossadnik, Schinke and Kaspar, 2016), so the AIP method was chosen for this research as more appropriate.

As the groups were consistent, the AIPs were calculated through both WAM and WGM. The WAM is provided in Table 4 to better demonstrate the priorities decided in the AHP. The total sum of the priority indexes in each topic equals one, so the preference is easily visible. The topics themselves were not compared as most of the respondents assessed the measures in only one of the topics. The prioritisation of measures resulted in a very similar set of priorities determined by both groups. The priorities set out for the employment and labour market were practically identical. There was a minor difference in priorities for education, and a slightly different order was assigned to priorities in social inclusion and social-health services but with minor differences (Table 4). When comparing priorities according to the individual criteria, the biggest differences appeared in the perception of feasibility of measures 2.2 and 2.3. The analysts saw the *Promotion of equal access* more feasible (0.29) than the specialists did (0.18) and, vice versa, the *Support of life-long learning* is considered more feasible by the specialists (0.39) than by the analysts (0.25). Measure 3.3, *Improving access to social security services and healthcare*, showed similar differences for two criteria. The relevance and feasibility were assessed high by the analysts (0.39 resp. 0.36) while the specialists assigned lower values to both criteria (0.22 resp. 0.29). The other values did not indicate substantial differences.

Table 4 – Weighted Arithmetic Mean of Aggregated Individual Priority / and Ranking for Individual Measures (by Group)

Topic – Measure	Analysts	Specialists	Overall
1. Employment and Labour Market			
1.1 Access to employment	0.29 / 1.	0.30 / 1.	0.29 / 1.
1.2 Modernisation of labour market institutions and services	0.28 / 2.	0.26 / 2.	0.28 / 2.
1.3 Promoting gender balanced labour market participation	0.20 / 4.	0.18 / 4.	0.19 / 4.
1.4 Support of adaptation	0.23 / 3.	0.26 / 2.	0.24 / 3.
2. Education			
2.1 Improving the quality of education	0.44 / 1.	0.37 / 2.	0.41 / 1.
2.2 Promoting equal access	0.35 / 2.	0.38 / 1.	0.36 / 2.
2.3 Support of life-long learning	0.21 / 3.	0.25 / 3.	0.23 / 3.
3. Social Inclusion and Social/Health Services			
3.1 Promoting active inclusion	0.18 / 4.	0.19 / 4.	0.19 / 4.
3.2 Support of socio-economic integration of MRC*	0.21 / 3.	0.27 / 2.	0.25 / 3.
3.3 Improving access to social security services and healthcare	0.37 / 1.	0.25 / 3.	0.29 / 1.
3.4 Promoting social integration of people at risk of poverty	0.24 / 2.	0.29 / 1.	0.27 / 2.

Notes: MRC* Marginalised Roma Communities

As the pilot exercise carried out in 2019 had the identical structure of measures for topics 2 and 3, it is possible to compare the priorities (Table 5). The original priorities did not highlight the quality of education to the same extent, and the main emphasis was on inclusion and life-long learning. We assume that the current change in the order of priorities may have been influenced by the global COVID-19 pandemic, and the long period of online education for Slovak schools focused the attention of the analysts on its quality (measure 2.1). A similar perception could be reflected in the higher priority of the measure related to the health sector (3.3). Both education and health are currently considered the main governmental priorities. The priorities set by the specialists highlight the same measures but more proportionally, with about the same weight. This provides approximately the same division of priorities in the social inclusion topic as in 2019 (Table 5).

Table 5 – Weighted Arithmetic Mean of Aggregated Individual Priority / and Ranking for Individual Measures in 2021 and 2019

Topic – Measure	Analysts 2021	Specialists 2021	Overall 2021	Overall 2019
2. Education				
2.1 Improving the quality of education	0.44 / 1.	0.37 / 2.	0.41 / 1.	0.26 / 3.
2.2 Promoting equal access	0.35 / 2.	0.38 / 1.	0.36 / 2.	0.46 / 1.
2.3 Support of life-long learning	0.21 / 3.	0.25 / 3.	0.23 / 3.	0.38 / 2.
3. Social Inclusion				
3.1 Promoting active inclusion	0.18 / 4.	0.19 / 4.	0.19 / 4.	0.17 / 4.
3.2 Support of socio-economic integration of MRC*	0.21 / 3.	0.27 / 2.	0.25 / 3.	0.30 / 1.
3.3 Improving access to social security services and healthcare	0.37 / 1.	0.25 / 3.	0.29 / 1.	0.28 / 2.
3.4 Promoting social integration of people at risk of poverty	0.24 / 2.	0.29 /1.	0.27 / 2.	0.24 / 3.

