Exploring the uncharted territory of B2B customer experience: Towards a novel approach for measurement and analysis

Beatriz Seabra de Matos Pereira

Master's Dissertation

Supervisor at FEUP: Prof. Doutor Miguel Alves Pereira
Supervisor at BA Glass: Luísa Lopes

Masters in Industrial Engineering and Management

2023-06-26

Abstract

In today's fast-growing and highly competitive business landscape, standing out from the crowd has become increasingly difficult. As a result, the creation of an exceptional Customer Experience (CX), a pivotal factor capable of setting companies apart, has become a critical objective. However, effectively designing, monitoring, measuring, and managing CX in Business-to-Business (B2B) environments, presents unique challenges due to its complexity and the scarcity of research available. To address this, this project develops an innovative procedure for measuring B2B CX using a Composite Indicator (CI) approach, integrating the Deck Cards Method Simos-Roy-Figueira and the Multi-Attribute Utility Theory. It is applied to a case study conducted within a multinational glass packaging company, and it involves the collaboration of employees, department managers, and sales directors. The B2B CX CI construction follows a three-step process, encompassing (1) the selection of variables; (2) the assessment of Key Performance Indicators (KPIs) weights, utility functions, and respective utilities, to obtain the departmental utility functions; and (3) the determination of department weights to ascertain the overall CX. This novel approach provides valuable insights to the company, enabling long-term monitoring and management of CX, while enhancing employee awareness on the topic, and fostering interdepartmental collaboration and cooperation. The findings contribute to the limited body of research on B2B CX measurement in the glass packaging industry while emphasizing the need for further investigation into specific aspects of B2B CX measurement.

Keywords: Customer experience, Business-to-Business, Composite indicator, Deck of Cards Method, Multi-Attribute Utility Theory

Resumo

No atual cenário empresarial, altamente competitivo e em rápido crescimento, uma empresa conseguir diferenciar-se e destacar-se das demais, tornou-se cada vez mais difícil. Consequentemente, a criação de uma experiência do cliente (CX) excepcional tornou-se um objetivo crítico, uma vez que é visto como fator fundamental na diferenciação das empresas. No entanto, desenhar, monitorizar, medir e gerir efetivamente a CX em ambientes *Business-to-Business* (B2B) apresenta desafios únicos devido à sua complexidade e e à escassez de investigação disponível. Endereçando este problema, este projeto desenvolve um procedimento inovador para medir a CX em ambientes B2B, através da criação de um indicador compósito (CI), integrando o Método das Cartas de Simos-Roy-Figueira e a Teoria da Utilidade Multiatributo. Este é aplicado a um caso de estudo, conduzido numa empresa multinacional produtora de embalagens de vidro, e conta com a colaboração de funcionários, diretores dos departamentos e diretores de vendas. A construção do B2B CX CI está organizada em três etapas: (1) a seleção de variáveis; (2) a avaliação dos pesos dos indicadores-chave de desempenho (KPIs), a determinação das funções utilidade dos KPIs e respectivas utilidades, para determinar as funções utilidade departamentais; e (3) a determinação dos pesos de cada departamento para obter a CX global. Esta nova abordagem permite que a empresa monitorize e gira a CX ao longo do tempo, assim como contribui para um aumento da sensibilização dos trabalhadores para a temática e, promove a colaboração e a cooperação interdepartamental. Para além disso, este trabalho contribui para o limitado corpo de investigação sobre a medição da CX em contexto B2B, em particular na indústria de embalagens de vidro, e salienta ainda a necessidade de mais investigação de ferramentas e métodos para a medição da CX em ambientes B2B.

Palavras-chave: Experiência do cliente, *Business-to-Business*, Indicador compósito, Método das Cartas, Teoria da Utilidade de Multi-Atributos

Acknowledgements

I would like to express my gratitude to all the people who have helped me to successfully complete this dissertation.

First, I would like to thank Prof. Doutor Miguel Alves Pereira for all the help, constant support, and tireless guidance throughout the work. And for always keeping me motivated and calm.

Second, I am sincerely grateful to my supervisor, Luísa Lopes, for her patience, encouragement, and guidance; and to my "FOAs", who welcomed, nurtured, and helped me from the first day. I would also like to thank all BA employees who made themselves available to help me whenever necessary. Without this support, this project would not have been possible.

I also would like to thank my internship colleagues, who were undoubtedly the best companions I could have had on this journey. My generous and brilliant friends, thank you for the stimulating discussions and all the fun we have had over the past few months.

Last, but certainly not least, I would like to thank my family and friends for their unconditional love and support. A special thanks to my mother and sister, the most important people in my life. Thank you for inspiring me to do my best and making me a better human.

"The more you lose yourself in something bigger than yourself, the more energy you will have."

Norman Vincent Peale

Contents

Abbreviations and Acronyms

AHP	Analytic Hierarchy Process
BA	BA Glass, S.A.
BAP	Budget Allocation Process
B2B	Business-to-Business
B2C	Business-to-Consumer
BOD	Benefit of the Doubt
CA	Conjoint Analysis
CI(s)	Composite Indicator(s)
CJ(s)	Customer Journey(s)
CJM	Customer Journey Mapping
CS	Customer Satisfaction
CSI	Customer Satisfaction Index
CX(s)	Customer Experience(s)
DCM-SRF	Deck Cards Method Simos-Roy-Figueira
DEA	Data Envelopment Analysis
DM(s)	Decision-Maker(s)
EBITDA	Earnings Before Interest, Taxes, Depreciation, and Amortization
ELECTRE	Elimination and Choice Expressing Reality
EXQ	Service Experience
EW	Equal Weighting
FA	Factor Analysis
GRA	Grey Relational Analysis
INDSAT	Customer Satisfaction in Industrial Markets
KPI(s)	Key Performance Indicator(s)
MACBETH	Measuring Attractiveness by a Categorical Based Evaluation Technique
MAUT	Multi-Attribute Utility Theory
MAVT	Multi-Attribute Value Theory
MCDM	Multi-Criteria Decision Making
NPS	Net Promoter Score
OECD	Organization for Economic Co-operation and Development
PCA	Principal Component Analysis
PROMETHEE	Preference Ranking Organization Method for Enrichment Evaluations
SAW	Simple Additive Weighting
SERVQUAL	Service Quality
SEVPERF	Service Performance
SMART	Simple Multi-Attribute Rating Technique
SRJ	Service Recovery Journey
TOPSIS	Technique for Order Preferences by Similarity to Ideal Solutions
UTA	Utility Theory Additive
WoM	Word-of-Mouth
WP	Weighted Product
ZOT	Zone of Tolerance

Nomenclature

Roman Symbols

u	Utility function of each KPI
g	Real performance of the KPI
w	Weight of each KPI
U	Utility function of each department
W	Weight of each department
CX	Overall customer experience utility function

Subscripts

i	Indexes per KPIs
j	Indexes per departments

List of Figures

[1] Available at https://media.cheggcdn.com/media/d51/d513a8ac-64c2-42d7-a063-9b92bb2422a1/phptcaTNp

List of Tables

Chapter 1

Introduction

Today, creating a remarkable Customer Experience (CX) is a major management objective. It is consensual, with 93% of business leaders agreeing, that providing a meaningful and reliable CX is a key source of competitive advantage. It is essential for improving overall business performance (Meyer & Schwager, 2007), contributing to higher Customer Satisfaction (CS) rates, lower customer churn rates, and higher revenues (Rawson, Duncan, & Jones, 2013). Not surprisingly, CX design, monitoring, and management are now focal points of academic research (McColl-Kennedy et al., 2015). Though, despite the numerous publications, the knowledge of CX remains limited, especially in Business-to-Business (B2B) environments (Keyser et al., 2020). Therefore, deepening the knowledge of this underexplored field is of the utmost interest.

Taking this into account, the present master's thesis proposes an innovative approach to measuring B2B CX, which is applied in a case study of a multinational glass packaging company. It was developed in BA Glass, S.A., within the Service Front Office department from February 2023 to mid-June 2023.

This initial chapter aims to present BA Glass, S.A., the context, the definition of the problem, the motivation of this project, and its main objectives. It also intends to highlight this work's relevance for the company and the scientific community and present the research methodology adopted. Lastly, an overview of the topics discussed in the following chapters of this dissertation is exhibited.

1.1 Project context

BA Glass, S.A. (hereafter referred to as BA) is a multinational company, headquartered in Portugal, and it is one of the biggest producers of glass containers in Europe. It develops, produces, and commercializes over 11 billion glass bottles and jars[1], annually, for the food and beverage industries.

[1]Retrieved from the company's annual report 2022. Available at `https://www.baglass.com/en/reports.php`

BA operates in a highly competitive market, competing simultaneously against other glass container manufacturers and with companies that produce bottles and jars using different materials, such as plastic, aluminum, Tetra Pak®, etc. These alternatives are typically lighter and cheaper than glass, thus, the switching risk of customers is high, reinforcing the importance of ensuring the high quality of both the products sold and services provided.

Over the years, BA has become a reference supplier due to the quality of its products and the ability to find innovative, creative, and flexible solutions together with the customers. BA soon realized that collecting and analyzing customers' feedback was critical to sustainable business growth and committed itself to providing them with the best experience possible. In 2011, the Customer Satisfaction Index (CSI) was implemented to measure CS and adopt strategies that contribute to its continuous improvement.

Last year, BA obtained the worst results ever in CSI. It is believed that the main causes of dissatisfaction were the continued rise in prices and a decrease in the service quality provided. On the one hand, to face the European energy crisis and the consequent surge of energy costs, BA increased the prices to its customers, which displeased them. On the other hand, some service problems negatively affected the quality of the services provided. However, it must be noted that satisfaction levels were much lower than expected.

In order to increase CSI in 2023, it is required to tackle these two issues. Firstly, regarding the prices, BA is currently reviewing them as the energy prices started to stabilize. Secondly, the quality of services provided must improve. Therefore, a complete understanding of the actual CX is required, which is not a simple task. Managing CX is, in fact, very complex. There are many variables involved, internal and external to the company, that impact CX; there are multiple actors, multiple encounters, processes that cross departments, etc.

Nowadays, there is no overall CX measurement instrument at BA. Although many Key Performance Indicators (KPIs) have been implemented in recent years following a digitalization process, they are being monitored individually by each department. Thus, despite knowing how well they are performing, they lack knowledge of the current situation of the other departments. This lack of interdepartmental visibility prevents a global vision of the company's performance and complicates the task of understanding where improvements can/should be made.

Therefore, to better manage the CX, it would be relevant to map the processual and experiential aspects of service processes, across all departments. And, on top of that, develop a unique measure, aggregating the main KPIs of each department, weighted by the degree of impact of different activities, which would also stimulate a joint, coordinated, and committed action.

1.2 Motivation

It is widely agreed that to succeed, a company must not only offer high-quality goods and services, but also create unique experiences and learn how to manage the customer's expectations and judgments, before, during, and after a purchase (Klaus, 2015). Due to this, many studies about CX were developed and published that have contributed to the development of this field. However,

much remains to be explored (Keyser et al., 2020). Some reports show that only one-third of CX initiatives are successful (Thompson, 2018), while others reveal that most CX programs fail to deliver, and companies see weak returns [2]. Additionally, a study conducted by Bain & Company revealed that while 80% of the companies believed they were delivering a "superior experience" to their customers, clients only agreed in 8% of the situations (Reichheld, Hamilton, Markey, & Allen, 2005). In 2017, a study showed that 71% of the companies elected CX as a top strategic objective, but only 13% rated themselves a 9 or 10 (out of 10) in terms of delivery (Davey, 2017).

At the same time, and along with this empirical evidence, a review of the literature shows that CX research on B2C environments is dominant, with very few studies published on B2B (Keyser et al., 2020). And, although some authors believe that many insights obtained from B2C research can be applied in B2B, it is required evidence of that. Moreover, scholars argue that comprehensive frameworks for B2B must be created since the ones available, in some situations such as Customer Journey (CJ) mapping, are overly simplified (Zolkiewski et al., 2017).

Therefore, this project is relevant for both BA and the scientific community. On the one hand, this quantitative instrument will allow the company to have a global view of the CX provided over time, enabling the implementation of strategies to assure CX is constantly improved; and it will help the company better understand the root causes of customers' dissatisfaction and the reasons for CSI's significant decrease. On the other hand, it is important to highlight that the evident lack of knowledge regarding the appropriate methods for measuring CX is even more significant in the glass packaging industry, with no study published on this matter. Thus, this work will meet the need for more studies on B2B CX and, in particular, within this industry. On top of that, this study is also original in what concerns the methodology adopted. To the author's knowledge, it is the first time a Composite Indicator (CI) is created to measure CX in a B2B setting.

1.3 Objectives

The main objective of this present project is to respond to the lack of knowledge on B2B CX, in particular, on quantitive tools to measure it. In order to achieve this, there are also more specific goals, as follows:

- Analyze the company CJ and list all the measured KPIs - identify the key touchpoints, the organizational activities, employees' responsibilities, the customers' goals (expectations), activities, needs, their pain points, etc.;
- Assess the impact of each department's activities on CX, and determine which KPIs are the best proxy measures to assess CX;
- Develop an alternative and innovative procedure to measuring B2B CX, over time, using a CI approach;

[2] Available at `https://www.forbes.com/sites/blakemorgan/2018/07/30/why-many-customer-experience-programs-are-still-failing-to-deliver/?sh=4c8c93de35d2`

- Analyze results and study the correlation between them and last year's CSI survey results;
- Create a dynamic dashboard for visualizing the overall B2B CX, the impact of each department, and each department's activities on CX.

Additionally, this project seeks to increase interdepartmental visibility, connecting all teams and fostering collaboration. It also aims to help manage customer expectations, decreasing the gap between the expected and perceived quality.

1.4 Research methodology

The research methodology used in this study allows for the fulfillment of the project purpose and goals and ensures the high quality and reliability of this study. It starts with the definition of the problem and the setting of the project's goals.

Then, a thorough examination of the relevant studies on CX and CI is conducted. At the same time, attending training sessions with all BA departments, job shadowing, and conducting interviews with employees from different departments and geographical divisions is crucial to understanding the bigger picture. All of this helps in the clarification of the impact of this project on both the company and the scientific community, resulting in the redefinition of the already established objectives. Following this, and as a consequence of it, CJ is mapped, the variables (KPIs) to be considered in the study are defined, and the methods to construct the B2B CX CI are selected. The data collection methods are also established at this stage of the project, and as they are participatory methods, the participants are chosen too.

Before starting the collection of data, the selected KPIs are validated by each department. After the approval, questionnaires are developed and, after the departments' validation, the preliminary phase starts which is followed by the meetings with department managers and, finally, the meeting with the sales directors. Meanwhile, each department gathers the required data to calculate the KPIs.

Once the data collection phase is over, the calculation of KPIs' weights and utility functions begins; the KPIs' obtention, and conversion into utilities. Then, the departments' utility functions are determined and, finally, the CX is computed. Afterward, the results are discussed with each department not only to validate the B2B CX CI but also to understand the results obtained, identify the root causes and points for improvement, and assess the feasibility of some recommendations aimed to improve CX.

Lastly, after having completed all these steps, dynamic dashboards and this report, which detail the most important findings, are created.

1.5 Dissertation structure

The present dissertation is structured in 5 chapters. Chapter 1 includes a brief presentation of the company, a description of the project framework (context, purpose, and goals), and the

methodology followed throughout the project.

Chapter 2 encompasses an in-depth analysis of the literature related to the main topics of this dissertation, CX and CIs. Regarding CX, there is a focus on mapping, measurement, monitoring, and management tools. And, concerning CIs there is a particular emphasis on weighting and aggregation methods.

Chapter 3 presents the methodology adopted in this project. Therefore, a description of the processes followed during the preparation, collection, and analysis of information phases is made. It also discusses the visualization techniques applied in this project.

Chapter 4, the case study, describes the implementation of the project methodology. It includes a brief overview of the glass packaging industry, the company, and the project; a description of the project actors (the analyst, stakeholders, and decision-makers); the examination of results and respective discussion, and the presentation of dynamic dashboards.

Finally, Chapter 5 highlights the main contributions, to both the company and the scientific community, and the results of this work, together with its major limitations. It also presents recommendations for future works.

Chapter 2

Literature review

This chapter presents the most relevant theories, methods, and gaps in the existing research regarding the main topics of this dissertation, CX and CIs.

It is divided into five subchapters. It begins with an overview of CX (Section 2.1), and proceeds with the presentation of CX design frameworks (Section 2.2) and with the analysis of the most commonly used CX measurement techniques (Section 2.3). Then, Section 2.4 outlines the use of CI to estimate the CX. It starts with a brief definition of CIs, which is followed by the presentation and explanation of several weighting and aggregation methods. Lastly, the current gaps in the B2B CX knowledge are highlighted (Section 2.5).

2.1 An overview of customer experience

Marketing theory and practice experienced a major evolution through a series of transformations from products to services and CXs (Maklan & Klaus, 2011). The first shift - the change from product brand to service-based relationship marketing - was highly motivated by the changes in the business environment. In particular, the intensifying competition, growing globalization of markets, and more demanding clients (Parasuraman, 1998). In this new reality, the product's quality was no longer enough to remain competitive, being fundamental to nurturing a good relationship with clients (Maklan & Klaus, 2011). Additionally, companies started constructing product-service systems, combining products and services to fulfill most of their customers' needs (Baines et al., 2007). More recently, firms started focusing on the customers' value in use and their experience, with products and services being seen as a whole rather than separate parts (Maklan & Klaus, 2011).

Creating a compelling CX has become a topic of particular interest for business executives (Keyser et al., 2020), and a top priority for marketing researchers (Lemon & Verhoef, 2016). And the positive impacts of CX, reported in several studies, such as the increase of CS (Liljander & Strandvik, 1997; Lemon & Verhoef, 2016; McColl-Kennedy et al., 2015), customer loyalty (Yu & Dean, 2001; Pullman & Gross, 2004; Mascarenhas, Kesavan, & Bernacchi, 2006) and confidence (Flanagan, Johnston, & Talbot, 2005) boosted the interest even further, which resulted

in a drastic increase in academic publications (Keyser et al., 2020). Consequently, diverse and distinct conceptualizations of CX were created, which raised concerns about the generalization and replication of some insights, as well as on the measurement tools and their validity (Becker & Jaakkola, 2020).

In order to overcome these challenges, Becker and Jaakkola (2020) developed an integrative view of CX, based on the most relevant CX research. They defined CX as "non-deliberate, spontaneous responses and reactions to offering-related stimuli embedded within a specific context". In other words, CXs arise from firms' multiple direct and indirect contacts with customers along the CJ (Meyer & Schwager, 2007). In these moments, organizations stimulate cognitive, emotional, physical, sensorial, spiritual, and social responses (Schmitt, 1999; Verhoef et al., 2009; Keyser et al., 2015). These are subjective and context-specific because customers' evaluative outcomes depend on individual, situational, and sociocultural context (Keyser et al., 2020). Moreover, CX is dynamic, evolving over time throughout the multiple contact points of the CJ (Verhoef et al., 2009; Lemon & Verhoef, 2016).

This broader view of CX highlighted the fact that companies cannot create it. They can only design, monitor, and manage interactions that impact such experiences. And to succeed, companies must first and foremost understand their customers and the differences between them, and clearly identify their needs. Over the years, tools and methods for managing customers have evolved. From *mass marketing*, in which firms try to satisfy all customers with a single offer; to *niche marketing*, in which the offers are tailored to specific and well-defined segments of the market, and, more recently, to *one-to-one marketing*, in which the marketing mix is tailored to the individual customer (Palmatier & Crecelius, 2019).

In B2B contexts, the most common approach for managing customers and, in particular, the differences between them, is through *segmentation* (niche marketing). As defined by Mitchell and Wilson (1998), it is "an ongoing and iterative process of examining and grouping potential and current customers whose needs are within-group similar but between-group different". It is considered the "core of good industrial marketing" (Bonoma & Shapiro, 1983) and it is seen as one of the most critical and difficult challenges to accomplish (Palmer & Millier, 2004; Boejgaard & Ellegaard, 2010).

In fact, effective B2B segmentation can have multiple benefits, such as developing/improving relationships with the most valuable customer segments, adding value to products and services through tailoring propositions, and creating barriers to competition (Dibb & Simkin, 2008). Additionally, Cortez et al. (2021) point out that segmentation can help the company find a balanced marketing mix, which is neither overgeneralized nor over-customized. Freytag and Clarke (2001) show it allows a smarter allocation of the firm resources, and LaPlaca (1997) demonstrates its importance regarding the development and evaluation of new approaches with respect to products and markets.

In addition to a thorough understanding of the company's customers, companies must select the best techniques to design and measure their CX, so it can be continuously improved. Over the years, academics and practitioners have developed an extensive toolbox. The most consensual

design tools are presented next, in Section 2.2, and the measurement techniques appear later, in Section 2.3.

2.2 Customer experience design frameworks

Starting with the CX mapping tools, in industrial markets, *service blueprinting* (Bitner et al., 2008; Patrício et al., 2008) and *CJ mapping* (Rawson et al., 2013) are amongst the most popular and used ones. They help companies to visualize and structure all the interactions with customers (Bitner et al., 2008).

2.2.1 Service blueprint

A service blueprint is a diagram developed to visually clarify the steps involved in a service delivery process, from back-office activities to front-facing customer interactions (see the example below, Figure 2.1) (Halvorsrud, Kvale, & Følstad, 2016; Lemon & Verhoef, 2016).

The key components of a service blueprint are *customer actions*, *frontstage actions* (i.e. activities over which customers have visibility), *backstage actions* (i.e. tasks that impact frontstage actions but over which the customer has no visibility), and *support processes*. In the diagram, a division between these components is made using three lines. The *line of interaction* aims to highlight where the direct contacts happen between customer-firm, the *line of visibility* separates the tasks that are visible and invisible to clients, and the *line of internal interaction* separates the employees who contact customers from the ones that do not have direct interactions with them.

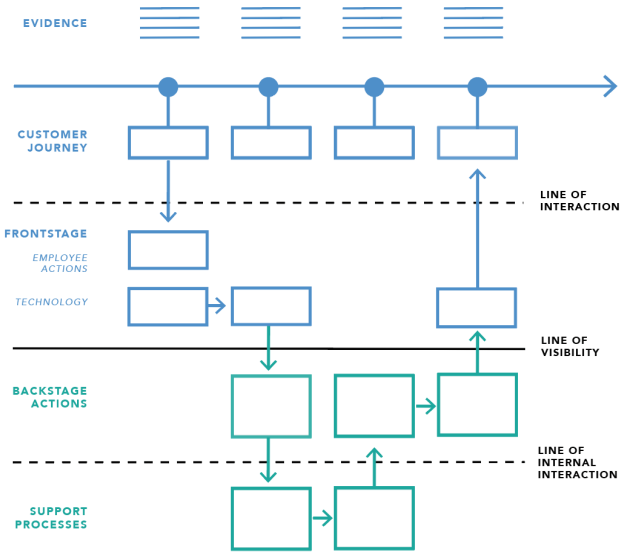


Figure 2.1: Service blueprint example. Source: Gibbons (2017)

The biggest criticism regarding service blueprinting is that it is not sufficiently customer-focused. Thus, it should only be used as a starting point for CJ mapping, which requires additional customer input (Lemon & Verhoef, 2016).

2.2.2 Customer journey map

The CJ can be described as a sequence of events the customer goes through that form CX. Mapping and analyzing it helps firms to detect weaknesses, monitor pain points, find opportunities for optimization, understand how to prioritize actions to enhance CX and how cross-departmental efforts should be bridged, and comprehend the time evolution of the customers' objectives, expectations, and behaviors (Lemon & Verhoef, 2016; Tueanrat et al., 2021). An example of a CJ map template can be found in Figure A.1, Appendix A.1.

2.2.2.1 Stages

The CJ is usually conceptualized in three main stages, *pre-purchase, purchase*, and *post-purchase* (Lemon & Verhoef, 2016).

1. **Prepurchase** - This stage includes all the interactions a customer has with the company, prior to the purchase transaction (Lemon & Verhoef, 2016). Thus, it starts when the customer recognizes the need and, ends when it is satisfied. Between these moments, the customer searches for, considers, and evaluates the alternatives (Puccinelli et al., 2009).

2. **Purchase** - The purchase stage encompasses all customer interactions with touchpoints during the purchase moment. Hence, it includes choice, order placement, payment, pickup, and delivery (Lemon & Verhoef, 2016). It is usually the stage with the shortest duration.

3. **Postpurchase** - Finally, this stage covers all the experiences customers go through after the purchase. It encompasses customers' behaviors such as usage and consumption of the products or services, postpurchase engagement, and service requests (Lemon & Verhoef, 2016). Researchers believe this stage is relevant in customer engagement or disengagement, which explains the number of studies focused on service recovery, product returns, repurchases, customer loyalty, and Word-of-Mouth (WoM) (Lemon & Verhoef, 2016; Keyser et al., 2020).

2.2.2.2 Touchpoints

CJs are described as a succession of touchpoints across different stages (Lemon & Verhoef, 2016). They are critical to CX since without them there are no experiences (Hoffman & Novak, 2018). Touchpoints are defined as the moments where customers contact/interact with the organization, which leads to specific experiences. They vary in terms of *control* and *nature*, as demonstrated in Figure 2.2.

Control refers to the fact it can be either the firm controlling the interaction (firm-controlled); or the client, a partner, or other external entity (non-firm-controlled). And *nature* concerns the way the brand/firm is represented in the touchpoint, which can be human (e.g., frontline employees), digital (e.g., website, firm's portal), physical (e.g., store), or a combination thereof.

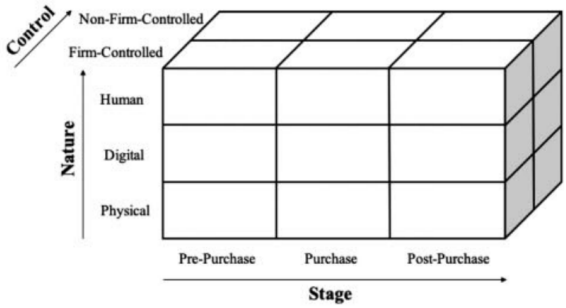


Figure 2.2: Touchpoints. Source: Keyser et al. (2020)

It is also important to highlight that touchpoints have different degrees of importance, i.e. they are not equal in value. Thus, firms should assess the importance of each touchpoint and focus on the most impactful ones (Witell et al., 2020).

2.2.2.3 Participants

Besides touchpoints, CJ also depends on the participants in each touchpoint (Lundin & Kindström, 2023). In B2B settings, the multiplicity of interactions amongst different actors, from different functional and organizational units, with different roles and in different hierarchical levels makes CJ highly complex (Witell et al., 2020; Lemon & Verhoef, 2016).

Regarding B2B CJs, participants can be divided into three categories: *suppliers*, *customers*, and *third parties*, who are part of the larger business ecosystem (e.g. consultants, competing suppliers, outsourced carriers) (Zolkiewski et al., 2017; Witell et al., 2020; Lundin & Kindström, 2023). They may take a more *active or passive role* (Lemon & Verhoef, 2016; McColl-Kennedy et al., 2019; Witell et al., 2020) and this can vary across the CJ. Additionally, interactions may occur among *individuals* (operating at different hierarchical levels) or *collectives* (such as functional units), or a *combination of them* (e.g. the sales manager of the supplier firm negotiates prices with the client's procurement department) (Witell et al., 2020).

Another important aspect is to define and monitor relationship control. It must take into account the suppliers' strategy, whether it is to increase efficiency through standardization or to improve CX through customization and, what are the customers' expectations. B2B relationships are usually long-term, involving an ongoing process of exchanges, collaboration, and cooperation between supplier and customer (Witell et al., 2020; Lundin & Kindström, 2023).

2.2.2.4 Service failures and the service recovery journey

Service failures, i.e. situations in which the firm is not able to perform the service as promised, can occur in every stage of the CJ and may trigger different customer reactions (Vaerenbergh et al., 2019). In B2B environments, they are particularly critical since they not only affect the customers' operations but also spread to their clients via "domino effects" (Zhu & Zolkiewski, 2015). Service failures usually result in customer dissatisfaction and negative WoM (T. C. Johnston & Hewa,

1997) and, in case of repeated failures, the likelihood of customers switching service providers is increased (Spreng et al., 2009).

Thus, it is imperative for companies to develop organizational responses to these situations. These responses can be integrated into the Service Recovery Journey (SRJ), which is the journey that encompasses the events that happen during the recovery process. SRJ is divided into three phases: *pre-recovery*, *recovery*, and *post-recovery* (Vaerenbergh et al., 2019). A framework of the SRJ with examples of strategies the company may implement is presented in Figure A.2, Appendix A.1.

Regarding this topic, researchers are also discussing the *service recovery paradox* (See Figure A.3, in Appendix A.1, that illustrates this phenomenon). It is believed that CS in B2B relationships can exceed pre-failure levels when there is a successful recovery. However, there is no clear evidence that this phenomenon happens always in all types of service failures and contexts (Michel & Meuter, 2008).

2.2.2.5 Customer's zone of tolerance

When managing CJ and SRJ, it is recommended the analysis of the customer's zone of tolerance (ZOT). This concept was first introduced by Parasuraman et al. (1991), and it is "the gap between two levels of customer expectations" - the adequate (the minimum service the customer considers acceptable) and the desired. Providing a service below the adequate level will result in customer dissatisfaction. Within the ZOT the customers will be satisfied, and above the desired level, they might even be delighted. See the schematic customer's ZOT diagram in Figure 2.3.

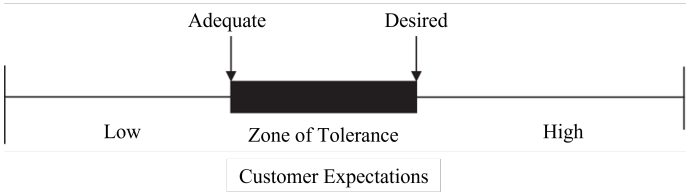


Figure 2.3: Diagram of the customer's zone of tolerance. Adapted from: Parasuraman et al. (1991)

It is important to highlight that each customer has their own ZOT which may change over time. Aspects like price, competition, or specific service attributes might expand or contract customers' ZOT (Zeithaml, Bitner, & Gremler, 2006), and service failures usually result in a narrower ZOT (R. Johnston, 1995).

2.3 Customer experience measurement

In order to foster a superior CX, it is crucial to measure it first. "What gets measured, gets managed", as Peter Drucker once said. The development of a measurement tool that captures accurately the complexities of CX, across numerous touchpoints, along the CJ is, nowadays, one of the biggest challenges facing researchers (Lemon & Verhoef, 2016; Zolkiewski et al., 2017).

Despite the vast interest, importance, and number of publications on the topic, there is a lack of specific measurement tools for CX. A recent literature review conducted by Patti et al. (2020) suggests that a complete understanding of CX may only be obtained using three categories of metrics: *perception-based*, *operational*, and *outcomes-based*.

2.3.1 Perception-based metrics

These metrics concern the customers' feelings, thoughts, and opinions toward the organization, being usually focused on behaviors such as loyalty, trust, forgiveness, satisfaction, advocacy, and effort (Patti et al., 2020). The most common perception-based metrics are *CS* and *willingness to recommend* (Meyer & Schwager, 2007; Homburg & Rudolph, 2001; Reichheld, 2003), *service quality* (Parasuraman et al., 1985; Szmigin, 1993), *service encounter quality* (Jayawardhena et al., 2007), *relationship quality* (Naudé & Buttle, 2000; Ulaga & Eggert, 2006) and *service experience* (Maklan & Klaus, 2011). These are further explained below.

– **Net Promoter Score (NPS)** - Reichheld (2003) created NPS and described it as "the one number companies need to grow" arguing that knowing the customers' willingness to recommend the company's products or services is the best proxy and route to firm growth. Customers rate the likelihood of recommending between 0 (not at all likely) and 10 (extremely likely), answering a simple question "How likely are you to recommend our company/products/services to a colleague?". NPS is then determined by subtracting the percentage of customers who are detractors (i.e. people who would not recommend; rate from 0-6) from the percentage who are promoters (i.e. people who would recommend; rate from 9-10).

  Despite being one of the most widely used metrics because of its simplicity, it is highly criticized. Firstly, Rossiter (2017) argues that having the intention to recommend is very different from actually recommending. Secondly, he also points out that individuals may also give negative recommendations and NPS does not measure that, since detractors only consider people that would not recommend, which is different from a bad recommendation. Thirdly, there is no empirical evidence so far, that proves NPS to be a better predictor of growth in comparison with other satisfaction metrics (Keiningham et al., 2007, 2017).

– **Customer Satisfaction in Industrial Markets (INDSAT)** - Even though CS plays an important role in industrial markets, it has received little attention so far. INDSAT is one of the very few methods which assesses it. The seven dimensions of this satisfaction scale are (i) satisfaction with products; (ii) satisfaction with salespeople; (iii) satisfaction with product-related information; (iv) satisfaction with order handling; (v) satisfaction with technical services; (vi) satisfaction with interaction with internal staff; (vii) satisfaction with complaint handling (Homburg & Rudolph, 2001). Its biggest problem is its low practicability (29-item scale).

– **Service Quality (SERVQUAL)** - SERVQUAL is conceptualized as "the customers' overall judgment or attitude about the quality of the service" (Parasuraman et al., 1985). SERVQUAL

was created to measure the discrepancy between customers' expectations and perceptions regarding the service offered by the firm. It is a 22 item-scale divided into five dimensions: (i) tangibles: the appearance of physical, equipment, personnel, and communication materials; (ii) reliability: the ability to perform as promised, (iii) assurance: employees' competence and knowledge and their ability to inspire trust and confidence, (iv) responsiveness: the willingness to help customers and provide prompt service, and (v) empathy: treating customers with care (Parasuraman et al., 1985). The biggest criticism of SERVQUAL is that its dimensions are difficult to replicate in different contexts (Cronin. Jr & Taylor, 1992).

- **Service Performance (SERVPERF)** - SERVPERF is a variant of the SERVQUAL scale, that measures quality as an attitude, not satisfaction. Similarly to SERVQUAL, the model consists of five dimensions: (i) tangibles, (ii) reliability, (iii) assurance, (iv) responsiveness, and (v) empathy, but it has not only one but two sets of 22 item statements for the importance and perception sections (Cronin. Jr & Taylor, 1992). Because of being so focused on performance, it fails to measure accurately CS and CX.

- **Service Experience (EXQ)** - EXQ is one of the very few methods that assess CX. It is a multi-item scale, composed of 19 items divided into four dimensions: (i) product experience, (ii) outcome focus, (iii) moments-of-truth, and (iv) peace-of-mind (Maklan & Klaus, 2011; Klaus, 2015; Kuppelwieser & Klaus, 2021).

  Although EXQ has been proven to be valid in many contexts (Klaus, 2015), and the insights obtained are highly valuable, it has low practicability and requires more resources and investment in terms of data collection (Kuppelwieser & Klaus, 2021).

2.3.2 Operational metrics

Operational metrics measure business performance, being a great complement to perception-based metrics. The most widely used are time-based, such as *average wait time, response time, average resolution time*, because customers consider time to be a valuable resource. Other operational metrics focus on effectiveness and efficiency, such as *first contact resolution* (i.e. the total number of incidents that were solved on first contact in comparison to the total number of incidents), *service level*, etc. (Patti et al., 2020).

2.3.3 Outcomes-based metrics

According to Zolkiewski et al. (2017), measuring CX in B2B settings requires a more strategic approach. Rather than thinking and focusing on CS and loyalty, managers need to learn more about the components of the CJs and how they impact the tactical and strategic goals of their clients. Some examples of outcomes-based metrics include the *number of complaints*, *number of incoming calls*, *market share*, *customer churn rate*, *retention rate*, etc.

2.3.4 Selected categories of metrics and customer journey stages

Considering the project purpose and main goals, and taking into account the data available and project restrictions (namely the time available and the impossibility of contacting customers), the metrics selected fall into either *operational* or *outcomes-based* category.

Regarding the CJ stages, this study will only encompass the *purchase* and *postpurchase* stages. The ideal would be to include the three, but since there is no data collected during the pre-purchase stage, it was not possible.

2.4 Customer experience monitoring using composite indicators

As demonstrated in previous sections, measuring B2B CX is highly complex, and in this specific setting, the few metrics available do not gather consensus from the scientific community or practitioners. Therefore, developing and analyzing new approaches are of the utmost interest. It is in this context that the idea of creating a CI to measure B2B CX arises. Firstly, a CI can help capture the multiple dimensions the CX has, and provide a comprehensive assessment of it. Secondly, during the CI construction, different weights can be assigned to indicators or dimensions, based on their relative importance, which ensures that indicators that have a greater impact on overall CX are given more significance in the composite measure. Thirdly, by condensing information, the CI facilitates clear and concise communication of results to stakeholders, making it easier to communicate and align efforts across teams and departments, and also making it easier to track changes over time.

This section aims to introduce the CI concept and present the procedure and the weighting and aggregation methods to construct it.

The term indicator refers to data that summarizes a system's characteristics or highlights its current state (Saisana & Tarantola, 2002). A CI is the "mathematical combination (or aggregation) of a group of individual indicators" (Nardo et al., 2008). CIs are often used when a single indicator cannot capture all the related issues, such as when the goal is to measure multi-dimensional concepts (e.g. competitiveness, industrialization, sustainability, public health, technological development).

CIs are increasingly recognized as a useful tool (Gibari et al., 2019) and their use is growing around the world (Nardo et al., 2008), mainly due to their potential. CIs allow the summarization of complex or multi-dimensional issues, facilitate the comparison of complex dimensions, and the interpretation of indicators, facilitate communication, promote accountability, and assess progress over time on complex issues (Saisana & Tarantola, 2002).

Even though the list of advantages of using CI is extensive, this tool has also received some criticism. Authors point out CIs can easily send misleading messages, if poorly constructed or misinterpreted; they may invite drawing simplistic conclusions; and they may mask significant flaws in some dimensions of the phenomenon, making it more challenging to determine the best course of action (Nardo et al., 2008).

The construction of CIs has been approached from a multitude of viewpoints, in fact, the literature offers a variety of methodological approaches (Gibari et al., 2019). Following the European Commission guidelines, the first steps concern the definition of the theoretical framework and the data selection, analysis, and treatment, which includes the imputation of missing data, multivariate analysis, and normalization. Then, indicators should be weighted and aggregated in line with the underlying theoretical framework. Afterward, the statistical coherence and robustness of the CI should be assessed. Lastly, it is crucial to select the visualization technique that shows the CI results in a clear and accurate manner that helps or enhances interpretability (Nardo et al., 2008).

Since it is imperative to meaningfully combine various dimensions measured on different scales when creating a CI, both the weighting model and the aggregation procedures are considered critical decisions (Nardo et al., 2008). According to Greco et al. (2019), these are the steps of CI construction where more criticism appears but also where a promising future lies. Thus, gaining a better understanding of each weighting and aggregation method, namely their benefits and drawbacks, is crucial for CI development.

2.4.1 Weighting methods

Weighting consists of attributing a weight (i.e. a kind of coefficient) to every criterion that composes the CI. The weight should reflect the relevance of each criterion in comparison to the others.

In order to obtain the variables' weight there are multiple techniques. The simplest is Equal Weighting (EW), which considers all variables to have the same weight. It is used in many applications because of its simplicity, alleged objectivity, and the fact it does not require agreement between decision-makers (DMs). Nevertheless, many authors argue that it has many flaws, namely being oversimplistic, and not conveying a realistic image of the issue in analysis (Greco et al., 2019).

The alternatives are usually divided into two groups: the ones based on *statistical models*, such as Principal Component Analysis (PCA), Factor Analysis (FA), Data Envelopment Analysis (DEA), and Benefit of the Doubt (BOD); and the ones based on *participation*, such as Budget Allocation Process (BAP), Analytic Hierarchy (AHP), Conjoint Analysis (CA) (Nardo et al., 2008), and the DCM-SRF. Table A.1 in Appendix A.2 shows some examples of CIs created with these weighting methods.

In the literature, there is no consensus on the best approach. On the one hand, researchers agree that statistical-based methods are more objective than participatory. While in the statistical ones, the weights emerge from data, thus, are immune to manipulation (Ray, 2008); in the participatory, a group of DMs evaluates the indicators' importance/impact based on their perceptions, which makes decisions subjective (Nardo et al., 2008). However, it is believed that by making the subjective judgment as transparent as possible, the likelihood of manipulation decreases and, consequently, results are better. On the other hand, Saisana and Tarantola (2002) showed that statistical relationships between variables do not always reflect their actual influence on one another.

Thus, it is required great attention when determining the significance of an indicator or dimension based on what the data "consider" to be true (Decancq & Lugo, 2013).

Further criticism appearing in the literature is focused on the specific properties of the methods within these groups, which are examined individually below.

2.4.1.1 Statistical-based methods

Principal Component Analysis (PCA). PCA is a multivariate technique for analyzing quantitative data and it was first proposed by Karl Pearson (1901). The aim of this method is to capture as much variance in the original variables as possible with the fewest number of components so that the dimensionality of the data set is reduced, which simplifies its analysis (Pearson, 1901). PCA has a few requirements that must be met in order to obtain meaningful results. The data set should be large, and free of outliers, and variables should be continuous and linearly correlated (Greco et al., 2019).

In addition to being objective and not manipulable, PCA has been proven to be a good tool to solve the double-counting problem. However, it should be used with caution as it is very sensitive to the construction of data, i.e. the weights are likely to change if data revisions and/or updates are done (Nicoletti et al., 2000). It is also very sensitive to the presence of outliers and small data samples (Nardo et al., 2008).

Data Envelopment Analysis (DEA) and Benefit of the Doubt (BOD). DEA is a popular non-parametric approach that uses linear programming to estimate/evaluate the efficiency of a set of comparable units (Charnes et al., 1978). Firstly, an efficiency frontier is determined either through benchmarking or by a hypothetical DM (Korhonen et al., 2001). Then, the weights are calculated by dividing two distances: the distance between the origin and the actual observed point and between the origin and the projected point in the frontier.

BOD is the name of the adjusted DEA formulation, most commonly used in the CIs field. In BOD the weights are obtained by dividing the unit's actual performance by its benchmark or target performance (Nardo et al., 2008).

Among the advantages of these methods is their objectivity, their ability to encompass a multiplicity of inputs and outputs and the incentive they give to the units lagging behind (Nardo et al., 2008). Regarding the disadvantages, it is difficult to incorporate preference information in these methods (Pereira et al., 2020) and they are sensitive to the selection of inputs and outputs. Additionally, the best performer will not see its progress reflected in the CI, which does not encourage continuous improvement nor the *status-quo* enhancement (Nardo et al., 2008).

2.4.1.2 Participation-based methods

Budget Allocation Process (BAP). BAP consists of selecting a set of DMs, giving them a "budget" and asking them to distribute it to the indicators. The allocation is based on their knowledge

and subjective assessment of the relative importance of the various indicators. Weights are then calculated as average budgets.

This process is relatively straightforward, transparent, and quick (Nardo et al., 2008). Nonetheless, it should only be used for a small set of indicators (ideally, fewer than 10), since inconsistency problems may arise with more indicators and a large number of indicators can cause serious cognitive stress (Saisana & Tarantola, 2002; Nardo et al., 2008).

Analytic Hierarchy Process (AHP). AHP is a multi-criteria decision method that was introduced by Saaty (1980). The complex problem under analysis is decomposed into a hierarchy consisting of three levels: the ultimate goal, the criteria, and the alternatives. DMs are asked to do pairwise comparisons among criteria, having to decide how much more important one criterion is compared to another. For that, a 9-point scale is used, that ranges from 1 -'equally important' to 9 - 'much more important'. The weights reflect the trade-off between different indicators. During their evaluation, an inconsistency ratio is calculated, to measure the degree of inconsistency of DMs judgments (Saaty, 1980). Thus, AHP is less prone to inconsistency problems than BAP. But like in BAP, when there is a large number of indicators, it exerts cognitive stress on DMs. Additionally, another disadvantage is the computational effort required due to the high number of pairwise comparisons (Nardo et al., 2008; Greco et al., 2019).

Conjoint Analysis (CA). CA was developed by Green and Rao (1971), and it is widely used in the marketing field. It is usually seen as the opposite of the AHP as it adopts a top-down approach. First, it seeks the evaluation of a set of alternative scenarios and then decomposes them according to the individual indicators. Some of CA's major drawbacks are its overall complexity and the requirements of a large sample (Nardo et al., 2008).

Deck Cards Method Simos-Roy-Figueira (DCM-SRF). This technique was first proposed by Simos (1989) and evolved over the years. In 2002, J. Figueira and Roy (2002) suggested a different approach for data collection and data processing and also a new method to normalize the weights minimizing the rounding off errors. And, in 2021, Corrente et al. (2021) presented a new way to improve the reliability of the values elicited with the methodology and provided procedures to handle inconsistency or incompleteness, or imprecise information supplied by the DM.

The DCM-SRF has been employed in a variety of contexts (for a list of applications see Siskos and Tsotsolas (2015)), has received strong support from DMs and it is believed that the data produced is significant from the DM's preference point of view (Corrente et al., 2021; Pereira et al., 2020). Even though it is intuitive and simple both in data collection and computation, some authors argue it has some robustness issues (J. Figueira & Roy, 2002).

The first step of this method is information collection. It consists in giving DMs a set of 'playing cards', each one associated with a criterion, and asking them to rank the cards in ascending order, from the least important to the most important. In cases of equal importance (i.e., the same weight), DMs should create subsets holding the cards together. After this, and since there may

be smaller or bigger differences between the different levels of importance, DMs are given white cards to introduce between two successive cards or subsets. No cards mean the difference between weights is minimal, one card means the difference is twice the minimum, etc. The greater the difference, the greater the number of white cards. After this, DMs must define how many times the weight of the criterion ranking first is higher than the weight of the criterion ranking in the lowest position. Once all these steps are complete, weights can be calculated. More details regarding this procedure are presented in Chapter 3.

2.4.2 Aggregation methods

The step following the weighting is aggregation. There are numerous methods suggested in the literature, as well as categorization suggestions that aim to help the analyst choose the best method. For example, Nardo et al. (2008), suggest three distinct categories: *linear*, *geometric*, and *multi-criteria*. Whereas Munda (2005) recommends a division into two groups, the *'compensatory'* and the *'non-compensatory'*. The latter is supported by the idea that weights can be interpreted in two different ways: as 'trade-offs', i.e. it is possible to offset a disadvantage on some criteria by a sufficiently large advantage on another criterion; or as 'importance coefficients' (Nardo et al., 2008). According to this categorization, the *linear* and *geometric methods* are included in the *'compensatory'* approaches category, while other *multi-criteria decision making* (MCDM) methods are included in the *'non-compensatory'* (Nardo et al., 2008). However, it must be highlighted MCDM methods can be designed to be either compensatory or non-compensatory based on the research objectives, the characteristics of the variables being evaluated, and the preferences of the decision-makers.

2.4.2.1 Compensatory aggregation

Starting with the compensatory methods, the two most widely used are *linear aggregation*, also known as Simple Additive Weighting (SAW), which is the simplest method; and the *geometric aggregation*, also known as Weighted Product (WP). In these methods, the global performance of an alternative is calculated as the weighted sum (in SAW) or product (in WP), of its values for each criterion. This entails a total compensation among the various criteria for the SAW method and a partial compensation for the WP method. This compensability is one the most serious criticisms these methods raise when considering their application on CI construction (Nardo et al., 2008).

2.4.2.2 Non-compensatory aggregation

Regarding the MCDM methods used in the CI construction, they usually fall into one of the following categories: the *value and utility-based methods*, the *outranking methods*, or the *distance functions-based methods* (Gibari et al., 2019). See examples in Table A.2, Appendix A.2.

Value and utility-based methods. The aim of these methods is to create a means of associating a real number with each alternative and to produce a preference order for the alternatives, based on

DM's value judgments (Belton & Stewart, 2002; Azapagic & Perdan, 2005). Within this group, the two most used methods are Multi-Attribute Utility Theory (MAUT) and Multi-Attribute Value Theory (MAVT), being the work of Keeney and Raiffa (1976) a key reference.