Notes: MRC* Marginalised Roma Communities

Based on the assumption of the economic theory that the rational decision maker is consistent, we measured consistency. The CRT is also confirmed as a substantial independent predictor of a group of rational thinking tasks (Toplak, West and Stanovich, 2014). The cognitive reflection score was thus measured together with consistency. The comparison of the consistency in the groups shows the same values, although the cognitive score is more than 20% higher for the analyst group (Table 6). Both groups scored high in the CRT, as the average score of common population was 4.2 (Sirota et al., 2021). It is known that people with a lower CRT are significantly more likely to be subject to overconfidence (Noori, 2016). In our experiment, the inconsistent respondents scored as high as the group of specialists (6.7) but their overconfidence was nearly twice as high as for the consistent respondents (2.6).

Table 6 – CRT Score and Overconfidence (by Group)

Main Group Features	Analysts	Specialists	Overall	Inconsistent respondents
CRT score	8.42	6.95	7.00	6.70
Overconfidence	1.58	1.90	1.38	2.60

The attempt to identify a correlation between consistency and both CRT score and the number of correct answers estimated by respondents was not successful (Table 7). Overconfidence was also calculated to test possible correlation with total consistency. The correlation could not be identified, either; first, because the overestimation for 80% of respondents was in the range (–2 to 2) – that is, the

sample was rather homogenous with only minor deviations. Second, the small sample, which complied with the size necessary for the AHP, was not sufficiently representative and large enough to detect a correlation. The results also suggest that the selection of experts and further selection of consistent experts, keeps the CRT score high and overconfidence at a similar level.

Table 7 – Correlation Matrix for Total Consistency, Confidence, and CRT Score (by Group)

	Analysts	**Specialists**	**Overall**
Total consistency (TC)	0.040	0.038	0.039
Correlation			
TC - CRT score	-0.285	-0.511	-0.428
TC - estimated correct CRT answers	0.202	-0.361	-0.139
TC - overconfidence	0.557	0.177	0.335

4 CONCLUSION

The authors used this unique opportunity to combine practical policy exercises with scientific research on high-stakes decisions through the use of AHP. High-stakes decision makers are not easily accessible for research, so their participation due to the support and involvement of the MIRDI is highly appreciated. The newly tested approach for strategic decision-making with the support of internal and external experts was beneficial. The original 'foggy' decision-making process in high-stakes strategic public policies was replaced by a transparent and operational mechanism utilising up-to-date research methodologies. This proved to be a functional managerial innovation that made it possible to carry out a participatory decision-making process with relevant stakeholders.

In addition, for the first time ever, the decision-making of expert groups was compared, and it was confirmed that groups of analysts and specialists are similar as regards their basic measured characteristics. The groups were equally (in)consistent, with slightly higher CRT score for analysts and slightly lower overconfidence compared to specialists. However, comparison of their preferences showed some differences and confirmed slightly different set of the priorities. We can assume that higher priorities assigned to quality of education and health measures by analysts were quite likely driven by the focus of governmental priorities highlighted by the COVID-19 pandemic. Comparison of the priorities set out in 2019 and 2021 reflects the changes in 2 "COVID" years.

These results indicate that the selection process for experts is very important. The consistent respondents in both groups produced similar results. The assumed relationship between consistency and CRT score and/or overconfidence was not

detected, as the consistent decision makers proved to be well calibrated. This suggests the possible application of consistency as a calibration tool, but this requires further research. Similar conclusions can be drawn from the attempt of a few respondents to assess more than one topic, which suggests that high expertise in a topic, leads to higher consistency, but if more topics are selected the inconsistent replies are more frequent. This hypothesis needs verification in a well-designed experiment. Although for our experiment the size and composition of the expert groups was in line with the general practice in Slovakia, it did not enable generalisation of our findings. The main limitations of our research are the small sample of decision makers; while this sample size complies with the requirements of AHP, it is not sufficient to confirm the statistical validity of our findings. As it is not practically feasible to increase the size of the groups, future research could benefit from the further repetition/application of the group decision-making, which is specifically meant for high-stakes decision-making in public policy.

REFERENCES

Bačová, V., 2010. *Rozhodovanie a usudzovanie. Pohľady psychológie a ekonómie I.* Bratislava: Ústav experimentálnej psychológie SAV.

Baláž, V., 2009. *Riziko a neistota. Úvod do behaviorálnej ekonómie a financií.* Bratislava: VEDA.