– MAUT involves the determination of partial utility functions to calculate a global utility function U, while MAVT involves the determination of partial value functions to calculate a global value function V. In both cases, additive and multiplicative models are used to obtain the global function (Keeney & Raiffa, 1976; Belton & Stewart, 2002; Azapagic & Perdan, 2005; Jansen, 2011). It should be noted that whereas MAUT can be used both for situations of certainty and uncertainty, MAVT cannot take uncertainty into account (Azapagic & Perdan, 2005).

In addition to these methods, the application of the Utility Theory Additive (UTA), the Simple Multi-Attribute Rating Technique (SMART), and the Measuring Attractiveness by a Categorical Based Evaluation Technique (MACBETH) has been growing.

Outranking methods. Similarly to the value- and utility-based approaches, they build a preference relation among alternatives evaluated on several attributes. These methods, developed in France, are based on pairwise comparisons between alternatives. They usually involve two phases: firstly, given two options, the DM decides whether "alternative a is at least as good as alternative b" and, secondly, it is determined how all the pairwise outranking assessments can be combined.

Within this family, the Elimination and Choice Expressing Reality (ELECTRE) and Preference Ranking Organization Method for Enrichment Evaluations (PROMETHEE) are the most used methods (Gibari et al., 2019).

– The ELECTRE method was developed by B. Roy (1968) and it selects an alternative as the best action(s) from a given set of actions. It is mainly applied in choosing, ranking, and sorting problems. ELECTRE has evolved over time and today ELECTRE's family includes ELECTRE I, ELECTRE IS, ELECTRE II, ELECTRE III, ELECTRE IV, and ELECTRE TRI, etc. (B. Roy, 1991).

– The PROMETHEE method was introduced by Brans and Vincke (1985) and it performs a pairwise comparison of alternatives to rank them with respect to some criteria. Similarly to ELECTRE, PROMETHEE's family grew over time including now PROMETHEE I, PROMETHEE II, PROMETHEE III, PROMETHEE IV, PROMETHEE V, and PROMETHEE VI (Brans et al., 2005).

Outranking methods have gained popularity due to their non-compensatory nature and their simplicity in incorporating uncertainty into the evaluation of differences between alternatives. However, some authors warn that these methods lack theoretical backing (Azapagic & Perdan, 2005).

Distance functions-based methods. According to Diaz-Balteiro et al. (2017), the idea behind the distance functions-based methods is that instead of maximizing a function of the preferences of the DM, one should minimize the distance between an alternative and a point or points enjoying 'good preferential properties'. These methods, when used to construct CIs, in addition to the assessment of weights, require an evaluation of the corresponding reference levels.

Examples of distance functions-based methods are the goal programming method, compromise programming method, the Technique for Order Preferences by Similarity to Ideal Solutions (TOPSIS), Grey Relational Analysis (GRA), etc. (Gibari et al., 2019).

It is important to highlight that the methods described above are the most common in the literature, but there are many CIs that were constructed with different aggregation approaches, using for example, *DEA based methods*, *Mazziotta-Pareto Index (MPI), Penalty for a Bottleneck, Mean-Min Function, Directional BOD, etc.* (Greco et al., 2019). In addition to these examples, recent studies have shown the potential of the Choquet integral-based approach especially to model interdependency relationships between criteria. In fact, although in most cases an independence assumption is made, criteria are not always independent, and sometimes, the combination of two criteria may generate a synergistic or redundant effect as a whole, which should be taken into account (Pereira et al., 2020).

2.4.3 Data collection methods

As shown, some weighting and aggregation methods require the participation of DMs (such as AHP, DCM-SRF, MAUT, etc.). And, collecting DMs preferences can be very challenging, especially when they are numerous. For these situations, the most recommended methods are:

- The *Delphi method*, which consists of sending questionnaires to a selected DM panel until a consensus is reached. The analyst must create a questionnaire and send it to all DMs. Then, he/she must analyze the answers, and if there is no agreement, proceed to the next round, i.e. create and send an updated questionnaire (Lin et al., 2020). This process will only end when a consensus is reached. This technique is very easy to implement and it has little room for errors or biases. Nevertheless, it is a slow process since it relies on DMs responses, and the number of rounds may be high, which besides being time-consuming, can be tiring;

- *Focus groups*, that are a way to collect data through DMs interactions. The analyst must carefully select a small group of specialists to discuss one topic and draw conclusions. This approach allows the obtention of insightful data, but it is time-consuming and during the discussion, some people can be manipulated or biased by other participants;

- The *nominal group technique* is similar to focus groups. It involves gathering a group of people to discuss and reach a consensus. Though, instead of starting immediately discussing the topic, everyone silently writes down their opinions and ideas. Then these are collected and discussed, and most of the time through voting, they reach a conclusion. With this tool,

everyone is encouraged to contribute and give their opinions. Even so, as with focus groups, manipulation may also happen and time is still a problem.

Another commonly used technique is structured interviews.

2.4.4 Selected methods

As demonstrated in Section 2.4.1 and Section 2.4.2, there are multiple weighting and aggregation approaches, each has both benefits and drawbacks, there is no 'one-size-fits-all' solution, and none is clearly superior in all situations. Thus, when constructing the CI, the analyst must choose the better weighting and aggregating system for that particular purpose.

Starting with the selection of the weighting method, it was decided that the most appropriate option would be a participatory method. First, the company wants to engage its employees and requires their participation. Raising employees' awareness towards CX, increasing intra- and interdepartmental visibility, fostering ownership, and gathering the knowledge and expertise of the different areas, as mentioned in Section 1.3, are some of the goals of this project. Second, data collection and processing within the company is somewhat immature, so conclusions based solely on this data were likely to be inaccurate. Moreover, with the statistical analysis of the data some relationships between variables could be found, that could not reflect their actual influence on one another, as studied by Saisana and Tarantola (2002).

Regarding the participatory methods, the most suitable for this study is the DCM-SRF. B2B CXs are highly complex, and their measurement involves a large number of indicators, thus using AHP or BAP would cause serious cognitive stress to DMs, and using CA, in addition to its complexity, would require a large sample. In turn, the DCM-SRF is a very simple method, convenient for the DM to express his/her preferences, easy to understand and apply, and previous studies have demonstrated the high receptivity of DMs to the use of the method (Corrente et al., 2021; Pereira et al., 2020).

In what concerns the aggregation method, the literature shows that in situations where weights should be interpreted as "importance coefficients", non-compensatory aggregation procedures are ideal and strongly suggested in the CI construction (Munda, 2005). But, although outranking-based methods are powerful techniques for ordinal classification and are usually good tools for building a state CI (ordinal scale), the need to analyze the temporal evolution of the CX requires the construction of a cardinal scale (J. R. Figueira et al., 2023). And despite being possible to adapt outranking-based methods to produce such a cardinal scale (J. R. Figueira, Greco, & Roy, 2022), it is a complex process, hard to explain. On the contrary, MAUT is simple enough for interaction and communication with the employees, which is particularly relevant since this study's participants are not familiar with MCDM methods. Therefore, taking into account both the purpose of this study and the company's requirement to include the most employees possible, the most suitable is MAUT.

It was also decided to use an additive model to construct the global utility function. The weighted linear additive preference function is the most commonly applied aggregation method

(Von Winterfeldt & Edwards, 1993). With it, the overall evaluation is calculated by multiplying the weight by the attribute utility for each attribute and summing these weighted attribute utilities over all attributes (Payne, Bettman, & Johnson, 1993).

In order to collect the information required, two methods were selected: focus groups and individual interviews. The first aims to gather employees' opinions, in a preliminary stage and, to rank the departments according to their impact, to obtain the overall CX, at the last stage. And, the interviews focus on KPIs ranking, utility function creation, and target setting. The other methods were discarded either because of being time-consuming, and time is a project constraint (the case of the Delphi method), or because of being impossible to combine with the selected weighting method (the nominal group technique does not work well with the DCM-SRF).

2.5 Knowledge gap

Although CX has attracted much attention and the volume of research is enormous (Witell et al., 2020), most publications are focused on the B2C context; very little research is conducted in B2B environments (notable exceptions include McColl-Kennedy et al. (2019); Zolkiewski et al. (2017)). In the particular case of the glass packaging industry, CX studies are nonexistent. In fact, marketing studies within this industry are generally scarce. Therefore, the need to deepen the knowledge of B2B CX is even further enhanced in it.

In addition to this, it is important to understand if the insights generated from B2C research may apply in B2B settings and how they differ (Keyser et al., 2020). Some authors have already pointed out relevant differences, such as the level of complexity of both the offerings and contact points, that involve multiple front-office and back-office actors interacting in various ways (Håkansson et al., 2009), with different objectives depending on their role (Mikolon et al., 2015; S. Roy et al., 2019; Witell et al., 2020). On top of that, whereas a great experience in B2C must be "engaging, robust, compelling and memorable", according to Gilmore and Pine II (2002), positive B2B CX must be "trouble-free and reassuring" (Meyer & Schwager, 2007).

Furthermore, as highlighted previously, most of the CX metrics are proxy measures and have been widely criticized (Zolkiewski et al., 2017). Therefore, a work like this focused on developing an alternative and innovative approach to measuring B2B CX is of profound interest. B2B CX CI is an original metric that implements for the first time, to the author's knowledge, the conceptual framework created by Patti et al. (2020). Unlike prior studies, this adopts a broad view, considering both operational and outcome-based metrics instead of focusing on the customers' perceptions.

Moreover, in spite of being applied in a wide variety of fields such as sustainability, environment, business, energy, tourism, human development, water service, health, logistics, etc. (see Gibari et al. (2019)), in the marketing field, there are not many examples of CIs (Table A.3, Appendix A.2, shows some of the few examples). And, regarding CX, in particular, the use of CI is almost nonexistent. Forrester's CX Index[1] is the only example, and it is more focused on the B2C environment.

[1] See `https://www.forrester.com/research/cx-index/`

Lastly, but also worth mentioning, is the DCM-SRF-based questionnaire. During the preliminary phase, group meetings were supported by a questionnaire that resulted from an adaptation of the DCM-SRF, in which the method was illustrated by an original example and pictures. It is therefore a new and practical approach to the method that can be replicated in other studies, specially in the ones where meetings in-person are difficult to have.

2.6 Summary

Considering the central objective of this dissertation, to measure the B2B CX, this chapter started by examining the most relevant works, and methods in the existing research regarding CX. This has highlighted the knowledge gap concerning CX measurement in B2B environments, i.e. there are few metrics and these are highly criticized. Therefore, it is important to develop a new approach. In line with this, the construction of a B2B CX CI, proved to be a good option, being introduced a handful of methods to construct it, alongside their strengths and weaknesses. Taking into account these, as well as some case study constraints, it has been defined: the stages of the CJ to include in the study - purchase and postpurchase; the metrics' categories - operational and outcomes-based category; and the methods to construct the CI - the DCM-SRF and the MAUT. The forthcoming chapter will explain the methodology of this new approach to obtain the B2B CX CI, combining all the above-mentioned components.

Chapter 3

Methodology

This chapter encompasses an in-depth analysis of the methodology followed in this project which is divided into four subchapters: *preparation*, *collection*, and *analysis of information*, and *presentation and visualization of results*, as shown in Figure B.1, Appendix B.

Firstly, the *preparation of information* subchapter describes the procedure followed to create the conceptual framework of this research. Section 3.1.1 details the strategy that enabled the selection of variables and presents them. Section 3.1.2 examines the preparation required for the implementation of the weighting and aggregating methods. Then, the *collection* subchapter reports how the information was gathered (Section 3.2), and the *analysis* subchapter, in Section 3.3, presents the methodology followed to calculate the KPIs and build the CI. Last but not least, Section 3.4 highlights the importance of data visualization and, explains the dashboards' construction.

3.1 Preparation of information

Bearing in mind the brilliant words of George Fueschel, 'Garbage in, Garbage out', it is appropriate to say the preparation stage is the most critical in CI construction. During this phase, it is crucial to understand and define the phenomenon to be measured, and its sub-components, selecting the dimensions of the overall composite and the individual indicators; and to plan the implementation of the weighting and aggregating methods.

Following the guidelines created jointly by the Organization for Economic Co-operation and Development (OECD) and the European Commission (Nardo et al., 2008), *experts* and *stakeholders* were engaged in this initial phase, which started with:

1. *Training sessions* with all the departments (13, in total), which was crucial to understanding the general *modus operandi* of the company, i.e. how the processes are organized, the participation of each actor in the process, the pain points, etc.;

2. *Job shadowing*, i.e. spending a day in each department observing employees performing their tasks, and interacting with other colleagues, departments, customers, and suppliers (5, in total). It was important to understand specific workflows and activities, the relationships between the different stakeholders, and to detect nonconformities in the processes;

3. *Interviews* with different people within the same department, from different geographies (17, in total), to clarify doubts and specific aspects and, to better understand what was really important and, thus, mandatory to include in this study.

This combination of training sessions, job shadowing, and interviews was essential to developing a clear understanding of the CJ, and structuring CX into various sub-groups, and dimensions, as shown in Figure 3.1. In this case, the sub-groups are 6 BA departments: Order Management Team (OMT), Service Front Office (FOA), Customer Service, Logistics, Transports, and Invoicing. In what concerns dimensions and criteria, a detailed explanation is presented in the next section.

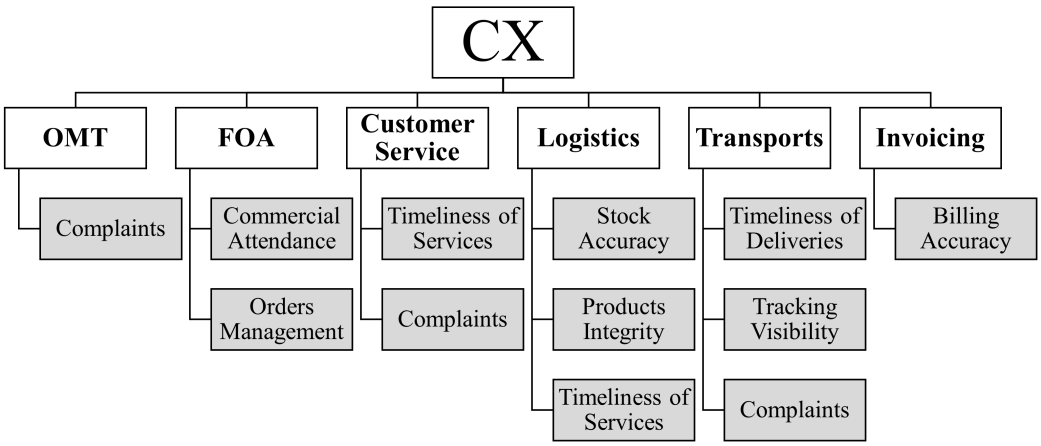


Figure 3.1: Decomposition of CX into departments and dimensions

Note that this preparation was especially important because unlike many existing CIs constructed based on the available literature (e.g. van Calker et al. (2006)), this CI is developed in an unexplored field with few existing works to rely on. Previous studies have, however, provided valuable insights into the identification of highly valued outcomes by customers.

3.1.1 Variable selection

This is a pivotal step in CI construction because its strengths and weaknesses largely depend on the quality of the underlying variables (Nardo et al., 2008). They should be selected according to their importance, analytical soundness, timeliness, and accessibility, among other criteria.

In this project, variable selection involved careful consideration of multiple factors. Firstly, in line with experts' and stakeholders' insights, the most meaningful and impactful indicators were listed[1]. Then, the redundant indicators were removed as well as the ones whose impact could not be directly linked to CX. Lastly, to enable a temporal analysis of CX, only the indicators with available data recordings were considered. It should also be stressed that the selected variables are more focused on the service provided, in alignment with the company's request.

[1]This list is not presented, for the sake of conciseness since over 100 KPIs were analyzed. The company has also restricted the disclosure of some of the indicators for confidentiality reasons.

In the following subsections, the responsibilities of each department are briefly described, along with the selected variables (criteria and indicators) and the respective preference direction. The latter took into account the KPI's impact on the CX. If the KPI positively impacts CX, the preferred direction is maximization, whereas if it negatively impacts CX, the preferred direction is minimization. This information is important for the utility function's construction, as explained next, in Section 3.1.2.2.

3.1.1.1 Order Management Team

The OMT is responsible for receiving, analyzing, and submitting orders in the system, changing/canceling orders due to the client's request/update or internal instructions, managing stock allocation, guaranteeing the issue of mandatory documents for order shipment, etc.

The work of this team has a major impact on the CX, as any errors in order processing, such as incorrect quantities, materials, or delivery dates, frequently lead to customer complaints. In line with this, it was decided to select *complaints* as the dimension to measure the impact of OMT in the CX and, as criteria, the different degrees of severity they may have (see Table 3.1). In what concerns the indicators, it was decided internally to use the average instead of a percentage since the goal is to eliminate administrative errors, which must not increase even if the number of orders significantly increases.

Table 3.1: Dimensions, criteria, and corresponding indicators - OMT

Dimensions	Criteria	Indicators	Preference direction
Complaints	Serious complaints *(g_1)*	Average monthly serious complaints	Minimization
	Medium complaints *(g_2)*	Average monthly medium complaints	Minimization
	Warning complaints *(g_3)*	Average monthly warning complaints	Minimization

3.1.1.2 Service Front Office department

Service Front Office is focused on managing customers' problems/requests or redirecting them internally. Front office assistants have daily interaction with customers. They help them by e-mail or by phone, control their purchase orders, give support to the logistic operation (support orders' fulfillment, i.e. support interactions with customers and transports), follow the customer stock and respective production plan, manage returns (i.e. support interactions with customers and transports), and also receive and forward the service complaints. Given these, it was deemed relevant to evaluate the quality of *commercial attendance* and *order management* (Table 3.2):

– The first focuses on monitoring the responsiveness of sales assistants, i.e. their ability to efficiently solve customers' problems and/or answer their questions. The *efficiency to attend customers' requests* refers to the situations in which sales assistants properly address customers' needs within the first call. The *waiting time calls* is the time customers wait for a sales assistant to answer a phone call and *waiting time e-mails* is the total time to solve an occurrence;

– The second concerns order processing-related issues with the indicators *order changes* and *order cancellations* focused only on changes and cancellations that happen due to internal decisions. Regarding the *short-time deliveries*, it refers to situations in which the customer places the order at very short notice, yet BA is able to deliver.

Table 3.2: Dimensions, criteria, and corresponding indicators - FOA

Dimensions	Criteria	Indicators	Preference direction
Commercial Attendance	Efficiency to attend customers' requests *(g_4)*	Number of requests met in the 1st call per 100 calls	Maximization
	Waiting time to answer e-mails *(g_5)*	Average time to close a case (in days)	Minimization
	Waiting time to answer calls *(g_6)*	Average time to answer a phone call (in seconds)	Minimization
Orders Management	Order changes *(g_7)*	Number of order changes per 100 orders	Minimization
	Order cancellations *(g_8)*	Number of order cancellations per 100 orders	Minimization
	Short-time orders delivered *(g_9)*	Number of short-time orders delivered per 100 short-time orders placed	Minimization

3.1.1.3 Customer Service department

BA Customer Service is responsible for the customers' technical support, i.e., assuring both the quality of the products and services. It is also responsible for managing complaints, analyzing and following up on them, and creating action plans to solve customers' problems.

Since a complaint is a very serious situation, guaranteeing both a quick resolution and non-repetition, is crucial to a good CX and to maintaining a great relationship with customers. Therefore, it was decided to include the following dimensions in this study: *timeliness of services* and *recurrent service complaints*. The first is assessed by the *average time to solve a service complaint*, while the second is measured by the *percentage of customers with multiple service complaints*, which refers to customers who have had multiple service issues, in a given period because of service problems.

Table 3.3: Dimensions, criteria, and corresponding indicators - Customer Service

Dimensions	Criteria	Indicators	Preference direction
Timeliness of Services	Timeliness to support customers *(g_{10})*	Average time to solve a service complaint (in days)	Minimization
Recurrent Service Complaints	Customers with recurrent service complaints *(g_{11})*	Number of customers with multiple service complaints per 100 customers with deliveries	Minimization

3.1.1.4 Logistics department

It is responsible for transporting the final products from the factory to the warehouses, maintaining and managing the stock (ensuring its integrity), and preparing and loading the trucks.

Although this department has no direct contact with the customer, its activities have a direct impact on CX. For example, if a wrong quantity is loaded, customers will submit a delivery non-compliance complaint; if pallets are not properly secured during loading, they will arrive damaged, or if a truck loading time is exceeded, it is likely to arrive late, etc.

As this department is mainly visible when problems arise, once normal functioning is expected, it was decided to measure CX through *complaints*. They were divided into three dimensions: *stock accuracy, products integrity and timeliness of services* and, similarly to the OMT complaints, the different degrees of severity were taken into consideration, as shown in the table below.

Table 3.4: Dimensions, criteria, and corresponding indicators - Logistics

Dimensions	Criteria	Indicators	Preference direction
Stock Accuracy	Serious, medium, and warning complaints due to stock inaccuracy *(g_{12}, g_{13}, g_{14})*, respectively	Number of serious, medium and warning complaints due stock inaccuracy per 100 complaints	Minimization
Products Integrity	Serious, medium, and warning complaints due to damaged pallets *(g_{15}, g_{16}, g_{17})*, respectively	Number of serious, medium and warning complaints due to damaged pallets per 100 complaints	Minimization
Timeliness of Services	Serious, medium, and warning complaints due to time non-compliance *(g_{18}, g_{19}, g_{20})*, respectively	Number of serious, medium and warning complaints due to time non-compliance per 100 complaints	Minimization

3.1.1.5 Transports department

This department focuses on efficiently and effectively managing transportation and it supports order fulfillment and returns. Thus, to assess its impact on CX, the *timeliness of deliveries*, the *tracking visibility*, and the *complaints* were analyzed.

Regarding the *timeliness of deliveries*, which concerns the time compliance of the deliveries, it was decided to consider *deliveries on-time*, *late deliveries*, and *early deliveries*. Despite the potential correlation between them, all three situations were included because they are perceived in distinct ways, having different impacts on CX. In what concerns *tracking visibility*, that examines if tracking and monitoring of deliveries by customers is possible, the chosen KPIs were: *visibility over the delivery process*, *confirmation of deliveries*, and *efficiency to inform customers of delays*. Lastly, similarly to OMT and Logistics, the evaluation of *complaints* took into consideration the different degrees of severity.

It should be noted that BA contracts suppliers to carry out deliveries, so some incidents are not the responsibility of the company. However, BA must assure they provide the best service so

that the CX is the best. It should also be noticed that customers do not make the separation of BA from the carrier since they are paying BA.

Table 3.5: Dimensions, criteria, and corresponding indicators - Transports

Dimensions	Criteria	Indicators	Preference direction
Timeliness of Deliveries	Delivery time compliance *(g_{21})*	Number of deliveries on-time per 100 deliveries	Maximization
	Delivery time non-compliance, earlier arrivals *(g_{22})*	Number of early arrivals per 100 deliveries	Minimization
	Delivery time non-compliance, late arrivals *(g_{23})*	Number of late arrivals per 100 deliveries	Minimization
Tracking Visibility	Visibility over the delivery process *(g_{24})*	Number of deliveries with real-time visibility tracking per 100 deliveries	Maximization
	Confirmation of deliveries *(g_{25})*	Number of deliveries confirmed per 100 deliveries	Maximization
	Efficiency to inform customers of delays *(g_{26})*	Number of delivery delays' warnings per 100 delivery delays	Maximization
Complaints	Serious, medium, and warning complaints (g_{27}, g_{28}, g_{29}), respectively	Number of serious, medium and warning complaints per 100 complaints	Minimization

3.1.1.6 Invoicing department

This department is responsible for processing the invoices. The customer expects BA to issue invoices with no errors and, if they occur, that they are corrected as soon as possible. The same applies to situations where there are quality and service problems that originate returns and require the issuing of credit notes. In this study, since there is no data available to calculate time-related indicators, *billing accuracy* was the dimension selected, that is assessed by the criterion *wrong invoices*, as can be seen in the table below.

Table 3.6: Dimensions, criteria, and corresponding indicators - Invoicing

Dimensions	Criteria	Indicators	Preference direction
Billing Accuracy	Wrong invoices (g_{30})	Number of wrong invoices per 100 invoices issued	Minimization

3.1.2 Preparation of weighting and aggregation methods' implementation

As shown in the previous section, the evaluation of CX in each department is done through a set of n criteria. And, each criterion is operationalized by an indicator. To obtain the CI, first, the weights for the criteria are calculated; then, a utility function for each criterion in each department is constructed; and, lastly, the utility functions are aggregated.

3.1.2.1 The DCM-SRF for determining weights

In order to determine the weights of criteria, using the DCM-SRF proposed by J. Figueira and Roy (2002), the procedure must be as follows:

1. The analyst must prepare a *set of cards* representing considered criteria, with their names on the corresponding cards (and, if necessary, some additional information) and provide it to the DM;

2. The analyst should also prepare and provide the DM with a *set of blank cards*;

3. The analyst must then ask the DM to rank the criteria cards from the criteria with the lowest to the one with the highest weight (if some are tied, cards must be placed in the same ranking position);

4. The analyst should inform the DM that two consecutive positions in the ranking may be more or less close. In order to model this *closeness*, DM can *insert blank cards* in the consecutive intervals of the ranking. The greater the number of blank cards, the bigger the difference in criteria weights. It is important to highlight that no blank cards between criteria in two consecutive positions do not mean they have the same weight, but that the difference between them is minimal;

5. Lastly, the analyst must ask DM how many times the weight of criterion(a) ranking first is higher than the weight of criterion(a) ranking in the lowest position. This number is known as *ratio z*.

See a schematic example of this method in Figure B.2, Appendix B.1.

3.1.2.2 The DCM-SRF for assessing utilities

In what concerns the assessment of utilities, the utility values of MAUT are the levels of a common interval scale, in general, within the range [0,1]. The conversion of the original criterion scales into a single common interval scale must involve the use of a method that takes the DM's preference intensity between consecutive intervals of the scale into account. This project adopts the methodology used by Bottero et al. (2018) to define an interval scale based on the DCM-SRF studies of J. Figueira and Roy (2002). Therefore, the steps described below must be followed:

1. Consider a discrete scale of criterion g, $E_g = \{l_1, ..., l_2, ..., l_k, ..., l_t\}$, where $l_1 \prec l_2 \prec ... \prec l_k \prec ... \prec l_t$ ($\prec$ implies a 'ranked strictly lower' relation) ;

2. Define two reference levels, l_p and l_q, and define their utilities. It is frequent to use $u\ (l_p) = 0$ and $u\ (l_q) = 1$;

3. Ask the DM to insert a certain number of blank cards, e_k between every two successive levels of the ranking l_k and l_{k+1}, with $k = 1, ..., t\text{-}1$, such that:

   $l_1\ e_1\ l_2 ... l_p\ e_p\ l_{p+1}\ e_{p+1} ... l_k\ e_k\ l_{k+1} ... l_{q\text{-}1}\ e_{q\text{-}1}\ l_q ... l_{t\text{-}1}\ e_{t\text{-}1}\ l_t$;

4. Consider only the levels in between l_p *and* l_q and compute the unit valuation:

$$\alpha = \frac{u(l_q) - u(l_p)}{h} \text{ ,where } h = \sum_{k=p}^{q-1} e_k + 1; \tag{3.1}$$

5. Compute the utility value *u* (l_k) for each level k, with k = 1,..., t as follows:

$$u(l_k) = \begin{cases} u(l_p) - \alpha(\sum_{j=k}^{p-1} e_j + 1), \text{for } k = 1,..., p\text{-}1 \\ \\ u(l_p) + \alpha(\sum_{j=p}^{k-1} e_j + 1), \text{ for } k = p + 1,..., q - 1,..., q + 1 ,..., t \end{cases} \tag{3.2}$$

6. Repeat the process for every criterion.

Nevertheless, this procedure is suitable for discrete scales, which is not the case of any criterion in this study. Thus, in this case of continuous criteria, the levels l_1,..., l_t can be considered as reference values whose utility values can be assessed with the aforementioned procedure. Note that the number of equidistant discrete levels, corresponding to the performances of each criterion should be defined by the DM. Once utilities *u*(l_1),..., *u*(l_t) have been assessed, the utility values assigned to the intermediate performances of criterion g_i in a given subinterval ($g_i^l < g_i < g_i^u$), are defined by linear interpolation. Thus, they are calculated through the application of the following formula:

$$u_i(g_i) = u_i(g_i^l) + \frac{g_i - g_i^l}{g_i^u - g_i^l}(u_i(g_i^u) - u_i(g_i^l)) \tag{3.3}$$

Moreover, note that the preference order of the discrete scale E of each criterion g_i varied according to the objective of minimizing or maximizing a given criterion.

3.1.2.3 Aggregation of utility functions

Lastly, the aggregation step. After obtaining the utility function and the weight of each indicator, the aggregated utility function of each department is obtained using a weighted linear additive function, as follows:

$$U_j(g_1, g_2, ..., g_n) = \sum_{i=1}^{n} w_i u_i(g_i) \qquad j \in \{1,2,3,4,5,6\} \tag{3.4}$$

where U_j is the multi-attribute utility function for the jth aspect (1 = OMT, 2 = FOA, 3 = Customer Service, 4 = Logistics, 5 = Transports, 6 = Invoicing), $u_i(g_i)$ is the utility corresponding to the ith KPI and the w_i represent the weight for the ith KPI ($w_i \geq 0$ and $\sum_{i=1}^{n} w_i = 1$).

To aggregate all the departments and obtain the overall CX, the CX must be calculated as follows:

$$CX(U_1, U_2, U_3, U_4, U_5, U_6) = \sum_{j=1}^{6} W_j U_j \tag{3.5}$$

where W_j represents the weight for the jth aspect of CX ($W_j \geq 0$ and $\sum_{j=1}^{6} W_j = 1$).

3.2 Collection of information

The second phase of this study consisted of two major moments: the first focused on gathering the data to calculate the KPIs (details in Section 3.2.1) and, the second, on collecting information to construct the CI (details in Section 3.2.2).

3.2.1 Data collection to calculate the KPIs

The data collection was done by each department, which exported the data for the years 2021, 2022, and 2023 to Microsoft Excel, and sent the raw data by e-mail. It is important to note that these specific periods were chosen due to limited data availability before 2021.

3.2.2 Data collection to construct the composite indicator

The construction of this CI involved the participation of multiple stakeholders and it followed a 3-step procedure: firstly, during a *preliminary phase*, a series of meetings with several groups of people from different departments and geographies were conducted. Secondly, *meetings with the departments' managers*, to gather the information to calculate the KPIs' weights and utility functions. And, finally, a *meeting with the sales' directors*, to aggregate the department's utility functions. Details about each step are presented below.

3.2.2.1 Preliminary phase

These meetings consisted of an adaptation of the focus group technique in which a group of employees, carefully selected based on their experience and expertise, discuss the impact of some of their department's activities on the CX. Due to the geographical dispersion of the teams, the meetings were held online, via Microsoft Teams, and as a way of orienting the group discussions, it was decided to elaborate one questionnaire per department, in English (the official language of the company), using Microsoft Forms, and based on the DCM-SRF method. To guarantee all the participants understood the method, the first part of the questionnaire consisted of a real example of a decision-making situation, followed by a detailed explanation of the DM's reasoning, accompanied by an illustration (see Figure B.3 and B.4, in Appendix B.1).

The meetings were structured in three serial stages. Firstly, the initial greetings, followed by the presentation of the project - context and objectives, an explanation of the meeting's structure, and what was expected from the participants. Secondly, the introduction of the questionnaire and of the DCM-SRF - analysis of the initial example, detailed explanation of the method, clarification of doubts, and presentation of each KPI's definition. After this, participants were asked to discuss and reach a consensus regarding the relative impact of each indicator, applying the DCM-SRF reasoning.

After each meeting, a follow-up was sent for a chat created in Microsoft Teams for each group, with a summary of the meeting's conclusions. The participants were asked to validate them and

also to share what they believed would be interesting to consider in this study, whenever they remember something.

3.2.2.2 Application of the DCM-SRF to weight KPIs and to determine their utility functions

A series of meetings with the manager(s) of each department were convened (see Table B.2, Appendix B.2). They were aimed to collect the information for calculating the KPIs' weights, and their utility functions and to set the performance targets for each KPI.

Likewise the preliminary phase's meetings, these ones also had three moments. First, the initial greetings, the project presentation, an overview of the meeting's structure, and the definition of the manager's role during the meeting. Secondly, the explanation of the DCM-SRF, and of each KPI that would be under analysis. Lastly, the implementation of the DCM-SRF both for the determination of KPIs'weights, following the procedure described in Section 3.1.2.1, and then, for the obtention of utility functions, detailed in Section 3.1.2.2. Additionally, in some situations, qualitative labels were added to the levels, to facilitate the utility functions' creation.

After that, managers were asked to set two target levels for each KPI, so that a 3-color scale could be created (*red* corresponding to activities that *need significant improvement*, *yellow* to the ones that *need improvement*, and *green* to those that are already *satisfactory*).

The manager's inputs were written down on paper and then recorded in Microsoft Excel. And, after each meeting, a follow-up was sent with a summary of the meeting's conclusions (i.e. the results of the weights, the utility functions, and the targets set), with a validation request.

3.2.2.3 Application of the DCM-SRF to determine departments' weights

In order to aggregate the utility functions of the various departments, it was necessary to hold a meeting with the 4 sales directors, who have a broader perspective of the department's impact on CX. It was convened online and was structured in 3 parts, following the structure of the meetings with the departments' managers. The only difference between these meetings was in the third step, since in this one, DecSpace[2], a web application, was elected to support the decision process. Thus, a brief presentation of this tool was also done.

3.3 Analysis of information

Similarly to the collection of information, the analysis of information had two phases. The first one focused on obtaining the KPIs, and the second, on the CI's construction.

3.3.1 Analysis of information to calculate KPIs

The raw data provided by the departments were processed and the KPIs were calculated, using Microsoft Excel. The KPIs were calculated for all the quarters between 2021 and 2023, and annually for 2021 and 2022.

[2]See `http://app.decspacedev.sysresearch.org/`

3.3.2 Analysis of information to construct the composite indicator

3.3.2.1 Determination of weights

In order to obtain the weights of criteria using the DCM-SRF, DecSpace was used. The procedure consisted of:

1. Creating the criteria, i.e. creating the cards to which the criteria correspond, and specifying their scale units and their desired directions (minimization or maximization);

2. Ranking the criteria, according to the DMs' decisions, and, then, inserting the blank cards;

3. Finally, defining *ratio z*, the number of decimal places, and the "weight type" (normalized or non-normalized).

By following these steps, and executing the workflow, DecSpace generated the results, that appeared in the corresponding output module in a table containing the name of the criteria and the respective non-normalized and normalized (in accordance with the "weight type" defined).

3.3.2.2 Determination of the utility functions

The determination of the utility functions was done in Microsoft Excel. For each indicator, a table was created in which the information gathered in the meetings with the departments' managers was inserted, i.e. the *values of reference levels*, the *number of levels*, and the *number of blank cards*. In this table, a column was programmed to calculate the value of the utilities, i.e. to convert the KPIs into utilities, using the Equation (3.3).

3.4 Presentation and visualization of results

After the CI construction, it is important to identify the best tools to present the results, taking into account the data analyzed, and to assure the communication of the most information in a clear and accurate manner. For this project, *7 dynamic and connected dashboards* were created (a dashboard for the global overview, and 6 individual dashboards, one for each department for a more detailed analysis). All of them include a header containing the title and traffic light icons to indicate the overall current performance and, either the current performance of each department (in the global overview dashboard) or of each KPI within each department's dashboard. In the dashboard's body, at least a *line chart*, to visualize the evolution over time is presented; along with a table to see the real performances of each period. The user can click on the buttons to select the period(s), i.e. year(s) and quarter(s). Additionally, colored target lines were added to the line graphs, making it easier to see the status' evolution. The footer was transformed into a navigation bar so by clicking on the icons, users are directed to the correspondent dashboard.

3.5 Summary

This chapter addresses the procedures to prepare, collect and analyze the information in the main stages of B2B CX CI construction. Therefore, it examines the selection of variables; the DCM-SRF to assess the KPIs' and departments' weights, and to determine the KPIs' utility functions and respective utilities; and, lastly, it analyzes the MAUT to determine departments' utility functions and to ascertain the overall CX. The next chapter implements the described methodology.

Chapter 4

Case study: European glass packaging manufacturer

This chapter covers the empirical application of the methodology presented in Chapter 3. Firstly, an overview of the glass packaging industry, the company, and the project is provided in Section 4.1. Secondly, Section 4.2 details the stakeholders and the selected DMs. Lastly, the main results are presented and discussed in Section 4.3. The main dashboard is also presented in Section 4.3.4.2.

4.1 Overview

4.1.1 An overview of the glass packaging industry

Glass has been part of human life for thousands of years and, today, it is present in millions of homes. It is a very popular choice and one of the most preferred packaging materials for end-users concerned about the health and environment[1].

According to Glass Alliance Europe, glass packaging is the largest sector of the European glass industry[2] and the competition is fierce. Nowadays, this industry is facing big challenges. After the difficult years of the COVID-19 pandemic; the inflation and energy crisis have significantly increased production costs, undermining this material's competitiveness. Additionally, studies predict that the market growth rate will be slowed by fluctuations in raw material prices and intense competition from other packaging materials. Moreover, it is believed that the strategies used in the past to keep competitive will no longer be enough. The history of this industry shows the focus has been on productivity (producing more in less time and using fewer raw materials), quality, and, more recently, sustainability; but this change of market paradigm requires a new approach. Enhancing CX, till now unexplored within this industry, can be a key source of competitive advantage.

[1] See https://feve.org/about-glass/facts-product-details/

[2] See https://www.mordorintelligence.com/industry-reports/europe-glass-packaging-market

4.1.2 An overview of the company

As mentioned in Chapter 1, BA produces food and beverage glass packaging in Europe. The company has grown consistently over the years and, in 2022, despite the energy crisis and inflation, had very positive results with an EBITDA (Earnings Before Interest, Taxes, Depreciation, and Amortization) exceeding 300 million euros. Nevertheless, in the same period, the company obtained the worst CSI results ever. For the first time, BA obtained a negative NPS, which is critical in particular considering the highly competitive industry in which the company operates. Besides all the glass container producers, the group faces significant competition from the packaging producers, mainly based on price, innovation, quality, delivery, and customer service.

The competitors' pressure as well as the understanding of the importance of CX, triggered BA to act. The company is determined to improve this year's CSI and the CX. Processes that impact customers are being reviewed and improved, a series of activities are being developed to strengthen relationships, there is a focus on guaranteeing high-quality products and services, and efforts to create innovative, lightweight products with appealing designs and brilliant color schemes. Complementarily, the creation of a CX measurement instrument will help BA to assess its CX and to spot the most critical areas where improvements must be made.

4.1.3 An overview of the project

Nowadays, both B2C and B2B companies are focused on guaranteeing the high quality of their offerings (both goods and services), and their experiences, so they set themselves apart from competitors and gain a competitive advantage in the market. However, monitoring and managing CXs is hard due to their high complexity. In addition to this, as the literature review demonstrated there is a knowledge gap in the B2B CX field of study, and in particular, in what concerns quantitative measurement instruments, which makes this task even more difficult.

This project addresses this problem by developing a new procedure to measure the CX, in a B2B setting, using a CI approach, which as demonstrated in Chapter 2, is an adequate technique for complex and multi-dimensional concepts like CX. To the author's knowledge, it is the first time a CX CI is created and tailored to a glass packaging manufacturing company.

4.2 Actors

In every decision-making process, there are at least three types of actors involved, the analyst, stakeholders, and DMs. In the following subsections, the actors within these categories, that participated in this project, are presented in detail.

4.2.1 Analyst

Starting with the analyst, in this research, this person was responsible for identifying the problem, the stakeholders, and the most relevant criteria, selecting and applying the most suitable weighting and aggregating methods, and, lastly, interpreting the results and drawing conclusions.

4.2.2 Stakeholders

Every organization has a system of key stakeholders with whom it establishes and manages relationships. Stakeholders can be *internal*, i.e. people with a direct relationship with the company, for example, employees, and owners; or *external*, such as customers, suppliers, creditors, and shareholders, i.e. those who although do not work directly with a company, are affected somehow by its actions and its business outcomes (Hitt, Ireland, & Hoskisson, 2016). The most important stakeholders of BA are presented below, as well as the impact they have had on the project and/or the project will have on them.

4.2.2.1 Internal Stakeholders

Employees are crucial to any business's success or failure. In BA, employees create, manufacture, and sell products. Their hard work, enthusiasm, dedication, and commitment are responsible for the business growth and CS. In this project, their collaboration was essential. Firstly, they were crucial to gather all the information required to select the most relevant criteria. Secondly, as DMs, their experience and in-depth knowledge of customers have greatly enriched this work (more details in Section 4.2.3). Lastly, after implementing this CI, the real benefit the company can get from it will come from the employees' ability to improve the CX based on data evidence.

4.2.2.2 External Stakeholders

Customers are at the heart of this project since its goal is to create a metric to measure and monitor the CX so that it can be continuously improved. Although they were not directly involved in this project, because it was not possible to contact them, among the stakeholders, they are the ones who will be most affected by this study.

Suppliers of raw materials, materials for production support, and transportation services are critical to BA's business. Any disruption in their business can significantly impact BA's operations and consequently, the CX.

Shareholders, the ones who own BA's shares, support the continuous investment and modernization of BA's operations, helping the company grow sustainably. The goal of BA as of any company is to create value for its shareholders. Thus, since this project aims ultimately to improve the CX and increase CS, it is expected that it will help increase revenues, therefore, creating more value for shareholders.

4.2.3 Decision-makers

As described in detail in Chapter 3, the DCM-SRF is a participatory method. Thus, to obtain the CI it is required the involvement of DMs. It was decided that the most appropriate DMs for determining the weights of KPIs, their utility functions, and setting the respective targets were the departments' managers, as they have a deep knowledge of all the daily operations within their department, i.e. of the tasks, interactions with customers, problems, pain points, etc. (see

more details about the DMs in Table B.2, Appendix B.2). Regarding the determination of the departments' weights, the 4 sales directors, who have the broadest vision and are closest to the customers, were the selected ones.

It should be noted that it was important to have the participation of both the departments' managers and the sales directors because their knowledge level is different. The departments' managers possess extensive knowledge within their own departments but lack a global perspective. On the other hand, sales directors have a global view but lack depth, making it impossible for them to assess certain details. Thus, by combining them, it was possible to get the complete picture.

4.3 Results and Discussion

This subchapter presents and discusses this study's findings.

4.3.1 Determination of weights and utility functions, per department

As explained, meetings with the departments' managers were conducted to gather the required information for calculating the KPIs' weights and their utility functions and to set the performance targets for each KPI. The results exhibited next, appear in the same order in which values were collected.

4.3.1.1 Determination of KPIs' weights and departments' utility functions

First, the determination of the KPIs' weights:

- As shown in Section 3.1.2.1, OMT, Logistics, and Transports departments have KPIs based on complaints. Thus, in order to harmonize the evaluation of these KPIs, it was decided that the Customer Service manager[3] should decide the weights of this group of KPIs. See the results in Table C.1, Annex C;

- Regarding FOA and Transports, the departments with the larger number of variables, to avoid DMs' cognitive stress and to facilitate their decisions, instead of comparing all KPIs with each other, DMs started by comparing the dimensions, and then KPIs within each dimension. The weight of each KPI was obtained by multiplying the weight of the corresponding dimension and the weight obtained with the KPIs' comparison within that dimension. To see FOA's results please consult Tables C.2 and C.3 of Annex C, and for the Transports', Table C.5, Table C.6 and Table C.7;

- As for Customer Service, the expert considered that the 2 KPIs under analysis had the same impact on CX, thus, the weights are fifty-fifty;

[3]Remember that the Customer Service department is responsible for managing complaints, analyzing, and following up on them. Therefore, members within this department have a very clear idea of the impact of the different complaints on the customer.

– Regarding Logistics, the different types of complaints were weighted by 2 of the 3 managers, and the results are shown in Table C.4;

– Lastly, the Invoicing department, since there was only one KPI, the utility function of this department is equal to the KPI's utility function.

All the results presented in Tables C.1 to C.7, Appendix C.1, were calculated using DecSpace. With them, it was possible to obtain the utility functions for each department, as shown below:

$$U_1(g_1,g_2,g_3) = 0.6410 * u_1(g_1) + 0.2942 * u_2(g_2) + 0.0641 * u_3(g_3) \tag{4.1}$$

$$\begin{aligned} U_2(g_4,g_5,g_6,g_7,g_8,g_9) =& 0.2500 * u_4(g_4) + 0.1250 * u_5(g_5) + 0.1250 * u_6(g_6) \\ &+ 0.1000 * u_7(g_7) + 0.2000 * u_8(g_8) + 0.2000 * u_9(g_9) \end{aligned} \tag{4.2}$$

$$U_3(g_{10},g_{11}) = 0.5000 * u_{10}(g_{10}) + 0.5000 * u_{11}(g_{11}) \tag{4.3}$$

$$\begin{aligned} U_4(g_{12},...,g_{20}) =& 0.5030 * [0.6410 * u_{12}(g_{12}) + 0.2949 * u_{13}(g_{13}) + 0.0641 * u_{14}(g_{14})] \\ &+ 0.4467 * [0.6410 * u_{15}(g_{15}) + 0.2949 * u_{16}(g_{16}) + 0.0641 * u_{17}(g_{17})] \\ &+ 0.0503 * [0.6410 * u_{18}(g_{18}) + 0.2949 * u_{19}(g_{19}) + 0.0641 * u_{20}(g_{20})] \end{aligned} \tag{4.4}$$

$$\begin{aligned} U_5(g_{21},...,g_{29}) =& 0.0812 * u_{21}(g_{21}) + 0.0290 * u_{22}(g_{22}) + 0.2898 * u_{23}(g_{23}) \\ &+ 0.0476 * u_{24}(g_{24}) + 0.0190 * u_{25}(g_{25}) + 0.1333 * u_{26}(g_{26}) \\ &+ 0.2564 * u_{27}(g_{27}) + 0.1180 * u_{28}(g_{28}) + 0.0256 * u_{29}(g_{29}) \end{aligned} \tag{4.5}$$

$$U_6(g_{30}) = 1.0000 * u_{30}(g_{30}) \tag{4.6}$$

4.3.1.2 Determination of KPIs' utility functions

As explained in Section 3.1.2.2, each KPI is activated by a utility function. Therefore, 30 utility functions were created and are shown in Tables C.8 to C.26, Appendix C.2 [4].

As it can be observed, each table has 8 columns. The information inserted in column 1 (*number of levels*), column 2 (*values of reference levels*), and column 3 (*number of blank cards*) was provided by the DM(s). The fourth column was obtained by adding 1 to each level's number of blank cards and then summing all. The fifth column's value was reached by subtracting the lowest utility from the highest (in this case, it was equal to 1 for all KPIs) and dividing this result by the value of column 4. To calculate the values of columns 6, 7, and 8, Equation (3.3) was applied.

It can also be seen that all the tables have a row with a reference level labeled as 'Entity#1', and it intends to be the row in which the real performance of the KPI enters, to then be converted into a utility. The cell in column 7 (the linear interpolation column), was programmed to verify between which levels the actual performance lies, and then execute Equation (3.3).

[4] Note that in the annexes, there are only 19 tables, because different KPIs have the same utility function, and for the sake of conciseness, it was decided to not repeat tables providing the same information. Thus, Table C.9, Table C.18 and Table C.19 in Appendix C.2, correspond to the utility functions of more than one KPI, being those KPIs duly referenced in the tables' titles.

In addition to these tables, utility plots were created to graphically represent the utility functions. With these illustrations is possible to visualize both the DMs' preferences in defining the levels with the blank cards and the monotonicity of the functions. Below, the utility plots of g_{21} and g_{22} are exhibited [5]. On the left, it is possible to observe a monotonically increasing function, since the preferred direction of g_{21} is maximization; and on the right, a monotonically decreasing function, since the preferred direction of g_{22} is minimization.