Baláž, V., Dokupilová, D. and Filčák, R., 2021. Participatory multi-criteria methods for adaptation to climate change. *Mitigation and Adaptation Strategies for Global Change*, 26(4), pp.1-22. DOI: 10.1007/s11027-021-09955-4.

Bazerman, M.H. and Moore, D.A., 2013. *Judgment in Managerial Decision Making*. Wiley.

Bernasconi, M., Choirat, C. and Seri, R., 2014. Empirical properties of group preference aggregation methods employed in AHP: Theory and evidence. *European Journal of Operational Research*, 232(3), pp.584-592. DOI: 10.1016/j.ejor.2013.06.014.

Berrittella, M., Certa, A., Enea, M. and Zito, P., 2011. An Analytic Hierarchy Process for the Evaluation of Transport Policies to Reduce Climate Change Impacts. [FEEM Working Paper No. 12.2007] *SSRN Electronic Journal*. Available at SSRN: <https://ssrn.com/abstract=962379> DOI: 10.2139/ssrn.962379.

Bhushan, N. and Rai, K., 2007. *Strategic Decision Making: Applying the Analytic Hierarchy Process.* Springer Science & Business Media.

Burgman, M.A., McBride, M., Ashton, R., Speirs-Bridge, A., Flander, L., Wintle, B., Fidler, F., Rumpff, L. and Twardy, Ch., 2011. Expert Status and Performance. *PLoS ONE,* 6(7), e22998. DOI: 10.1371/journal.pone.0022998.

Čavojová, V. and Hanák, R., 2014. Assessing fluid rationality and its relation to cognitive styles. In: T. Ahram, W. Karwowski and Marek, T., eds. *Proceedings of the 5th International Conference on Applied Human Factors and Ergonomics*. Krakow: AHFE. pp.142-153.

Chu, P. and Kuand-Han, L. J., 2002. Note on consistency ratio. *Mathematical and Computer Modelling*, 35, pp.1077-1080.

Danesh, D., Ryan, M. and Abbasi, A., 2017. A Systematic Comparison of Multi Criteria Decision Making Methods for the Improvement of Project Portfolio Management in Complex Organisations. *International Journal of Management and Decision Making*, 16(1), pp.280-320.

Dawes, R., 1979. The Robust Beauty of Improper Linear Models in Decision Making. *American Psychologist*, 34, pp.571-82.

Dokupilová, D., Baláž, V., Kurincová Čavojová, V., Ballová Mikušková, E. and Gombitová, D., 2020. Identifying major policy challenges and policy interventions via expert methods application of the Delphi and AHP methods in preparation of the Partnership Agreement for the Slovak Republic in period 2021-2027. *Národohospodářský obzor*, 20(3), pp.361-377. DOI: 10.2478/revecp-2020-0017.

Dunning, D., Johnson, K., Ehrlinger, J. and Kruger, J., 2003. Why people fail to recognize their own incompetence. *Current Directions in Psychological Science*, 12(3), pp.83-87. DOI: 10.1111/1467-8721.01235.

Emrouznejad, A. and Marra, M., 2017. The state-of-the-art development of AHP (1979–2017): A literature review with a social network analysis. *International Journal of Production Research*, 55(22), pp.6653-6675.

Franek, J. and Kresta, A., 2014. Judgment scales and consistency measure in AHP. *Precedia Economics and Finance*, 12, pp.164-173.

Frederick, S., 2005. Cognitive Reflection and Decision Making. *Journal of Economic Perspectives*, 19(4), pp.25-42. DOI: 10.1257/089533005775196732.

Gilmour, P. and Corner, J.L., 1998. The Role of the Expert's Decision Making Skills in Management. *33rd Annual Operational Research Society of New Zealand,* pp.195-204.

Hutton, R.J.B. and Klein, G., 1990. Expert decision making. *Systems Engineering*, 2(1), pp.32-45.

Ishizaka A. and Labib A., 2011. Review of the main developments in the analytic hierarchy process. *Expert Systems with Applications*, 38(11), pp.14336-14345.

Juanchich, M., Sirota, M. and Bonnefon, J. F., 2020. Anxiety-induced miscalculations, more than differential inhibition of intuition, explain the gender gap in cognitive reflection. *Journal of Behavioural Decision Making*, 33, pp.427-443. DOI: 10.1002/bdm.2165.

Kahneman, D., 2011. *Thinking Fast and Slow*. New York: Farrar, Straus and Giroux.

Kruger, J. and Dunning, D., 1999. Unskilled and unaware of it: How difficulties in recognizing one's own incompetence led to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6), pp.1121-1134.