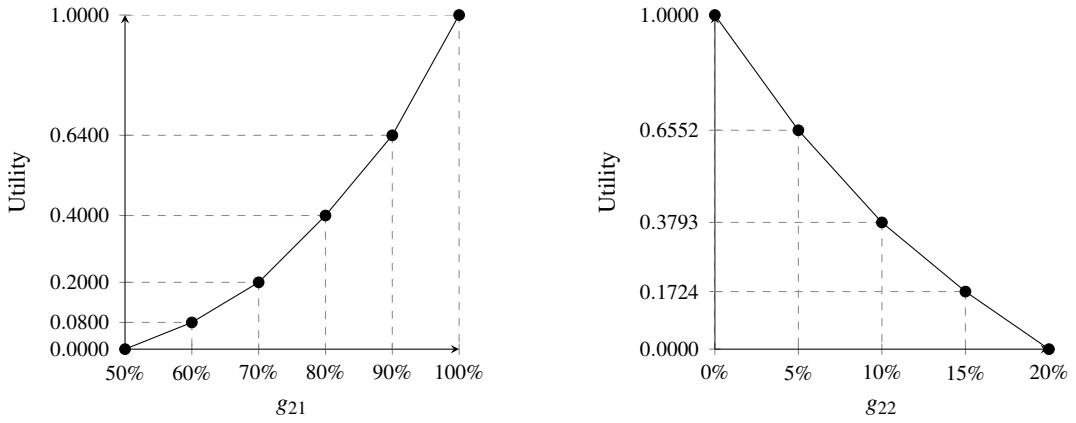


Figure 4.1: Utility plot of criteria g_{21} (on the left) and g_{22} (on the right)

4.3.2 Determination of the overall customer experience's utility function

In order to obtain the weights to aggregate the department's utility functions, the procedure detailed in Section 3.1.2.1 was followed in a meeting conducted with the 4 sales directors.

The DMs considered FOA to be the most impactful department since it is the one that has daily contact with customers, and the relationship established between customers and sales assistants plays a major role in their CX. In what concerns Logistics and Transports, it is believed they have the same impact as from the customer's perspective they are not separate departments, and failures in their activities have similar consequences. Regarding Customer Service, it has a major impact on CX when there is a quality or service problem that originates a complaint. But, since these situations are not so frequent, most customers never contact this department. However, it should be noted that it has more responsibilities apart from being the "firefighter". It is expected that the Customer Service team visits customers, develops projects with them, etc. Hence sales directors decided to rank it in the third position. Lastly, it was deliberated that OMT and Invoicing were the departments with the smallest impact. Customers rarely associate problems in their orders with OMT and, when it comes to Invoicing, usually problems are solved between the customer's purchasing team and BA Invoicing team in a smooth way.

Below, Table 4.1 presents the results, calculated using DecSpace.

[5] For the sake of conciseness, it was decided to present just these two examples as they are sufficient to show the DMs' preferences in defining the levels with the blank cards and the monotonicity of the functions.

Table 4.1: Determining the normalized weights of each criterion for $w = 1$ and $z = 8$

Rank r	Criteria in rank r	Number of blank cards	Non-normalized weights	Normalized weights
1	FOA	2	8.00	0.3167
2	Logistics, Transports	0	5.38	0.2130
3	Customer Service	3	4.50	0.1781
4	OMT, Invoicing	0	1.00	0.0396

Rewriting the Equation (3.5), the overall CX utility function obtained is:

$$CX = 0.0396 * U_1 + 0.3167 * U_2 + 0.1781 * U_3 + 0.2130 * U_4 + 0.2130 * U_5 + 0.0396 * U_6 \quad (4.7)$$

It is worth mentioning that sales directors stressed the importance of including more departments in this CI, namely Returnable Packaging, and Planning and Production, because of their great impact on CX. Though, as mentioned, they had to be excluded due to a lack of data available.

4.3.3 Calculation and analysis of departments' utilities

After obtaining the KPIs' utility functions, the KPIs and respective targets were converted into utilities[6]. Then, the KPIs' utilities were introduced in the departments' utility functions, shown in Section 4.3.1.1, and the departments' utilities were calculated.

The results of each department's utilities, per period[7], as well as the respective targets, are presented next along with the respective analysis.

4.3.3.1 Order Management Team

Table 4.2: Utilities OMT

	2021					2022					2023	
Period	Q1	Q2	Q3	Q4	Annual	Q1	Q2	Q3	Q4	Annual	Q1	...
Utilities	0.9887	0.2949	0.2949	0.8677	0.6001	0.8677	0.9686	0.8839	0.6642	0.7843	0.9480	

Status Green: $U_1 \in]0.6808; 1.0000]$; Status Yellow: $U_1 \in]0.3921; 0.6808]$; Status Red: $U_1 \in [0.0000; 0.3921]$

As can be seen in the table above, the worst periods are 2021Q2 and 2021Q3, followed by a notorious recovery in the subsequent quarter. In the year after, there was a decrease from 2022Q2 to 2022Q3 and 2022Q4, with a change in the status, to yellow, in the last quarter. These results accurately reflect what happened. In 2021, there was a change in the team's leader, and the number of complaints had a significant increase. In order to revert this situation, complaints started being

[6]Due to confidentiality, neither the KPIs nor the individual utilities of each KPI are disclosed in this work.
[7]Note that: Q1 refers to Quarter 1, Q2 to Quarter 2, Q3 to Quarter 3, and Q4 to Quarter 4.

typified and analyzed with more detail and care. This was crucial to the identification of tasks that could and should be automated, reducing manual work and the probability of human errors. Additionally, system alerts were implemented, which allowed the early detection of errors. This helped to eliminate some types of complaints and significantly reduce others. The implementation of these happened by the end of 2021Q3, which justifies the improvement of the last quarter.

Regarding the decrease verified in 2022Q3, there were internal problems that justify these results. A member of the team was underperforming (and was responsible for most of the complaints the department was receiving). Thus, there was a replacement, but during the adjustment period, the problems did not disappear immediately. However, as demonstrated by the 2023Q1 result, the problems were solved and the performance has improved to a satisfactory level.

4.3.3.2 Service Front Office

Continuing the analysis with the FOA department. As can be observed the third and fourth quarters of both years, 2021 and 2022, were slightly worse than the first ones, being Q3 the one with the lowest results.

Table 4.3: Utilities FOA

	2021					2022					2023	
Period	Q1	Q2	Q3	Q4	Annual	Q1	Q2	Q3	Q4	Annual	Q1	...
Utilities	0.5714	0.5048	0.4414	0.3637	0.4708	0.4375	0.4578	0.4434	0.3976	0.4345	0.6109	

Status Green: $U_2 \in]0.5293; 1.0000]$; Status Yellow: $U_2 \in]0.1659; 0.5293]$; Status Red: $U_2 \in [0.0000; 0.1659]$

The results of Q3 can be explained by the sales volume increase in conjunction with a decrease in the number of front office assistants, due to holiday leave. With the volume of work increasing and being divided by fewer people, the quality of the service provided is affected. Nevertheless, there are strategies being implemented to minimize these oscillations.

In what concerns the decrease verified during Q4, this is mainly due to the company's sales strategies in reaction to customers' behaviors. With the year's turn, prices increase so lots of customers order large quantities at the end of Q4, to save money, creating piles of stock in their warehouses. To avoid that, BA implements some sales restrictions/rules, that consequently result in an increase in order changes and cancellations. Moreover, the slight decrease verified in some periods of 2022 can be partially justified by the decrease in the number of *short-time orders delivered*, which was a consequence of internal decisions. Both the sales rules and the short-time orders limitation impacted negatively the U_2.

Furthermore, results were also impacted negatively by the fact that some sales assistants did not close the cases immediately after solving them. It is believed the average time calculated does not correspond to reality in most cases. To understand better that impact, U_2 was calculated excluding it. And, the results showed that apart from Q4, all the other periods in yellow changed to green. Still, this KPI has a significant impact on CX and should be kept; sales assistants should

start registering the closure of the cases in time (this was stressed at the beginning of this year, and it is already possible to see an improvement in 2023Q1).

In addition to this, it is important to notice that in the past few years, there were some changes in the team, i.e. there were many people leaving and others arriving, which originated some instability in the service provided since the new members took some time to be fully integrated and delivering according to expectations.

4.3.3.3 Customer Service

Proceeding with Customer Service, the results, presented in the table below, show that performance is better in the first and last quarter of each year.

Table 4.4: Utilities Customer Service

	2021					2022					2023	
Period	Q1	Q2	Q3	Q4	Annual	Q1	Q2	Q3	Q4	Annual	Q1	...
Utilities	0.4653	0.3093	0.0000	0.4259	0.3473	0.4848	0.2704	0.0000	0.4243	0.3750	0.4680	

Status Green: $U_3 \in]0.7149; 1.0000]$; Status Yellow: $U_3 \in]0.4486; 0.7149]$; Status Red: $U_3 \in [0.0000; 0.4486]$

This can be justified by a lower volume of orders, which decreases the pressure and consequently the stress of the company and enables a more stable and better service.

Traditionally, at the beginning of the year and at the end, the sales volume is smaller. In Q1, customers order less and less frequently, either because they are depleting their stock of the previous year, or because the volume of their operations is lowered during this period. In Q4, the decrease happens usually because of two reasons. On the one hand, beverage consumption decreases, and consequently beverage packaging sales. On the other hand, companies start preparing for Christmas campaigns in advance, usually buying big quantities during Q3, reaching their production peak during Q3 or in early Q4. During Q4 their production commonly slows down, as well as their orders.

In the second and third quarters, there is a significant increase in sales, and, therefore, the number of occurrences typically increases. The problems are a bit aggravated because it is usually during these periods that the Customer Service team is smaller, because of holidays, thus decreasing BA's ability to manage and solve complaints. Furthermore, there are environmental factors during these periods, such as high temperatures, that contribute to an increase in the number of complaints. More details on this issue, appear in Section 4.3.3.4 and Section 4.3.3.5.

Besides that, taking a closer look at Q2 and Q3 results, it is clear that 2022 had worst results than 2021. This is explained by a failure in the implementation of a change in the complaints' treatment process, and by an abrupt change of team members. In spite of numerous training sessions, the complexity of the new process (less clear and intuitive than the older), and the difficulty felt and not overcome by the people, dictated its failure. Thus, by the end of Q3, the process

change was reverted, which justifies the improvement verified in 2022Q4 and 2023Q1. Yet, the results are far from reaching the green status being clear the need for improvement.

It is important to highlight that although the annual averages of 2021 and 2022 are similar, the real improvement was actually significant. From 2021 to 2022 the complaints' management procedure changed. Prior to 2022, one complaint could include many incidents. But from 2022, each complaint started being associated with only one incident. As each incident started generating individual complaints, the total number would be expected to increase with this change. However, the increase was very low, which shows an improvement.

4.3.3.4 Logistics

As mentioned in Section 3.1.1.4, this department's utility function is based on complaints regarding *stock accuracy, product integrity, and timeliness of services*. Similarly to FOA and Customer Service, there are performance oscillations between the different quarters, with Q3 being the most critical period.

Table 4.5: Utilities Logistics

	2021					2022					2023	
Period	Q1	Q2	Q3	Q4	Annual	Q1	Q2	Q3	Q4	Annual	Q1	...
Utilities	0.7766	0.4297	0.4814	0.8083	0.5527	0.9765	0.8790	0.4372	0.7740	0.7761	0.8189	

Status Green: $U_4 \in]0.6667; 1.0000]$; Status Yellow: $U_4 \in]0.2652; 0.6667]$; Status Red: $U_4 \in [0.0000; 0.2652]$

This can be explained by several reasons. First, during Q3, the temperature inside trucks is very high (it easily reaches 70ºC/80ºC), which affects the foil protecting the pallets. With the heat, the foil becomes softer, then when the truck cools down, the foil hardens. This affects the foil's properties, reducing its ability to ensure the stability of the pallets, consequently, increasing the risks of pallets damage. In addition to this, as mentioned, in Q3 the sales volume is higher and the number of employees lower. Thus, employees have to work fast being more prone to errors.

Besides Q3, in general, the results are really positive (7 of the 11 periods have green status). Over the past few years, procedures were developed and implemented concerning the loading process, so that everything runs smoothly, without product damages or delays. This contributed to a significant improvement in the *timeliness of deliveries*, and *products integrity* (pallets are well-secured with corners and belts). Although accidents may occur, they are rare if the procedures are properly followed.

Moreover, regarding *stock accuracy*, the routines to ensure its precision are well-developed and mature. However, its quality is greatly influenced by the workers. In fact, internally it was verified that every time the warehouse responsible changed, the stock accuracy decreased. This helps to explain some of the oscillations between the periods.

4.3.3.5 Transports

As highlighted in Section 3.1.1.5, BA contracts carriers to effectuate the deliveries and, to assure the best CX is provided BA must contract the best suppliers. In order to make an optimal choice, the company evaluates each supplier annually, being *operational performance*, *supplier complaints*, and *price*, some of the parameters assessed. The results are taken into account when signing new contracts, thus it is expected an increase in the results between years. However, it is not so linear since there are many external factors, such as strikes (e.g. truck drivers' strikes, port strikes, etc.), waiting time to cross borders (especially critical in Southeastern Europe, with trucks waiting more than 1 day to cross borders), etc. that may happen and are not controllable.

Table 4.6: Utilities Transports

	2021					2022					2023	
Period	Q1	Q2	Q3	Q4	Annual	Q1	Q2	Q3	Q4	Annual	Q1	...
Utilities	0.5013	0.3991	0.5149	0.5097	0.4667	0.4478	0.3707	0.4357	0.5473	0.4313	0.5935	

Status Green: $U_5 \in]0.7273; 1.0000]$; Status Yellow: $U_5 \in]0.4028; 0.7273]$; Status Red: $U_5 \in [0.0000; 0.4028]$

When comparing 2022 results with 2021, it is clear they got worse. In fact, 2022 was considered an *annus horribilis* for Transports. The beginning of the Ukraine war caused fuel prices to skyrocket. Consequently, truckers' strikes contesting the price increase, started to happen. In Spain, for example, a truckers strike lasted a month causing many disruptions in the supply chains[8] (partially justifying 2022Q2 results). Additionally, carriers decided to increase their prices to face their cost increase, which generated conflict situations due to price disagreement. Negotiation processes took too long, which created a major stalemate. On top of that, in Southeastern Europe there was a labor shortage problem because many truck drivers went to Ukraine to fight in the war, and others simply left the region.

Besides that, a closer look at the fluctuations throughout the years, shows a similar pattern to other departments, i.e. Q2 and Q3 are slightly worse; and the reasons are also the same, a combination of sales increase and fewer employees. However, in general, the values are low, and show there is much to be improved, especially with regard to communication (customers should always be informed of delays and early arrivals) and, the timeliness of deliveries. The transport team and suppliers must understand that some early arrivals cause major constraints for customers, and should not be seen as a positive thing.

In line with this, Transports department has already implemented some strategies to improve:

– Tracking visibility - BA has been negotiating with its suppliers so they adopt Sixfold[9], which is an information system that allows real-time visibility, sends alerts when a truck is delayed, etc. To incentivize carriers to start using Sixfold, BA is going to charge suppliers for manual

[8] See `https://www.thelocal.es/20220321/how-the-truck-drivers-strike-is-affecting-life-in-spain`

[9] See `https://sixfold.com/`

updates. Additionally, for those suppliers that are not keeping the tracking information updated, BA will delay its payments until it is done;

– Timeliness of deliveries - BA is using platforms that allow the creation of automatic reports that increase visibility over the suppliers' performance. This is expected to contribute to better delivery management by carriers.

4.3.3.6 Invoicing

Lastly, the Invoicing department. Since this department's utility function is based on a single indicator, a factor was applied to the results to guarantee confidentiality. Observing the table, it is possible to see that 2021 results were far better than 2022 and that at the beginning of this year, there is already a recovery.

Table 4.7: Utilities Invoicing

	2021					2022					2023	
Period	Q1	Q2	Q3	Q4	Annual	Q1	Q2	Q3	Q4	Annual	Q1	...
Utilities	0.9135	0.9554	0.9104	0.8246	0.9119	0.5493	0.0000	0.0000	0.1954	0.0000	0.8513	

Status Green: $U_6 \in]0.7857; 1.0000]$; Status Yellow: $U_6 \in]0.3929; 0.7857]$; Status Red: $U_6 \in [0.0000; 0.3929]$

Historically, prices were negotiated once a year, at the end of the year, and were valid throughout the following year. Thus, Q1 was typically the period with poor results[10], because as prices changed, the company's system needed to be updated, and mistakes in the insertion of the new values were usual.

In 2022 this reality changed. The inflation and energy crisis forced the company to increase its prices several times throughout the year. Prices started changing in March, but it was during April 2022 that the most significant changes happened, i.e. the number of price changes was the highest as well as the number of clients impacted, and the magnitude of the difference was the biggest. The high number of changes concentrated in a short period of time generated lots of wrong invoices, which justifies the bad results of 2022Q2. During the rest of 2022, prices were updated, more or less, on a monthly basis, and without stabilization, the performance continued to be weak. In addition to this, the strike in Spain, in March 2022, resulted in the delay of the deliveries to the following month. Some customers who were supposed to receive the deliveries in March only received them in April, and as prices changed, invoices had a new price. This generated customer dissatisfaction and anger since they did not agree to pay those orders with the new price, forcing BA to correct the invoices.

Within the company, and within the glass packaging industry, it is believed that the reality prior to 2022, of annual prices, is now history, i.e. it will not be repeated. In this new much more complex reality, prices will have multiple oscillations throughout the year, so it is necessary to

[10] 2021 results do not reflect that but Invoicing team members were not able to justify this.

implement strategies that facilitate price updates in the system, to avoid this type of error. Some automation mechanisms have already been implemented, being possible to see the improvements at the beginning of 2023.

4.3.4 Calculation and analysis of customer experience

To obtain the overall CX, the partial results of each department, in a given period, were inserted in Equation (4.7), being possible to reach the results presented in the table below.

Table 4.8: Utilities customer experience

	2021					2022					2023	
Period	Q1	Q2	Q3	Q4	Annual	Q1	Q2	Q3	Q4	Annual	Q1	...
Utilities	0.6114	0.4410	0.3997	0.5388	0.4880	0.5844	0.4977	0.3614	0.5170	0.4926	0.6489	

Status Green: $CX \in]0.6499; 1.0000]$; Status Yellow: $CX \in]0.3066; 0.6499]$; Status Red: $CX \in [0.0000; 0.3066]$

In line with the department's results, it can be observed that the overall CX is lower in the second and third quarters. This means that BA provides a poor experience during these periods. As explained this is mainly due to the combination of the sales increase and employees decrease (due to holiday leave). Nevertheless, these reasons cannot be used as an excuse. BA ought to develop strategies that allow a better reaction to these factors so that CX is more stable and higher throughout the year.

Additionally, it can be observed that, overall, the results of 2022 were insignificantly better than 2021. In fact, the difference is negligible. Nevertheless, the results of the first quarter of 2023 show an impressive upgrade. The 2023Q1 utility is the highest value within the analyzed period, almost reaching the status green. However, if the 2021 and 2022 trends persist, and strategies that contribute to a more significant improvement are not implemented, it is expected that status green will not be reached until 2024.

4.3.4.1 Comparison between the composite indicator results and the customer satisfaction index

The 2022 CSI results were the worse since this metric was implemented, with BA scoring a negative NPS in all divisions. Moreover, taking a closer look at each department's results, almost all obtained a negative score, except FOA and Production. That means customers recognized the quality of commercial attendance as well as of the BA products.

When comparing these results with the CI, it can be observed that 2022 results were higher than 2021. But, the magnitude of the difference was really small. Although results seem contradictory, they should be examined in light of the 2022 context - war, inflation, and energy prices. As prices increased, customers became more demanding, their expectations rose, and their tolerance for failures diminished. Consequently, they may have given a negative rating to BA because they felt the higher prices did not have a reflection in terms of quality and overall experience.

Additionally, it is important to highlight that the CSI included the evaluation of the quality of BA products, Returnable Packaging, Planning, Production, and Pricing, which did not enter in the CX CI due to lack of data, as explained in Section 3.1.1. Therefore, if there were significant issues in these areas that contributed to customers' dissatisfaction, with the CX CI it will not be possible to see that. This comment can be extended to other aspects customers do value and were not considered because of the data unavailability. Some examples of these are mentioned in Chapter 5.

Furthermore, unlike CX CI, in the CSI the aggregation method used was equal weighting, which as already explained is not the most adequate procedure for this situation.

On top of that, it is important to stress that the CSI questionnaire used to collect the data to obtain the CSI had some flaws, that may have led to inaccurate insights, such as:

- In the questionnaire it was not explicit that the scales were based on NPS (i.e. '0-6' - detractors, '7-8' - neutral, '9-10' - promoters) thus, it is believed that many people may have filled it thinking that from '0-5' it means negative, and from '5-10' positive;

- BA does not know if the person filling in the questionnaire answered all questions with knowledge of the facts or if some were random;

- This questionnaire was sent during Q3, the quarter that is typically more problematic, and if the problems are more recent the magnitude is perceived in a different way, thus, customers may have been more punitive;

- One of the first questions concerned the customers' satisfaction, and it is believed this could have biased the other answers. Customers were particularly angry with prices and the fact this question appeared first, may have impacted the other answers.

It is also imperative to bear in mind that CS is one of the many consequences the CX has. Therefore, this comparison does not enable the validation of this quantitative measurement tool. Although, it is clear the relationship between these two metrics.

4.3.4.2 Presentation and visualization of results

As explained in Section 3.4, to present the results, dynamic and connected dashboards were created. Below, it is exhibited the dashboard for the global overview. As it is possible to see, the header contains the traffic light icons to indicate the overall current performance and, the current performance of each department. In the dashboard's body, buttons allow the user to select the periods he/she wants to analyze, being possible to see the historical evolution (on the left) and the status of each department (on the right). The footer was transformed into a navigation bar, so by clicking on the icons, the user will be directed to the dashboard of the corresponding department.

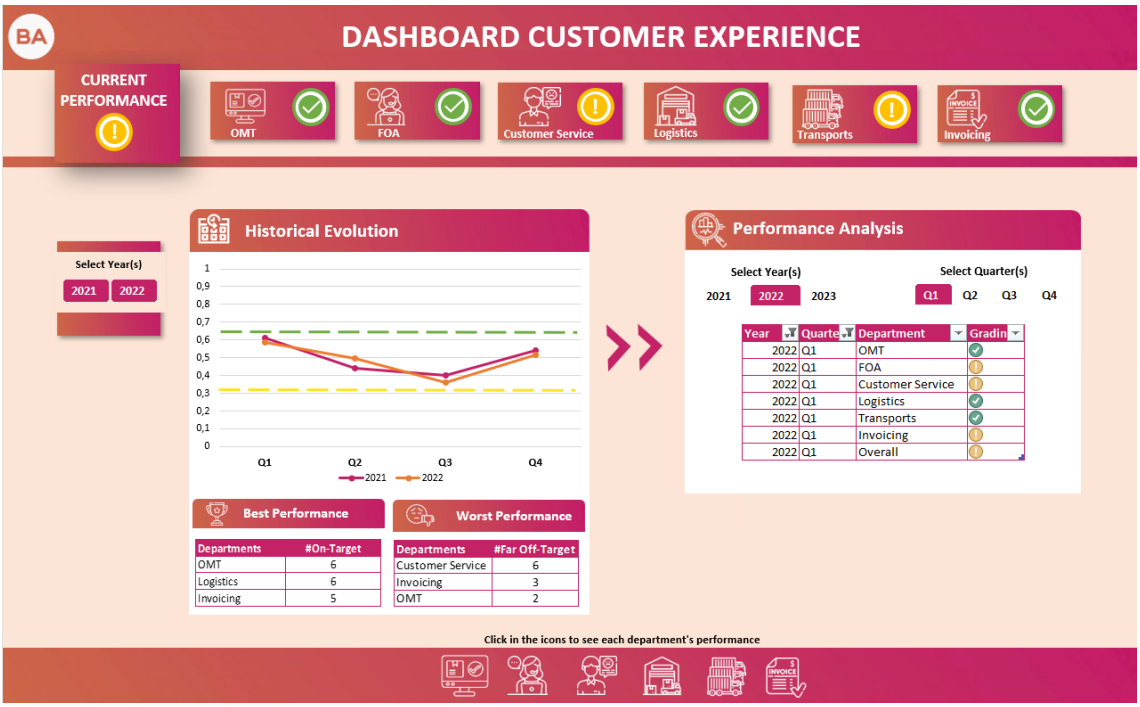


Figure 4.2: Dashboard - Customer experience

4.3.5 Interactions with decision-makers

In addition to the analysis and discussion of the results, it is also important to reflect upon the collection methodology followed. In general, both the participants of the preliminary meetings and the DMs considered the experience nice and pleasant, even without being familiar with MCDM methods. Although they expressed some difficulties at the beginning, they were easily overcome, and the implementation of the DCM-SRF was smooth and straightforward. More details are presented below.

1. **DCM-SRF-based questionnaires**

   – The questionnaires created to support preliminary meetings were very successful. In fact, the real-life example, followed by the illustrations, allowed participants to quickly understand the method and their role in that meeting, enabling them to promptly start discussing KPIs, and easily reach a consensus.

2. **Determination of KPIs' weights**

   – At first, DMs had some difficulty in evaluating KPIs from the customer's perspective (they were very hesitant);

   – Then, it was hard to rank KPIs since they were not being able to differentiate the impact (in their minds the differences between the activities' impact were really small);

   – At a certain point, they started feeling more comfortable and when they had to insert the blank cards at different intervals, DMs easily understood the idea behind those

differences. Though, there was some confusion concerning the difference between the insertion of 0 and/or 1 blank cards.

3. **Determination of KPIs' utility functions and targets**

   – For some DMs, providing the values of the two reference levels was hard since their sensitivity toward the KPIs is low. The same happened when defining the targets.

4. **Determination of departments' weights**

   – The sales directors felt more confident than the departments' managers when evaluating the departments' impact on CX, from the customer's perspective. Although they expressed the same confusion regarding the difference between the insertion of 0 and 1 blank cards;

   – As mentioned, the meeting was held online and DecSpace proved to be a useful tool for online decision-making meetings with multiple DMs.

4.4 Summary

This chapter presents the description of the implementation of the new approach to measure the B2B CX CI, with an emphasis on the calculations of weights, utility functions, and utilities. The obtained results are discussed in order not only to check the validity of the indicator but also to inspect possible areas for improvement. The next chapter encompasses the most relevant conclusions of this work, discusses its limitations, and proposes forthcoming considerations regarding the future of B2B CX CI.

Chapter 5

Conclusions and future remarks

5.1 Achievements

The main objective of this dissertation manifested in Chapter 1, was to develop a quantitative instrument to measure the CX, over time, in a B2B environment, more specifically, in the glass packaging industry. In order to accomplish this, a comprehensive analysis of existing literature was undertaken (Chapter 2). This made it possible to observe the lack of studies on B2B CX and, consequently, highlighted the importance and impact this study could have on both the company and the scientific community, not only by exploring the uncharted territory of B2B CX, but also by adopting a new and innovative approach to measure it. Additionally, the literature review allowed for the development of the B2B CX CI model, based on established methods, the DCM-SRF and MAUT, that were only applied in the construction of CI in other fields of study. The methodology developed was detailed in Chapter 3. In Chapter 4, this model was applied to a case study, being presented and analyzed the results of the aggregation of 30 KPIs, i.e. the results of the B2B CX CI, in different periods of time. In essence, it can be stated that the main objectives, outlined at the beginning of this project, were successfully achieved, given that:

– From the company's perspective:

  i. This quantitative instrument is the first metric company implements to measure the CX, and it allows BA to have a global view and a departmental view of the CX provided; and to monitor and manage CX, over time;

  ii. This work contributed to increasing the awareness of the BA employees involved in the study, regarding the importance of their actions on the CX; and to increasing inter-departmental visibility;

  iii. Considering the DMs, this project was crucial for them to realize how data collection and analysis can significantly impact their everyday activities. It is expected they start to monitor KPIs more closely, implement more KPIs, elaborate the best strategies to improve them, set more ambitious targets, etc.;

iv. The dynamic dashboards created make the visualization and analysis of the B2B CX CI's and departments' results easier.

– From the researchers' perspective:

i. This study, focused on an underexplored field of study, presented an alternative and innovative approach to measuring B2B CX;

ii. Within the glass packaging industry, to the author's knowledge, it was the first time CX was examined in detail, and a tailored metric to measure it was developed;

iii. Regarding the methods that were employed, it can be said that they worked well in an industrial and multicultural environment, where people were not familiar with MCDM methods.

5.2 Limitations

All studies have limitations, being crucial to acknowledge them and to perform a critical appraisal of their impact on the results (Price & Murnan, 2004). Next, limitations, in descending order of impact, are detailed, being provided the reasons for their existence, the impossibility of overcoming them, and their impact on the conclusions of this study.

– **Time** - The time to complete the literature review, apply the methodology, and gather and interpret the results was very short. As consequence:

i. Segmentation and geographical differences were excluded from the analysis. This decision had a significant impact on results since this CI does not allow BA to evaluate whether the current strategies for the different segments and geographies are maximizing the CX. However, to capture and measure the impact of those differences, utility functions, and targets should be tailored to each segment and geography which is highly complex and requires the analysis of a huge volume of data (since in BA, there are 7 segments and 4 geographical divisions), thus, being unfeasible in such a short time;

ii. Regarding the KPIs, in the beginning, it was planned that each department would collect and calculate them. However, this did not happen. As mentioned, departments only gathered data, and the analyst calculated 30 KPIs in 11 different periods. Thus, there was no time to analyze and statistically treat the data (namely, inspect missing data, verify the existence of outliers, calculate the average, standard deviation, maximum and minimum values, etc.), as should have been done. These actions would have enhanced the overall quality and validity of the findings;

iii. Dashboards could have been created in Microsoft Power BI or Tableau, which are considered better software for dashboard creation than Microsoft Excel. This was not the case due to the lack of time to learn how to use them.

- **Access to customers** - Assessing the impact of a particular activity and department on CX based on what the company's employees think of what the customer thinks is not the best approach. First, it is highly subjective and easily biased. Second, as highlighted, there is a gap between the company's and the customer's perceptions. However, since the beginning, the company established that it would not be possible to contact customers, and within the time available, it would have probably been impossible;

- **Lack of available data** - As mentioned in Section 3.1.1, some departments whose activities have a major impact on CX (such as Returnable Packaging, Planning, Production, and Pricing) were not included in this the study due to the lack of data. And, even the period analyzed 2021-2023 was shortened, because prior to 2021 there was very little data collected, and it was not possible to calculate most KPIs. This made it difficult to find a trend and a meaningful relationship between CX and CSI;

- **DMs sensitivity and subjectivity** - During the meetings with the departments' managers it was clear that some have low sensitivity regarding KPIs, being really difficult to establish targets, etc. Thus, some targets may not be the most appropriate. In addition to this, as highlighted in the literature review, participatory methods are highly subjective;

- **Departments' databases not connected** - Ideally, the CX CI would have been implemented in a platform that allowed a continuous update. However, it was not possible. Departments use different platforms to gather their data and they are not all connected nor all are instantly updated.

Despite some limitations, it should be highlighted this B2B CX CI is the first metric to measure CX in the glass packaging industry and the first implemented in the company, therefore, must be seen as a starting point for further analysis.

5.3 Recommendations and future work

Considering the results obtained, it is possible to identify two short-term actions that require little effort from the company and can have a positive and almost immediate impact on the CX. First, train, recognize, and develop reward policies for employees based on CX delivery; and second, improve communication between departments. Equally important, although requiring more effort from the company, is the development of strategies to mitigate fluctuations in the CX provided throughout the year.

In addition to these, and considering all the limitations observed throughout this project, duly recognized and identified in Section 5.2, is recommended to:

1. Validate the obtained results with customers, i.e. understand if the KPIs' and departments' weights, utility functions, and targets, reflect customers' expectations;

2. Include more KPIs in the CI, for example:

  - Include *perception-based metrics*, in line with Patti et al. (2020) suggestions;
  - Implement KPIs to measure the CX in the *pre-purchase stage*, and then include them in the CI, so it covers the entire CJ;
  - Implement the KPIs suggested in Table D.1, in Appendix D and include them in the CI [1].

3. Study segmentation and geographical differences, i.e. study how segmentation and geography affect customers' expectations and perceptions in relation to the weights of KPIs and departments, KPI's utility functions, and zones of tolerance. This may allow BA to develop more specific strategies that better suit their customers' needs and wants, resulting in an optimization of the use of resources;

4. Study the SRJ. Designing and monitoring the SRJ is crucial to turn negative CXs into positive ones, to rebuild customer trust, and to maintain long-term customer relationships, that characterize B2B customer-suppliers relationships. Although there are already processes implemented addressing the resolution of customer complaints, the monitoring of service recovery effectiveness is poor. To improve it, more KPIs should be implemented and a dashboard to visualize the complaints history of customers should be created;

5. Connect the databases of all departments in a platform that can be continuously updated and that allows the implementation of the CI.

Furthermore, is recommended the creation of a CI focused on measuring the company's performance. This is extremely important since it is crucial to balance CX improvements with the impact of those strategies on the company's performance. For example, customers value the ability of the company to deliver on short notice, but for the company, it is not good since it requires extra effort. Another example is production. The company wants to increase its production efficiency but the customer wants to have stock when needed. Therefore, it is expected that with the two measures, the trade-offs can be easily measured and managed. Moreover, it would also be interesting to examine the following topics in the future research:

1. **Customers' evolution** - As evidenced in the literature review, customers' needs and expectations change over time, as well as their zone of tolerance. Thus, it would be interesting to study, for example, the degree of impact of certain disruptions on customers' expectations. Or, how being a sustainable supplier is impacting/ will impact and to what extent the customers' choices and CXs. This may be of particular interest to BA, since it was recognized as "The Most Sustainable Company in the Glass Industry" [2], by World Finance, and it is investing a lot in research and development initiatives to reduce its environmental footprint[3]. Additionally, it would be interesting to analyze how much customers value their suppliers' innovations and proactivity, and if it changes over time;

[1] The Table D.1 shows only some of the KPIs recommended to the company.
[2] See `https://www.worldfinance.com/sustainability-awards`
[3] See `https://www.baglass.com/en/history.php`

2. **Develop a text mining model** - Text mining is a valuable approach for managing CX due to its ability to extract meaningful insights from large volumes of textual data. A text mining model may help businesses gain a deeper understanding of customer perceptions, preferences, and satisfaction levels, and can guide decision-making processes and aid in improving products, services, and overall CX. Moreover, by analyzing customer interactions, feedback, and preferences, businesses can deliver personalized experiences, targeted recommendations, and customized offers. In their article, McColl-Kennedy et al. (2019) provide a thorough explanation of how a linguistics-based text mining model can be developed;

3. **Chatbots** - Study the impacts of chatbots' implementation in a B2B context. On the one hand, chatbots can be highly beneficial, contributing to enhancing the overall CX, since they provide instant and efficient customer support; consistent and accurate information, by reducing the risk of human errors or inconsistencies; and more advanced chatbots can even utilize artificial intelligence algorithms to analyze customer data and provide personalized recommendations or solutions; etc. On the other hand, B2B relationships thrive on personal connections and human interactions, that for now, chatbots cannot simulate; B2B interactions often involve sensitive and confidential information, thus, there might be some data privacy and security concerns; and, due to the high complexity of B2B environments, chatbots may not be always capable of providing a solution. Therefore, assessing where and how chatbots can be effectively implemented in a B2B context, should be further examined;

4. **DCM-SRF-based questionnaires and the Delphi method** - Study whether the application of the Delphi method using the questionnaires based on DCM-SRF would be effective in a CX CI construction.

In summary, it can be concluded that the main objectives of this project have been successfully accomplished. Based on the obtained results, two short-term actions are identified and suggested. Additionally, it is recommended to validate the results with customers, include more KPIs in the CX measurement, study segmentation, and geographical differences, analyze SRJs, and connect departmental databases in a platform for continuous updating and monitoring. Furthermore, creating a performance-focused CI is advised to balance CX improvements with their impact on company performance, enabling better measurement and management of trade-offs. Future research suggestions also include studying the evolution of customers, developing a text mining model for CX analysis, assessing the implementation of chatbots in B2B contexts, and exploring the effectiveness of the Delphi method in CI construction using DCM-SRF-based questionnaires.

References

Andreassen, T. W., & Lindestad, B. (1998). Customer loyalty and complex services. the impact of corporate image on quality, customer satisfaction and loyalty for customers with varying degrees of service expertise. *International Journal of Service Industry Management*, *9*(1), 7-23. doi: 10.1108/09564239810199923

Antanasijević, D., Pocajt, V., Ristić, M., & Perić-Grujić, A. (2017). A differential multi-criteria analysis for the assessment of sustainability performance of european countries: Beyond country ranking. *Journal of Cleaner Production*, *165*, 213-220. doi: 10.1016/j.jclepro.2017.07.131

Attardi, R., Cerreta, M., Sannicandro, V., & Torre, C. M. (2018). Non-compensatory composite indicators for the evaluation of urban planning policy: The land-use policy efficiency index (lupei). *European Journal of Operational Research*, *264*(2), 491-507. doi: 10.1016/j.ejor.2017.07.064

Azapagic, A., & Perdan, S. (2005). An integrated sustainability decision-support framework part ii: Problem analysis. *International Journal of Sustainable Development and World Ecology*, *12*, 112-131. doi: 10.1080/13504500509469623

Baines, T. S., Lightfoot, H. W., Evans, S., Neely, A., Greenough, R., Peppard, J., ... Wilson, H. (2007). State-of-the-art in product-service systems. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, *221*(10), 1543-1552. doi: 10.1243/09544054JEM858

Becker, L., & Jaakkola, E. (2020). Customer experience: fundamental premises and implications for research. *Journal of the Academy of Marketing Science*, *48*, 630-648. doi: 10.1007/s11747-019-00718-x

Belton, V., & Stewart, T. J. (2002). Value function methods: Practical basics. In *Multiple criteria decision analysis: An integrated approach* (p. 119-161). Boston, MA: Springer US. doi: 10.1007/978-1-4615-1495-4_5

Bitner, M. J., Ostrom, A. L., & Morgan, F. N. (2008). Service blueprinting: A practical technique for service innovation. *California Management Review*, *50*(3), 66-94. doi: 10.2307/41166446

Boejgaard, J., & Ellegaard, C. (2010). Unfolding implementation in industrial market segmentation. *Industrial Marketing Management*, *39*(8), 1291-1299. doi: 10.1016/j.indmarman.2010.06.004

Bonoma, T. V., & Shapiro, B. P. (1983). *Segmenting the industrial market*. Lexington Books.

Bottero, M., Ferretti, V., Figueira, J. R., Greco, S., & Roy, B. (2018). On the Choquet multiple criteria preference aggregation model: Theoretical and practical insights from a real-world application. *European Journal of Operational Research*, *271*(1), 120-140. doi: 10.1016/j.ejor.2018.04.022

Brans, J. P., Mareschal, B., Figueira, J., Greco, S., & Ehrogott, M. (2005). PROMETHEE Methods. In *Multiple criteria decision analysis: State of the art surveys* (p. 163-186). New York:

Springer New York. doi: 10.1007/0-387-23081-5_5

Brans, J. P., & Vincke, P. (1985). A Preference Ranking Organisation Method: (The PROMETHEE method for multiple criteria decision-making). *Management Science*, *31*(6), 647-656. doi: 10.1287/mnsc.31.6.647

Breebaart, K. (2012). *CEM – The Service Recovery Paradox.* https://blogs.sap.com/2012/10/16/cem-the-service-recovery-paradox/. (Accessed on 13/03/2022)

Bruhn, M., & Grund, M. (2000). Theory, development and implementation of national customer satisfaction indices: The Swiss Index of Customer Satisfaction (SWICS). *Total Quality Management & Business Excellence*, *11*, 1017-1028. doi: 10.1080/09544120050135542

Charnes, A., Cooper, W., & Rhodes, E. (1978). Measuring the efficiency of decision-making units. *European Journal of Operational Research*, *2*(6), 429-444. doi: 10.1016/0377-2217(78)90138-8

Cherchye, L., & Kuosmanen, T. (2004). *Benchmarking Sustainable Development: A Synthetic Meta-index Approach* (Tech. Rep. No. RP2004-28). World Institute for Development Economic Research. Retrieved from https://ideas.repec.org/p/unu/wpaper/rp2004-28.html

Corrente, S., Figueira, J., & Greco, S. (2021). Pairwise comparison tables within the deck of cards method in multiple criteria decision aiding. *European Journal of Operational Research*, *291*(2), 738-756. doi: 10.1016/j.ejor.2020.09.036

Cortez, R. M., Clarke, A. H., & Freytag, P. V. (2021). B2B market segmentation: A systematic review and research agenda. *Journal of Business Research*, *126*, 415-428. doi: 10.1016/J.JBUSRES.2020.12.070

Cracolici, M. F., & Nijkamp, P. (2009). The attractiveness and competitiveness of tourist destinations: A study of southern italian regions. *Tourism Management*, *30*, 336-344. doi: 10.1016/j.tourman.2008.07.006

Cronin. Jr, J., & Taylor, S. (1992). Measuring service quality - a reexamination and extension. *The Journal of Marketing*, *56*, 55-68. doi: 10.2307/1252296

Dantsis, T., Douma, C., Giourga, C., Loumou, A., & Polychronaki, E. A. (2010). A methodological approach to assess and compare the sustainability level of agricultural plant production systems. *Ecological Indicators*, *10*(2), 256-263. doi: 10.1016/j.ecolind.2009.05.007

Davey, N. (2017). *Why 2017 has been a bad year for customer experience management.* https://www.mycustomer.com/experience/engagement/why-2017-has-been-a-bad-year-for-customer-experience-management. (Accessed on 10/03/2022)

Dawson, L., & Schlyter, P. (2012). Less is more: Strategic scale site suitability for concentrated solar thermal power in Western Australia. *Energy Policy*, *47*, 91-101. doi: 10.1016/j.enpol.2012.04.025

Decancq, K., & Lugo, M. A. (2013). Weights In Multidimensional Indices Of Wellbeing: An Overview. *Econometric Reviews*, *32*(1), 7-34. doi: 10.1080/07474938.2012.690641

Diaz-Balteiro, L., González-Pachón, J., & Romero, C. (2017). Measuring systems sustainability with multi-criteria methods: A critical review. *European Journal of Operational Research*, *258*(2), 607-616. doi: 10.1016/j.ejor.2016.08.075

Dibb, S., & Simkin, L. (2008). *Market segmentation success: Making it happen!* London: Routledge.

ECSI Technical Committee. (1998). *European customer satisfaction index: foundation and structure for harmonized national pilot projects* (Tech. Rep.). Report prepared for the ECSI Steering Committee, October.

Figueira, J., & Roy, B. (2002). Determining the weights of criteria in the electre type methods with

a revised simos' procedure. *European Journal of Operational Research*, *139*(2), 317-326. doi: 10.1016/S0377-2217(01)00370-8

Figueira, J. R., Greco, S., & Roy, B. (2022). Electre-score: A first outranking based method for scoring actions. *European Journal of Operational Research*, *297*(3), 986-1005. doi: 10.1016/j.ejor.2021.05.017

Figueira, J. R., Oliveira, H. M., Serro, A. P., Colaço, R., Froes, F., Robalo Cordeiro, C., ... Guimarães, M. (2023). A multiple criteria approach for building a pandemic impact assessment composite indicator: The case of covid-19 in portugal. *European Journal of Operational Research*, *309*(2), 795-818. doi: 10.1016/j.ejor.2023.01.025

Flanagan, P., Johnston, R., & Talbot, D. (2005). Customer confidence: The development of a "pre-experience" concept. *International Journal of Service Industry Management*, *16*, 373-384. doi: 10.1108/09564230510614013

Fornell, C. (1992). A national customer satisfaction barometer: The Swedish experience. *Journal of Marketing*, *56*(1), 6-21. doi: 10.2307/1252129

Fornell, C., Johnson, M. D., Anderson, E. W., Cha, J., & Bryant, B. E. (1996). The American customer satisfaction index: nature, purpose, and findings. *Journal of marketing*, *60*(4), 7-18. doi: 10.2307/1251898

Freytag, P. V., & Clarke, A. H. (2001). Business to business market segmentation. *Industrial Marketing Management*, *30*, 473-486. doi: 10.1016/S0019-8501(99)00103-0

Gibari, S. E., Gómez, T., & Ruiz, F. (2019). Building composite indicators using multicriteria methods: a review. *Journal of Business Economics*, *89*, 1-24. doi: 10.1007/s11573-018-0902-z

Gibbons, S. (2017). *Service blueprints: Definition.* https://www.nngroup.com/articles/service-blueprints-definition/. (Accessed on 30/03/2023)

Gilmore, J., & Pine II, B. (2002). Customer experience places: The new offering frontier. *Strategy & Leadership*, *30*, 4-11. doi: 10.1108/10878570210435306

Greco, S., Ishizaka, A., Tasiou, M., & Torrisi, G. (2019). On the methodological framework of composite indices: A review of the issues of weighting, aggregation, and robustness. *Social Indicators Research*, *141*, 61-94. doi: 10.1007/s11205-017-1832-9

Green, P. E., & Rao, V. R. (1971). Conjoint measurement for quantifying judgmental data. *Journal of Marketing Research*, *8*(3), 355–363. doi: 10.2307/3149575

Gwartney, J., Lawson, R., & Block, W. (1996). Economic freedom of the world: 1975-1995. *BC Canada: The Fraser Institute*.

Hajkowicz, S. (2006). Multi-attributed environmental index construction. *Ecological Economics*, *57*(1), 122-139. doi: 10.1016/j.ecolecon.2005.03.023

Håkansson, H., Ford, D., Gadde, L., Snehota, I., & Waluszewski, A. (2009). *Business in networks.* John Wiley & Sons Ltd.

Halvorsrud, R., Kvale, K., & Følstad, A. (2016). Improving service quality through customer journey analysis. *Journal of Service Theory and Practice*, *26*, 840-867. doi: 10.1108/JSTP-05-2015-0111

Hernandez-Perdomo, E., Mun, J., & Rocco, C. (2017). Active management in state-owned energy companies: Integrating a real options approach into multicriteria analysis to make companies sustainable. *Applied Energy*, *195*, 487-502. doi: 10.1016/j.apenergy.2017.03.068

Herssens, C., Jureta, I. J., & Faulkner, S. (2008). Dealing with quality tradeoffs during service selection. In *2008 International Conference on Autonomic Computing* (p. 77-86). doi: 10.1109/ICAC.2008.8

Hitt, Ireland, & Hoskisson. (2016). *Strategic management: Competitiveness & Globalization: Concepts and Cases* (12th ed.). Cengage Learning.

Hoffman, D. L., & Novak, T. P. (2018). Consumer and object experience in the Internet of Things: An assemblage theory approach. *Journal of Consumer Research*, *44*(6), 1178-1204. doi: 10.1093/jcr/ucx105

Homburg, C., & Rudolph, B. (2001). Customer Satisfaction in Industrial Markets: Dimensional and multiple role issues. *Journal of Business Research*, *52*, 15-33. doi: 10.1016/S0148-2963(99)00101-0

Jansen, S. J. T. (2011). The Multi-attribute Utility Method. In *The measurement and analysis of housing preference and choice* (p. 101-125). Dordrecht: Springer Netherlands. doi: 10.1007/978-90-481-8894-9_5

Jayawardhena, C., Souchon, A. L., Farrell, A. M., & Glanville, K. (2007). Outcomes of service encounter quality in a business-to-business context. *Industrial Marketing Management*, *36*(5), 575-588. doi: 10.1016/j.indmarman.2006.02.012

Johnston, R. (1995). The zone of tolerance. *International Journal of Service Industry Management*, *6*, 46-61. doi: 10.1108/09564239510084941

Johnston, T. C., & Hewa, M. A. (1997). Fixing service failures. *Industrial Marketing Management*, *26*(5), 467-473. doi: 10.1016/S0019-8501(96)00158-7

Keeney, R. L., & Raiffa, H. (1976). *Decisions with multiple objectives: preferences and value trade-offs*. Cambridge university press.