Larrick, R.P. and Feiler, D.C., 2015. Expertise in decision making. In: G. Keren, G. Wu, eds. 2015. *The Wiley Blackwell Handbook of Judgment and Decision Making*. John Wiley & Sons. pp.696-721.

Lockton, D., 2012, Cognitive biases, heuristics and decision-making in design for behaviour change. [Working Paper] *SSRN Electronic Journal*. Available at SSRN: <https://ssrn.com/abstract=2124557>. DOI: 10.2139/ssrn.2124557.

Mu, E. and Pereyra-Rojas, M., 2017. AHP Models with Sub-criteria. In: E. Mu and M. Pereyra-Rojas, eds., 2017. *Practical Decision Making*. Cham: Springer. pp.45-56.

Noori, M., 2016. Cognitive reflection as a predictor of susceptibility to behavioural anomalies. *Judgment and Decision Making*, 11(1), pp.114-120.

OECD, 2017. *Policy Advisory Systems: Supporting Good Governance and Sound Public Decision Making*. Paris: OECD Public Governance Reviews, OECD Publishing. DOI: 10.1787/9789264283664-en.

Oppenheimer, D.M. and Thomson, K.S., 2016. Investigation of an alternate form of the Cognitive Reflection Test. *Judgment and Decision making*, 11(1), p.99.

Ossadnik, W., Schinke, S. and Kaspar, R.H., 2016. Group Aggregation Techniques for Analytic Hierarchy Process and Analytic Network Process: A Comparative Analysis. *Group Decision and Negotiation*, 25(2), pp.421-457. DOI: 10.1007/s10726-015-9448-4.

Pennycook, G., Ross, R. M., Koehler, D. J. and Fugelsang, J. A., 2017. Dunning–Kruger effects in reasoning: Theoretical implications of the failure to recognize incompetence. *Psychonomic Bulletin and Review*, 24(6), pp.1774-1784.

Peregrin, S. and Karahuta, M., 2014. Skupinový analytický hierarchický proces a jeho využitie v manažmente spoločnosti/Group analytic hierarchy process and its usage in management of company. *Mladá Veda*, 2(3), p.60.

Primi, C., Morsanyi, K., Chiesi, F., Donati, M.A. and Hamilton, J., 2016. The development and testing of a new version of the cognitive reflection test applying item response theory (IRT). *Journal of Behavioral Decision Making*, 29(5), pp.453-469.

Saaty, T., 2008. Decision making with the analytic hierarchy process. *International Journal of Services Sciences*, 1(1), pp.83-98.

Saaty, T.L. and Özdemir, M.S., 2014. How Many Judges Should There Be in a Group? *Annals of Data Science*, 1(3-4), pp.359-368.

Saaty, T.L. and Peniwati, K., 2013. *Group decision making: drawing out and reconciling differences*. RWS publications.

Shanteau, J. and Stewart, R., 1992. Why Study Expert Decision Making? Some Historical Perspectives and Comments. *Organisational Behaviour and Human Decision Processes*, 53, pp.95-106.

Shanteau, J., Weiss, D., Thomas, R.P. and Pounds, J.C., 2002. Performance based assessment of expertise: How to decide if someone is an expert or not. *European Journal of Operational Research*, 136(2), pp.253-263.

Sičáková-Beblavá, E., 2015. *Behaviorálne základy verejnej politiky*. Bratislava: Univerzita Komenského v Bratislave.

Sinayev, A. and Peters, E., 2015. Cognitive reflection vs. calculation in decision making. *Frontiers in Psychology*, 6, p.532. DOI: 10.3389/fpsyg.2015.00532.

Sirota, M. and Juanchich, M., 2018. Effect of response format on cognitive reflection: Validating a two- and four-option multiple choice question version of the Cognitive Reflection Test. *Behaviour Research Methods,* 50, pp.2511-2522.

Sirota, M., Dewberry, Ch., Janchich, M., Valuš, L. and Marshall, A.C., 2021. Measuring cognitive reflection without maths: Development and validation of the verbal cognitive reflection test. *Journal of Behavioural Decision Making*, 34, pp.322-343. DOI: 10.1002/bdm.2213.

Sleboda, P. and Sokolowska, J., 2017. Measurements of Rationality: Individual Differences in Information Processing, the Transitivity of Preferences and Decision Strategies. *Frontiers in Psychology*, 8, p.1844. DOI: 10.3389/fpsyg.2017.0184

Šoltés, V. and Gavurová, B., 2014. The functionality comparison of the health care systems by the analytical hierarchy process method. *Ekonomie a Management,* 17(3), pp.1-18.