Keiningham, T. L., Ball, J., (née Moeller), S. B., Bruce, H. L., Buoye, A., Dzenkovska, J., ... Zaki, M. (2017). The interplay of customer experience and commitment. *Journal of Services Marketing*, *31*(2), 148-160. doi: 10.1108/JSM-09-2016-0337

Keiningham, T. L., Cooil, B., Andreassen, T. W., & Aksoy, L. (2007). A longitudinal examination of net promoter and firm revenue growth. *Journal of Marketing*, *71*(3), 39-51. doi: 10.1509/jmkg.71.3.039

Keyser, A. D., Lemon, K. N., Keiningham, T. L., & Klaus, P. (2015). A framework for understanding and managing the customer experience. *Marketing Science Institute working paper series*.

Keyser, A. D., Verleye, K., Lemon, K. N., Keiningham, T. L., & Klaus, P. (2020). Moving the Customer Experience Field Forward: Introducing the Touchpoints, Context, Qualities (TCQ) Nomenclature. *Journal of Service Research*, *23*(4), 433 - 455. doi: 10.1177/1094670520928390

Klaus, P. (2015). Customer experience: The origins and importance for your business. In *Measuring customer experience: How to develop and execute the most profitable customer experience strategies* (p. 1-21). London: Palgrave Macmillan UK. doi: 10.1057/9781137375469_1

Kodikara, P., Perera, B., & Kularathna, M. (2010). Stakeholder preference elicitation and modelling in multi-criteria decision analysis: A case study on urban water supply. *European Journal of Operational Research*, *206*(1), 209-220. doi: 10.1016/j.ejor.2010.02.016

Korhonen, P., Tainio, R., & Wallenius, J. (2001). Value efficiency analysis of academic research. *European Journal of Operational Research*, *130*(1), 121-132. doi: 10.1016/S0377-2217(00)00050-3

Kuppelwieser, V. G., & Klaus, P. (2021). Measuring customer experience quality: The EXQ scale revisited. *Journal of Business Research*, *126*, 624-633. doi: 10.1016/j.jbusres.2020.01.042

LaPlaca, P. J. (1997). Contributions to marketing theory and practice from industrial marketing management. *Journal of Business Research*, *38*, 179-198. doi: 10.1016/S0148-2963(96)00128-2

Lemon, K. N., & Verhoef, P. C. (2016). Understanding Customer Experience Throughout the Customer Journey. *Journal of Marketing*, *80*(6), 69 - 96. doi: 10.1509/jm.15.0420

Liljander, V., & Strandvik, T. (1997). Emotions in service satisfaction. *International Journal of Service Industry Management*, *8*(2), 148-169. doi: 10.1108/09564239710166272

Lin, S., Shen, L., Xiong, C., & Li, X. (2020). Multi-criteria group decision making and group agreement quotient analysis based on the delphi method. In *Complex, intelligent, and software intensive systems: Proceedings of the 13th international conference on complex, intelligent, and software intensive systems (cisis-2019)* (pp. 237–246).

Lundin, L., & Kindström, D. (2023). Digitalizing customer journeys in B2B markets. *Journal of Business Research*, *157*. doi: 10.1016/j.jbusres.2022.113639

Mahlberg, B., & Obersteiner, M. (2001). Remeasuring the HDI by Data Envelopment Analysis. doi: 10.2139/ssrn.1999372

Maklan, D. S., & Klaus, P. (2011). Customer experience: Are we measuring the right things? *International Journal of Market Research*, *53*(6), 771-772. doi: 10.2501/IJMR-53-6-771-792

Malkina-Pykh, I. G., & Pykh, Y. A. (2008). Quality-of-life indicators at different scales: Theoretical background. *Ecological Indicators*, *8*(6), 854-862. doi: 10.1016/j.ecolind.2007.01.008

Mascarenhas, O., Kesavan, R., & Bernacchi, M. (2006). Lasting customer loyalty: A total customer experience approach. *Journal of Consumer Marketing*, *23*, 397-405. doi: 10.1108/07363760610712939

McColl-Kennedy, J. R., Gustafsson, A., Jaakkola, E., Klaus, P., Radnor, Z. J., Perks, H., & Friman, M. (2015). Fresh perspectives on customer experience. *Journal of Services Marketing*, *29*, 430-435. doi: 10.1108/JSM-01-2015-0054

McColl-Kennedy, J. R., Zaki, M., Lemon, K. N., Urmetzer, F., & Neely, A. (2019). Gaining customer experience insights that matter. *Journal of Service Research*, *22*(1), 8-26. doi: 10.1177/1094670518812182

Melyn, W., & Moesen, W. (1991). Towards a synthetic indicator of macroeconomic performance: Unequal weighting when limited information is available, public economics research paper 17. *Center for Economic Studies, KU Leuven*.

Meyer, C., & Schwager, A. (2007). Understanding customer experience. *Harvard Business Review*, *85*, 116-157. Retrieved from `https://hbr.org/2007/02/understanding-customer-experience`

Michel, S., & Meuter, M. L. (2008). The service recovery paradox: true but overrated? *International Journal of Service Industry Management*, *19*(4), 441-457. doi: 10.1108/09564230810891897

Mikolon, S., Kolberg, A., Haumann, T., & Wieseke, J. (2015). The complex role of complexity: How service providers can mitigate negative effects of perceived service complexity when selling professional services. *Journal of Service Research*, *18*(4), 513-528. doi: 10.1177/1094670514568778

Mitchell, V. W., & Wilson, D. F. (1998). Balancing theory and practice: A reappraisal of business-to-business segmentation. *Industrial Marketing Management*, *27*, 429-445. doi: 10.1016/S0019-8501(98)00002-9

Munda, G. (2005). Multiple criteria decision analysis and sustainable development. In *Multiple criteria decision analysis: State of the art surveys* (p. 953-986). New York, NY: Springer New York. doi: 10.1007/0-387-23081-5_23

Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hoffmann, A., & & Giovannini, E. (2008). *Handbook on constructing composite indicators: Methodology and user guide*. Paris (France): OECD publishing. Retrieved from `www.oecd.org/publishing,http://213.253.134.43/oecd/pdfs/browseit/3008251E.PDF`

Naudé, P., & Buttle, F. (2000). Assessing relationship quality. *Industrial Marketing Management*, *29*, 351-361. doi: 10.1016/S0019-8501(00)00112-7

Nicoletti, G., Scarpetta, S., & Boylaud, O. (2000). Summary indicators of product market regulation with an extension to employment protection legislation. *OECD Economics Department Working Papers*(226). doi: 10.1787/215182844604

Palmatier, R. W., & Crecelius, A. T. (2019). The "first principles" of marketing strategy. *Academy of Marketing Science Review*, *9*, 5-26. doi: 10.1007/s13162-019-00134-y

Palmer, R. A., & Millier, P. (2004). Segmentation: Identification, intuition, and implementation. *Industrial Marketing Management*, *33*(8), 779-785. doi: 10.1016/j.indmarman.2003.10.007

Parasuraman, A. (1998). Customer service in business-to-business markets: an agenda for research. *Journal of Business & Industrial Marketing*, *13*, 309-321. doi: 10.1108/08858629810226636

Parasuraman, A., Berry, L. L., & Zeithaml, V. A. (1991). Understanding customer expectations of service. *MIT Sloan Management Review*.

Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1985). A conceptual model of service quality and its implications for future research. *Journal of Marketing*, *49*(4), 41-50. doi: 10.2307/1251430

Patrício, L., Fisk, R. P., & e Cunha, J. F. (2008). Designing multi-interface service experiences: The service experience blueprint. *Journal of Service Research*, *10*(4), 318-334. doi: 10.1177/1094670508314264

Patti, C. H., van Dessel, M. M., & Hartley, S. W. (2020). Reimagining customer service through journey mapping and measurement. *European Journal of Marketing*, *54*, 2387-2417. doi: 10.1108/EJM-07-2019-0556

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge University Press. doi: 10.1017/CBO9781139173933

Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *2*(11), 559-572. doi: 10.1080/14786440109462720

Pennoni, F., Tarantola, S., & Latvala, A. (2005). The European e-business readiness index. *Joint Research Centre (2003-2007)*.

Pereira, M. A., Figueira, J. R., & Marques, R. C. (2020). Using a Choquet integral-based approach for incorporating decision-maker's preference judgments in a data envelopment analysis model. *European Journal of Operational Research*, *284*(3), 1016-1030. doi: 10.1016/j.ejor.2020.01.037

Petrović, M., Bojković, N., Anić, I., Stamenković, M., & Tarle, S. P. (2014). An ELECTRE-based decision aid tool for stepwise benchmarking: An application over EU Digital Agenda Targets. *Decision Support Systems*, *59*, 230-241. doi: 10.1016/j.dss.2013.12.002

Pictet, J., & Bollinger, D. (2008). Extended use of the cards procedure as a simple elicitation technique for MAVT. application to public procurement in Switzerland. *European Journal of Operational Research*, *185*(3), 1300-1307. doi: 10.1016/j.ejor.2006.05.051

Price, J., & Murnan, J. (2004). Research limitations and the necessity of reporting them. *American Journal of Health Education*, *35*, 66-67. doi: 10.1080/19325037.2004.10603611

Puccinelli, N., Goodstein, R., Grewal, D., Price, R., Raghubir, P., & Stewart, D. (2009). Customer experience management in retailing: Understanding the buying process. *Journal of Retailing*, *85*, 15-30. doi: 10.1016/j.jretai.2008.11.003

Pullman, M., & Gross, M. (2004). Ability of experience design elements to elicit emotions and loyalty behaviors. *Decision Sciences*, *35*, 551 - 578. doi: 10.1111/j.0011-7315.2004.02611

.x

Puolamaa, M., Kaplas, M., & Reinikainen, T. (1996). *Index of environmental friendliness: a methodological study*. Statistics Finland.

Rawson, A., Duncan, E., & Jones, C. (2013). The truth about customer experience. *Harvard Business Review*, *91*, 90-98. Retrieved from `https://hbr.org/2013/09/the-truth-about-customer-experience`

Ray, A. K. (2008). Measurement of social development: an international comparison. *Social Indicators Research*, *86*(1), 1-46. doi: 10.1007/s11205-007-9097-3

Reichheld, F. F. (2003). The one number you need to grow. *Harvard Business Review*, *81*(12), 46-55.

Reichheld, F. F., Hamilton, B., Markey, R., & Allen, J. (2005). *Closing the delivery gap* (Tech. Rep.). London: Bain & Company. Retrieved from `https://media.bain.com/bainweb/PDFs/cms/hotTopics/closingdeliverygap.pdf`

Rossiter, J. R. (2017). Optimal standard measures for marketing. *Journal of Marketing Management*, *33*(5-6), 313-326. doi: 10.1080/0267257X.2017.1293710

Roy, B. (1968). Classement et choix en présence de points de vue multiples (la méthode ELECTRE). *Revue française d'informatique et de recherche opérationnelle*, *2*(8), 57-75. doi: 10.1051/ro/196802V100571

Roy, B. (1991). The outranking approach and the foundations of ELECTRE methods. *Theory and Decision*, *31*, 49-73. doi: 10.1007/BF00134132

Roy, S., Sreejesh, S., & Bhatia, S. (2019). Service quality versus Service experience: An empirical examination of the consequential effects in B2B services. *Industrial Marketing Management*, *82*, 52-69. doi: 10.1016/j.indmarman.2019.02.017

Saaty, T. (1980). *The analytic hierarchy process (AHP) for decision making*. McGraw-Hill, New York.

Saisana, M., & Tarantola, S. (2002). *State-of-the-art report on current methodologies and practices for composite indicator development* (Vol. 214). European Commission, Joint Research Centre. doi: 10.13140/RG.2.1.1505.1762

Schmiedeknecht, M. H. (2013). Environmental sustainability index. In *Encyclopedia of corporate social responsibility* (p. 1017-1024). Springer Berlin Heidelberg. doi: 10.1007/978-3-642-28036-8_116

Schmitt, B. (1999). Experiential marketing. *Journal of Marketing Management*, *15*(1-3), 53-67. doi: 10.1362/026725799784870496

Shanian, A., Milani, A., Carson, C., & Abeyaratne, R. (2008). A new application of electre iii and revised simos' procedure for group material selection under weighting uncertainty. *Knowledge-Based Systems*, *21*(7), 709-720. doi: 10.1016/j.knosys.2008.03.028

Simos, J. (1989). *L'évaluation environnementale: Un processus cognitif négocié* (Doctoral dissertation, École Polytechnique Fédérale de Lausanne, Lausanne). doi: 10.5075/epfl-thesis-823

Singh, R. K., Murty, H. R., Gupta, S. K., & Dikshit, A. K. (2007). Development of composite sustainability performance index for steel industry. *Ecological Indicators*, *7*(3), 565-588. doi: 10.1016/j.ecolind.2006.06.004

Siskos, E., & Tsotsolas, N. (2015). Elicitation of criteria importance weights through the simos method: A robustness concern. *European Journal of Operational Research*, *246*(2), 543-553. doi: 10.1016/j.ejor.2015.04.037

Spreng, R. A., Hui Shi, L., & Page, T. J. (2009). Service quality and satisfaction in business-to-business services. *Journal of Business & Industrial Marketing*, *24*(8), 537-548. doi: 10.1108/08858620910999411

Szmigin, I. T. D. (1993). Managing quality in business-to-business services. *European Journal of Marketing*, *27*, 2-21. doi: 10.1108/03090569310024512

Tarantola, S., Liska, R., Saltelli, A., Leapman, N., & Grant, C. (2004). *The internal market index 2004* (Tech. Rep. No. JRC28435). Italy: European Commission. Retrieved from `https://publications.jrc.ec.europa.eu/repository/handle/JRC28435`

Thompson, B. (2018). *An inconvenient truth: 93% of customer experience initiatives are failing.* `https://customerthink.com/an-inconvenient-truth-93-of-customer-experience-initiatives-are-failing/`.

Tomizawa, H., & Niwa, F. (1996). Evaluating overall national science and technology activity: General Indicator Of Science and Technology (GIST) and its implications for S&T policy. *Research Evaluation*, *6*(2), 83-92. doi: 10.1093/rev/6.2.83

Tueanrat, Y., Papagiannidis, S., & Alamanos, E. (2021). Going on a journey: A review of the customer journey literature. *Journal of Business Research*, *125*, 336-353. doi: 10.1016/j.jbusres.2020.12.028

Ulaga, W., & Eggert, A. (2006). Relationship value and relationship quality. *European Journal of Marketing*, *40*, 311-327. doi: 10.1108/03090560610648075

Ülengin, B., Ülengin, F., & Güvenç, Ü. (2001). A multidimensional approach to urban quality of life: The case of Istanbul. *European Journal of Operational Research*, *130*(2), 361-374. doi: 10.1016/S0377-2217(00)00047-3

Vaerenbergh, Y. V., Varga, D., Keyser, A. D., & Orsingher, C. (2019). The service recovery journey: Conceptualization, integration, and directions for future research. *Journal of Service Research*, *22*(2), 103-119. doi: 10.1177/1094670518819852

van Calker, K., Berentsen, P., Romero, C., Giesen, G., & Huirne, R. (2006). Development and application of a multi-attribute sustainability function for dutch dairy farming systems. *Ecological Economics*, *57*(4), 640-658. doi: 10.1016/j.ecolecon.2005.05.016

Verhoef, P., Lemon, K., Parasuraman, A. P., Roggeveen, A., Tsiros, M., & Schlesinger, L. (2009). Customer experience creation: Determinants, dynamics and management strategies. *Journal of Retailing*, *85*, 31-41. doi: 10.1016/j.jretai.2008.11.001

Von Winterfeldt, D., & Edwards, W. (1993). Decision analysis and behavioral research. In (p. 275). Cambridge, MA (USA) Cambridge Univ. Press.

Witell, L., Kowalkowski, C., Perks, H., Raddats, C., Schwabe, M., Benedettini, O., & Burton, J. (2020). Characterizing customer experience management in business markets. *Journal of Business Research*, *116*, 420-430. doi: 10.1016/J.JBUSRES.2019.08.050

Wu, Y., Zhang, J., Yuan, J., Geng, S., & Zhang, H. (2016). Study of decision framework of offshore wind power station site selection based on ELECTRE-iii under intuitionistic fuzzy environment: A case of China. *Energy Conversion and Management*, *113*, 66-81. doi: 10.1016/j.enconman.2016.01.020

Yu, T., & Dean, A. (2001). The contribution of emotional satisfaction to consumer loyalty. *International Journal of Service Industry Management*, *12*, 234-250. doi: 10.1108/09564230110393239

Zeithaml, V. A., Bitner, M. J., & Gremler, D. D. (2006). *Services marketing: Integrating customer focus across the firm*. McGraw-Hill/Irwin.

Zhu, X., & Zolkiewski, J. (2015). Exploring service failure in a business-to-business context. *Journal of Services Marketing*, *29*(5), 367–379. doi: 10.1108/JSM-02-2014-0055

Zolkiewski, J., Story, V., Burton, J., Chan, P., de Souza, A., Hunter-Jones, P., ... Robinson, W. (2017). Strategic B2B customer experience management: the importance of outcomes-based measures. *Journal of Services Marketing*, *31*, 172-184. doi: 10.1108/JSM-10-2016-0350

Appendix A

Literature Review

A.1 Customer experience

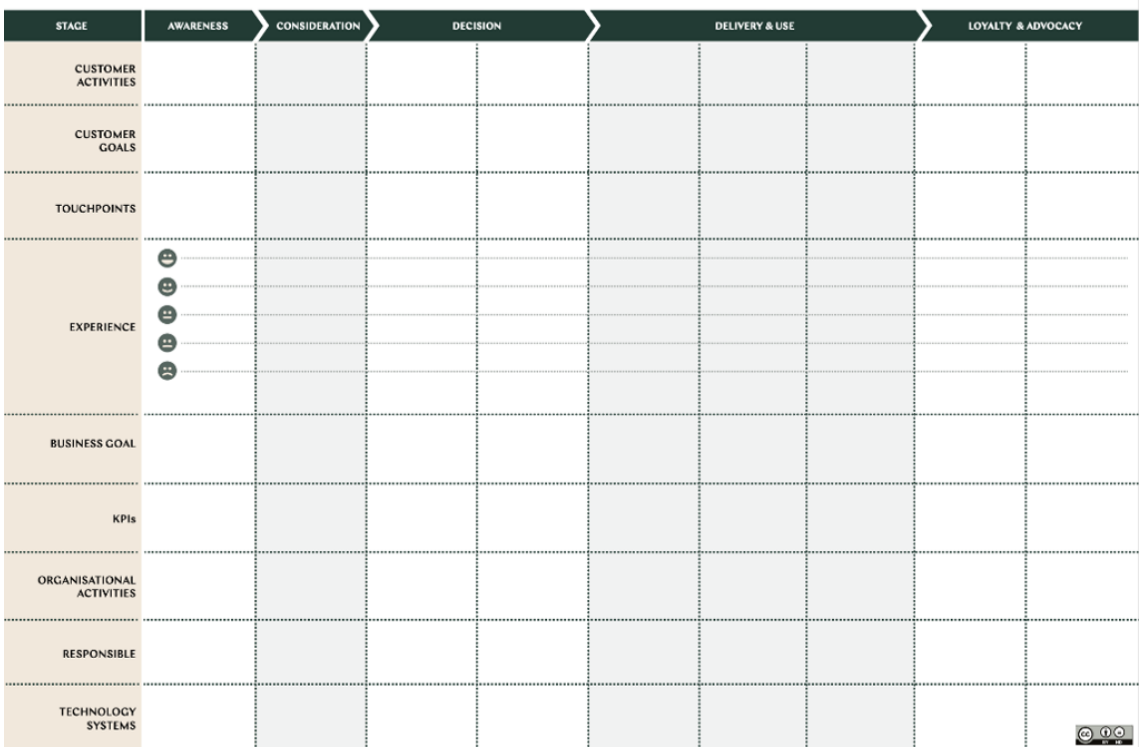
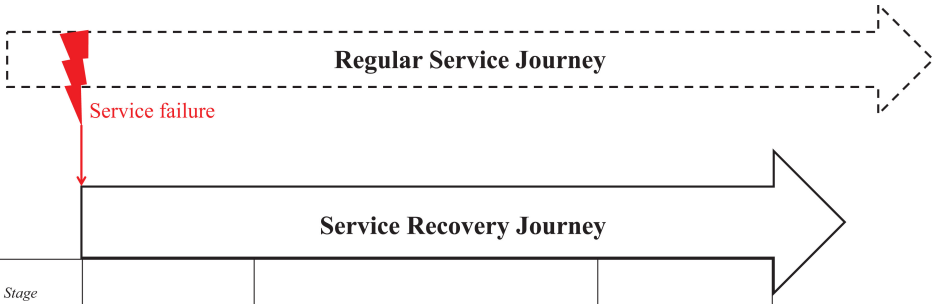


Figure A.1: Customer journey example [a]

[a] Available at https://media.cheggcdn.com/media/d51/d513a8ac-64c2-42d7-a063-9b92bb2422a1/phptcaTNp



Stage / *Type of organizational response*	**Pre-Recovery**	**Recovery**	**Post-Recovery**
Compensation	-	• Immediate monetary compensation • Delayed monetary compensation • New/exchanged goods • New/reperformed service • Apology	-
Favorable employee behavior	-	• Excuse • Justification • Referential account • Credibility feedback • Courtesy • Effort • Empathy • Willingness to listen	-
Organizational procedures	• Facilitation • Initiation	• Customer participation • Employee empowerment • Flexibility • Recovery time	• Follow-up • Process recovery communication

Figure A.2: Service recovery journey. Source: Vaerenbergh et al. (2019).