Subramanian, N. and Ramanathan, R., 2012. A review of applications of Analytic Hierarchy Process in operations management. International Journal of Production Economics, 138, pp.215-241. DOI: 10.1016/j.ijpe.2012.03.036

Tetlock, P.E., 2005. *Expert Political Judgment: How Good Is It? How Can We Know?* Princeton University Press.

Toplak, M.E., West, R.F. and Stanovich, K.E., 2011. The Cognitive Reflection Test as a predictor of performance on heuristics-and-biases tasks. *Memory & Cognition*, 39(7), pp.1275-1289. DOI: 10.3758/s13421-011-0104-1.

Toplak, M.E., West, R.F. and Stanovich, K.E., 2014. Assessing miserly information processing: An expansion of the Cognitive Reflection Test. *Thinking & Reasoning*, 20(2), pp.147-168. DOI: 10.1080/13546783.2013.844729.

Tversky, A., 1969. Intransitivity of preferences, *Psychological Review*, 76, pp.31-48.

Vargas, R.V., 2010. Using the analytic hierarchy process (AHP) to select and prioritize projects in a portfolio. In: Project Management Institute (PMI®), *Global Congress 2010-North America*. 12 October 2010. Washington, DC. Newtown Square, PA: Project Management Institute. Available at: <https://www.pmi.org/learning/library/analytic-hierarchy-process-prioritize-projects-6608> [10 December 2021].

Velasquez, M. and Hester, P.T., 2013. An Analysis of Multi-Criteria Decision-Making Methods. *International Journal of Operations Research*, 10(2), pp.55-66.

Vidal, L.M., Marle, F. and Bocquet, J.C., 2011. Using a Delphi process and the Analytic Hierarchy Process (AHP) to evaluate the complexity of projects. *Expert Systems with Applications*, 38(5), pp.5388-5405.

Wagenaar, W.A. and Keren, G.B., 1985. Calibration of probability assessments by professional blackjack dealers, statistical experts, and lay people. *Organizational Behaviour and Human Decision Processes*, 36(3), pp.406-416.

Wallenius, J., Dyer, J.S., Fishburn, P.C., Steuer, R.E., Zionts, S. and Deb, K., 2008. Multiple criteria decision-making, multi-attribute utility theory: Recent accomplishments and what lies ahead. *Management Science*, 54(7), pp.1336-1349.

Weinstein, B.D., 1993. What is an expert? *Theoretical Medicine and Bioethics*, 14(1), pp.57-73.

Welsh, M.B., Burns, N.R. and Delfabbro, P.H., 2013. The Cognitive Reflection Test: How much more than numerical ability? In: M. Knauff, M. Pauen, N. Sebanz and I. Wachsmuth, ed., *35th Annual Conference of the Cognitive Science Society*. Berlin, Germany, July 31-August 3, 2013. Austin, TX: Cognitive Science Society. pp.1587-1592.

Wu, Z. and Tu, J., 2020. Managing transitivity and consistency of preferences in AHP group decision making based on minimum modifications. *Information Fusion*, 67, pp.125-135. DOI: 10.1016/j.inffus.2020.10.012.

Zyoud, S.H. and Fuchs-Hanusch, D., 2017. A bibliometric-based survey on AHP and TOPSIS techniques. *Expert Systems with Applications*, 78(C), pp.158-181. DOI: 10.1016/j.eswa.2017.02.016.

ABOUT AUTHORS

Dagmar Gombitová[0000-0001-6204-9238] (D.G.) – Ing. Mgr., PhD. Student at the Institute for Forecasting, Centre of Social and Psychological Studies, Slovak Academy of Sciences (CSPS SAS), Bratislava, Slovakia, and Public Policy Institute, Faculty of Social and Economic Sciences, Comenius University Bratislava, Slovakia, e-mail: dagmar.gombitova@savba.sk.

Dušana Dokupilová[0000-0001-9813-2201] (D.D.) – RNDr. PhD., Researcher at the Institute for Forecasting, Centre of Social and Psychological Studies, Slovak Academy of Sciences (CSPS SAS), Bratislava, Slovakia, e-mail: dusana.dokupilova@savba.sk.

AUTHOR CONTRIBUTIONS

Conceptualization: D.G.; Methodology: D.G. and D.D.; Software, D.D.; Validation, D.D.; Formal analysis, D.D.; Investigation, D.G.; Resources, D.G.; Data curation: D.D.; Original draft preparation: D.G.; Review and editing: D.D.

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.