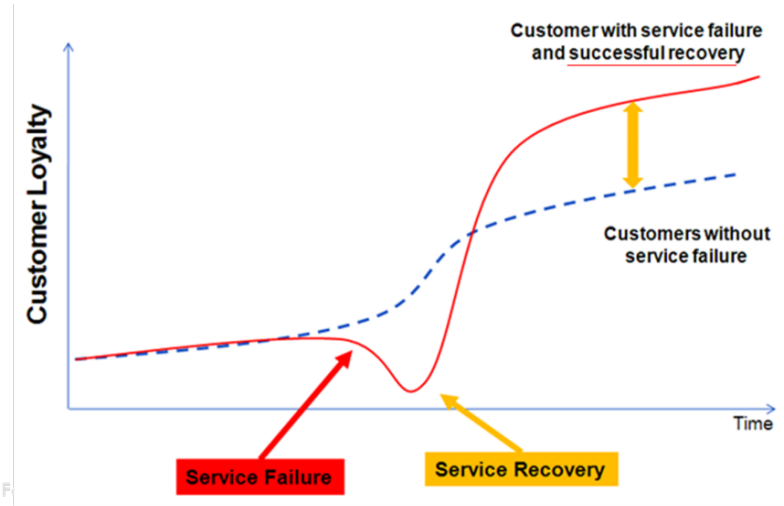


Figure A.3: Service paradox diagram. Source: Breebaart (2012)

A.2 Composite indicators

Table A.1: Examples of weighting methods

Method	Research Paper	CI or Application Field
PCA	Schmiedeknecht (2013)	Environmental Sustainability Index
	Tomizawa and Niwa (1996)	General Indicator of Science & Technology
	Nicoletti et al. (2000)	Indicators of product market regulation
DEA	Mahlberg and Obersteiner (2001)	Human Development Index
	Cherchye and Kuosmanen (2004)	Sustainable development
	Melyn and Moesen (1991)	Macro-economic performance evaluation
BAP	Pennoni et al. (2005)	e-Business Readiness Index
	Gwartney et al. (1996)	Economic Freedom Indices
	Tarantola et al. (2004)	Internal Market Index
AHP	Puolamaa et al. (1996)	Index of Environmental Friendliness
	Singh et al. (2007)	Composite Sustainability Performance Index (CSPI)
CA	Ülengin et al. (2001)	Indicator of quality of life in the city of Istanbul
	Malkina-Pykh and Pykh (2008)	Quality-of-life Indicator
DCM-SRF	Corrente et al. (2021)	Portuguese Pandemic - Impact Assessment
	Kodikara et al. (2010)	Urban Water Supply
	Shanian et al. (2008)	Material Selection

Table A.2: Examples of aggregation methods

Method	Research Paper	CI or Application Field
MAUT	van Calker et al. (2006)	Overall Sustainability function for Dutch dairy farming systems
	Hajkowicz (2006)	Multi-attributed environmental index construction
	Cracolici and Nijkamp (2009)	Assessment of the attractiveness and competitiveness of tourist destinations
MAVT	Dantsis et al. (2010)	Assessment of the sustainability level of agricultural plant production systems
	Pictet and Bollinger (2008)	Public procurement Switzerland
	Dawson and Schlyter (2012)	Site suitability for solar thermal power station
ELECTRE	Attardi et al. (2018)	Land-Use Policy Efficiency Index
	Petrović et al. (2014)	Benchmarking European Union member countries
	Wu et al. (2016)	Site selection of an off-shore wind power station
PROMETHEE	Antanasijević et al. (2017)	Overall sustainability performance
	Hernandez-Perdomo et al. (2017)	Rank a portfolio of multiple firms' projects
	Herssens et al. (2008)	Service selection process

Table A.3: Examples of customer satisfaction indexes

Research Paper	Index
Fornell (1992)	Swedish Customer Satisfaction Barometer (SCSB)
Bruhn and Grund (2000)	Swiss Index of Customer Satisfaction (SWICS)
Fornell et al. (1996)	American Customer Satisfaction Index (ACSI)
ECSI Technical Committee (1998)	European Customer Satisfaction Index (ECSI)
Andreassen and Lindestad (1998)	Norwegian Customer Satisfaction Barometer (NCSB)

Appendix B

Methodology

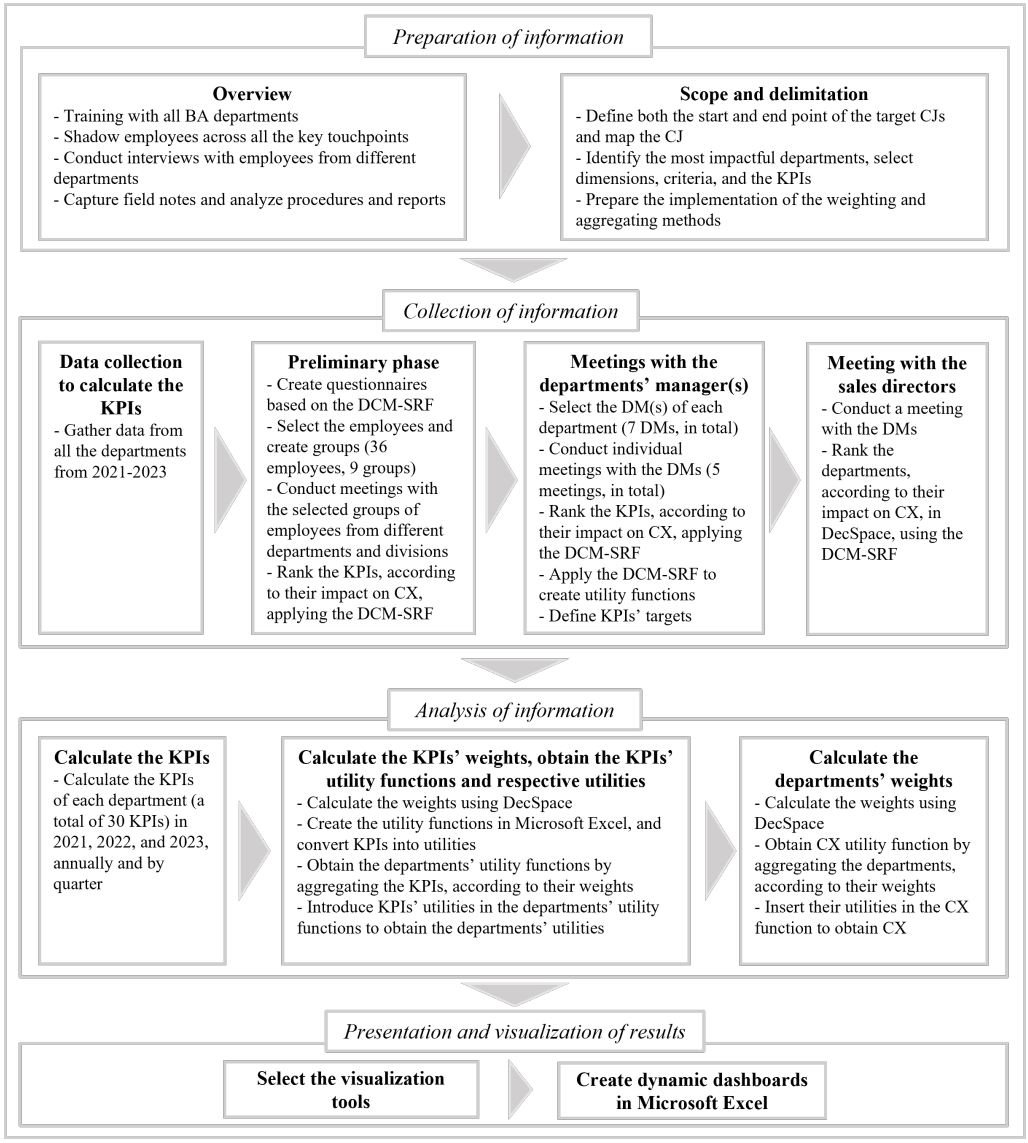


Figure B.1: Methodology diagram

B.1 Examples

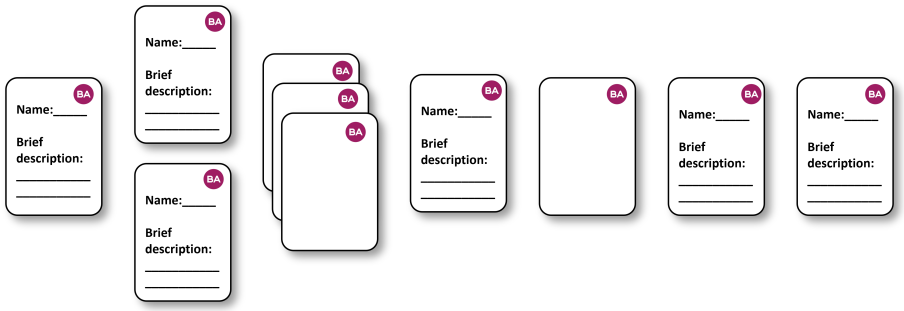


Figure B.2: Schematic example of a sample card played by DM in DCM-SRF. Adapted from: Corrente et al. (2021)

Example provided in the questionnaires

Situation: Ana wants to buy a car but cannot decide.
Decision criteria: *price*, *safety*, *cargo* and *passenger capacity*, *fuel type*, and *color*.
Note that:

– **Not all criteria have the same impact on her final decision.** For example, she considers price more important than color.

– **Differences in impact between criteria may be more or less significant.** For example, the impact of cargo and passenger capacity on the final decision is slightly bigger than color, but the impact of safety is much bigger than cargo and passenger capacity.

In order to choose the best car, according to her preferences, Ana was asked to:

1. Rank each criterion from the least impactful to the most. See the result in the image below.

   a) Interpretation: Ana considered price the most impactful criterion (rank 1) and color the least (rank 4). And she considered that safety and fuel type have the same impact (rank 2 for both).

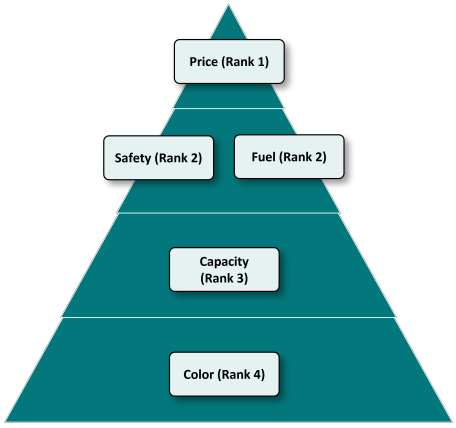


Figure B.3: Schematic example from DCM-SRF-based questionnaire - Part I

2. Assess differences in impact between criteria (the larger the number of levels between criteria, the bigger the difference). See the result in the image below.

   a) Two criteria in consecutive levels mean the difference is minimal.

   b) Criteria with 1 level between means the difference is twice the minimum.

   c) Criteria with 3 levels between means the difference is four times the minimum.

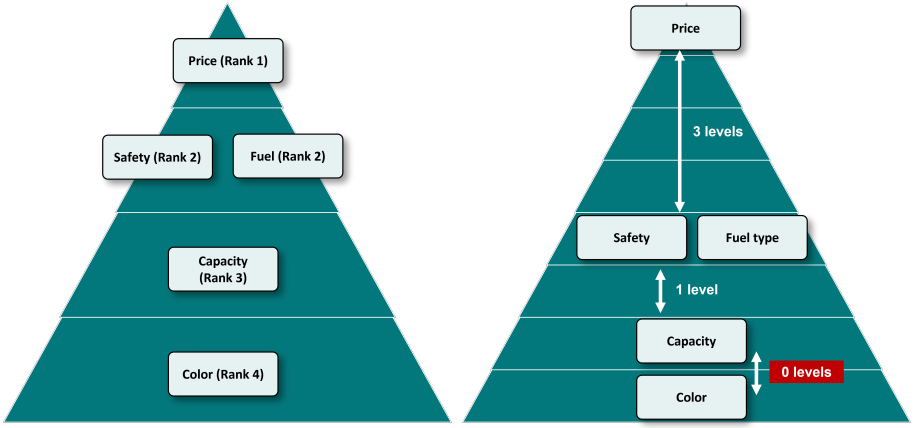


Figure B.4: Schematic example from DCM-SRF-based questionnaire - Part II

Please discuss the questions of this questionnaire according to the reasoning of Ana's example.

1. **Rank the criteria from the least impactful criterion to the most (rank 1)**, concerning the CX.

   Remember that if you consider that two criteria have the same impact, you must put them in the same rank position.

2. Second, **evaluate the differences in each criterion's impact**. The larger the number selected, the larger the difference in the impact.

   Remember that two criteria in consecutive levels mean the difference is minimal and the number of levels between is 0.

3. Third, **evaluate the difference between the items in the two extremes (the most impactful criterion and the least)**.

B.2 Participants

Table B.1: Preliminary stage - participants informations

Department	Division	#Group	Number of participants
FOA	Iberia	1	4
	Central Europe	2	4
	Southeastern Europe	3	4
Logistics	Iberia	4	4
	Central Europe	5	2
	Southeastern Europe	6	5
Transports	Iberia	7	5
	Central Europe	8	3
	Southeastern Europe	9	5
		Total	36

Table B.2: Presentation of departments' managers

Department	Position
OMT and Invoicing	Global Business Services Manager
FOA	Service Front Office Manager
Logistics	Distribution Manager Iberia
	Distribution Manager Southeastern Europe
Transports	Transport Manager Iberia
	Transport Manager Central Europe
Customer Service	Customer Service Manager

Appendix C

Case Study

C.1 Weights

Table C.1: Determining the normalized weights of *complaints* for $w = 1$ and $z = 10$

Rank r	Criteria in rank r	Number of blank cards, according to rank r	Non-normalized weights	Normalized weights
1	Serious Complaints	5	10.00	0.6410
2	Medium Complaints	3	4.60	0.2949
3	Warning Complaints	0	1.00	0.0641

Table C.2: Determining the normalized weights of *commercial attendance* criteria for $w = 1$ and $z = 3$

Rank r	Criteria in rank r	Number of blank cards, according to rank r	Non-normalized weights	Normalized weights	KPI weight * Dimension weight[a]
1	g_4	1	2.00	0.5000	0.2500
2	g_5, g_6	0	1.00	0.2500	0.1250

[a]It was determined that *commercial attendance* and *orders management* have the same impact on CX and, thus, both weigh 0.5000.

Table C.3: Determining the normalized weights of *orders management* criteria for $w = 1$ and $z = 2$

Rank r	Criteria in rank r	Number of blank cards, according to rank r	Non-normalized weights	Normalized weights	KPI weight * Dimension weight
1	g_8, g_9	0	2.00	0.4000	0.2000
2	g_7	0	1.00	0.2000	0.1000

Table C.4: Determining the normalized weights of Logistics for $w = 1$ and $z = 10$

Rank r	Criteria in rank r	Number of blank cards, according to rank r	Non-normalized weights	Normalized weights
1	Products Integrity	0	10.00	0.5030
2	Stock Accuracy	6	8.88	0.4467
3	Timeliness of Services	0	1.00	0.0503

Table C.5: Determining the normalized weights of Transports'dimensions for $w = 1$ and $z = 2$

Rank r	Criteria in rank r	Number of blank cards, according to rank r	Non-normalized weights	Normalized weights
1	Complaints, Timeliness of deliveries	2	2.00	0.4000
2	Tracking visibility	0	1.00	0.2000

Table C.6: Determining the normalized weights of *timeliness of deliveries* criteria for $w = 1$ and $z = 10$

Rank r	Criteria in rank r	Number of blank cards, according to rank r	Non-normalized weights	Normalized weights	KPI weight * Dimension weight
1	g_{23}	7	10.00	0.7246	0.2898
2	g_{21}	1	2.80	0.2029	0.0812
3	g_{22}	0	1.00	0.0725	0.0290

Table C.7: Determining the normalized weights of *tracking visibility* criteria for $w = 1$ and $z = 7$

Rank r	Criteria in rank r	Number of blank cards, according to rank r	Non-normalized weights	Normalized weights	KPI weight * Dimension weight
1	g_{26}	5	7.00	0.6667	0.1333
2	g_{24}	1	2.50	0.2381	0.0476
3	g_{25}	0	1.00	0.0952	0.0190

C.2 Utility functions

Table C.8: Utility conversion procedure for criterion g_1

Level	Performance	Blank cards	Number of units between levels	Unit valuation	Interval scale	Linear interpolation	Utilities
1/5	1.00	7			0.0000		0.0000
2/5	0.75	5			0.3810		0.3810
3/5	0.50	3			0.6667		0.6667
4/5	0.25	2	21	0.0476	0.8571		0.8571
Entity #1	Real					Eq. 3.3	Eq. 3.3
5/5	0.00				1.0000		1.0000

Table C.9: Utility conversion procedure for criteria g_2, g_3

Level	Performance	Blank cards	Number of units between levels	Unit valuation	Interval scale	Linear interpolation	Utilities
1/5	1.00	6			0.0000		0.0000
2/5	0.75	4			0.4118		0.4118
3/5	0.50	2			0.7059		0.7059
4/5	0.25	1	17	0.0588	0.8824		0.8824
Entity #1	Real					Eq. 3.3	Eq. 3.3
5/5	0.00				1.0000		1.0000

Table C.10: Utility conversion procedure for criterion g_4

Level	Performance	Blank cards	Number of units between levels	Unit valuation	Interval scale	Linear interpolation	Utilities
1/6	50.00%	3			0.0000		0.0000
2/6	60.00%	3			0.1481		0.1481
3/6	70.00%	4			0.2963		0.2963
4/6	80.00%	5	27	0.0370	0.4815		0.4815
Entity #1	Real					Eq. 3.3	Eq. 3.3
5/6	90.00%	7			0.7037		0.7037
6/6	100.00%				1.0000		1.0000

,

Table C.11: Utility conversion procedure for criterion g_5

Level	Performance	Blank cards	Number of units between levels	Unit valuation	Interval scale	Linear interpolation	Utilities
1/6	55	4			0.0000		0.0000
2/6	45	5			0.2174		0.2174
3/6	35	2			0.4783		0.4783
4/6	25	3	23	0.0435	0.6087		0.6087
Entity #1	Real					Eq. 3.3	Eq. 3.3
5/6	15	4			0.7826		0.7826
6/6	5				1.0000		1.0000

,

Table C.12: Utility conversion procedure for criterion g_6

Level	Performance	Blank cards	Number of units between levels	Unit valuation	Interval scale	Linear interpolation	Utilities
1/8	14	1			0.0000		0.0000
2/8	12	2			0.0625		0.0625
3/8	10	2			0.1563		0.1563
4/8	8	3			0.2500		0.2500
Entity #1	Real		32	0.0313		Eq. 3.3	Eq. 3.3
5/8	6	7			0.3750		0.3750
6/8	4	3			0.6250		0.6250
7/8	2	7			0.7500		0.7500
8/8	0				1.0000		1.0000

,

Table C.13: Utility conversion procedure for criterion g_7

Level	Performance	Blank cards	Number of units between levels	Unit valuation	Interval scale	Linear interpolation	Utilities
1/5	60.00%	3			0.0000		0.0000
2/5	45.00%	4			0.1538		0.1538
3/5	30.00%	5			0.3462		0.3462
4/5	15.00%	10	26	0.0385	0.5769		0.5769
Entity #1	Real					Eq. 3.3	Eq. 3.3
5/5	0.00%				1.0000		1.0000

,

Table C.14: Utility conversion procedure for criterion g_8

Level	Performance	Blank cards	Number of units between levels	Unit valuation	Interval scale	Linear interpolation	Utilities
1/4	15.00%	5			0.0000		0.0000
2/4	10.00%	7			0.3000		0.3000
3/4	5.00%	5	20	0.0500	0.7000		0.7000
Entity #1	Real					Eq. 3.3	Eq. 3.3
4/4	0.00%				1.0000		1.0000

,

Table C.15: Utility conversion procedure for criterion g_9

Level	Performance	Blank cards	Number of units between levels	Unit valuation	Interval scale	Linear interpolation	Utilities
1/5	4.00%	8			0.0000		0.0000
2/5	3.00%	9			0.3103		0.3103
3/5	2.00%	6			0.6552		0.6552
4/5	1.00%	2	29	0.0345	0.8966		0.8966
Entity #1	Real					Eq. 3.3	Eq. 3.3
5/5	0.00%				1.0000		1.0000

,

Table C.16: Utility conversion procedure for criterion g_{10}

Level	Performance	Blank cards	Number of units between levels	Unit valuation	Interval scale	Linear interpolation	Utilities
1/3	1.00%	10			0.0000		0.0000
2/3	0.50%	5	17	0.0588	0.6471		0.6471
Entity #1	Real					Eq. 3.3	Eq. 3.3
3/3	0.00%				1.0000		1.0000

,

Table C.17: Utility conversion procedure for criterion g_{11}

Level	Performance	Blank cards	Number of units between levels	Unit valuation	Interval scale	Linear interpolation	Utilities
1/6	20	1			0.0000		0.0000
2/6	16	2			0.1000		0.1000
3/6	12	3			0.2500		0.2500
4/6	8	4	20	0.0500	0.4500		0.4500
Entity #1	Real					Eq. 3.3	Eq. 3.3
5/6	4	5			0.7000		0.7000
6/6	0				1.0000		1.0000

,

Table C.18: Utility conversion procedure for criteria g_{12}, g_{15}, g_{18}, g_{27}

Level	Performance	Blank cards	Number of units between levels	Unit valuation	Interval scale	Linear interpolation	Utilities
1/3	0.010%	7			0.0000		0.0000
2/3	0.005%	3	12	0.0833	0.6667		0.6667
Entity #1	Real					Eq. 3.3	Eq. 3.3
3/3	0.000%				1.0000		1.0000

,

Table C.19: Utility conversion procedure for criteria g_{13}, g_{14}, g_{16}, g_{17}, g_{19}, g_{20}, g_{28}, g_{29}

Level	Performance	Blank cards	Number of units between levels	Unit valuation	Interval scale	Linear interpolation	Utilities
1/3	0.02%	5			0.0000		0.0000
2/3	0.01%	2	9	0.1111	0.6667		0.6667
Entity #1	Real					Eq. 3.3	Eq. 3.3
3/3	0.00%				1.0000		1.0000

,

Table C.20: Utility conversion procedure for criterion g_{21}

Level	Performance	Blank cards	Number of units between levels	Unit valuation	Interval scale	Linear interpolation	Utilities
1/6	50.00%	1			0.0000		0.0000
2/6	60.00%	2			0.0800		0.0800
3/6	70.00%	4			0.2000		0.2000
4/6	80.00%	5	25	0.0400	0.4000		0.4000
Entity #1	Real					Eq. 3.3	Eq. 3.3
5/6	90.00%	8			0.6400		0.6400
6/6	100.00%				1.0000		1.0000

,

Table C.21: Utility conversion procedure for criterion g_{22}

Level	Performance	Blank cards	Number of units between levels	Unit valuation	Interval scale	Linear interpolation	Utilities
1/5	20.00%	4			0.0000		0.0000
2/5	15.00%	5			0.1724		0.1724
3/5	10.00%	7			0.3793		0.3793
4/5	5.00%	9	29	0.0345	0.6552		0.6552
Entity #1	Real					Eq. 3.3	Eq. 3.3
5/5	0.00%				1.0000		1.0000

,

Table C.22: Utility conversion procedure for criterion g_{23}

Level	Performance	Blank cards	Number of units between levels	Unit valuation	Interval scale	Linear interpolation	Utilities
1/5	20.00%	8			0.0000		0.0000
2/5	15.00%	7			0.3103		0.3103
3/5	10.00%	7			0.5862		0.5862
4/5	5.00%	3	29	0.0345	0.8621		0.8621
Entity #1	Real					Eq. 3.3	Eq. 3.3
5/5	0.00%				1.0000		1.0000

,

Table C.23: Utility conversion procedure for criterion g_{24}

Level	Performance	Blank cards	Number of units between levels	Unit valuation	Interval scale	Linear interpolation	Utilities
1/5	0.00%	9			0.0000		0.0000
2/5	25.00%	7			0.2703		0.2703
3/5	50.00%	8			0.4865		0.4865
4/5	75.00%	9	37	0.0270	0.7297		0.7297
Entity #1	Real					Eq. 3.3	Eq. 3.3
5/5	100.00%				1.0000		1.0000

,

Table C.24: Utility conversion procedure for criterion g_{25}

Level	Performance	Blank cards	Number of units between levels	Unit valuation	Interval scale	Linear interpolation	Utilities
1/6	50.00%	1			0.0000		0.0000
2/6	60.00%	2			0.0833		0.0833
3/6	70.00%	4			0.2083		0.2083
4/6	80.00%	4	24	0.0417	0.4167		0.4167
Entity #1	Real					Eq. 3.3	Eq. 3.3
5/6	90.00%	8			0.6250		0.6250
6/6	100.00%				1.0000		1.0000

,

Table C.25: Utility conversion procedure for criterion g_{26}

Level	Performance	Blank cards	Number of units between levels	Unit valuation	Interval scale	Linear interpolation	Utilities
1/3	2.00%	10			0.0000		0.0000
2/3	1.00%	5	17	0.0588	0.6471		0.6471
Entity #1	Real					Eq. 3.3	Eq. 3.3
3/3	0.00%				1.0000		1.0000

,

Table C.26: Utility conversion procedure for criterion g_{30}

Level	Performance	Blank cards	Number of units between levels	Unit valuation	Interval scale	Linear interpolation	Utilities
1/5	4.00%	10			0.0000		0.0000
2/5	3.00%	7			0.3929		0.3929
3/5	2.00%	5			0.6786		0.6786
4/5	1.00%	2	28	0.0357	0.8929		0.8929
Entity #1	Real					Eq. 3.3	Eq. 3.3
5/5	0.00%				1.0000		1.0000

,

Appendix D

Future Work

Table D.1: Examples of suggested KPIs

Department	Suggested KPIs	Comments
OMT	Number of order changes whose changes impacted customers	Currently, it is only recorded the existence of changes and who carried them; but not its reasons, if they impacted the customers, the magnitude of the impact, etc.
FOA	Number of proactive follow-up calls after solving a complaint	Nurturing the relationship after a service failure is crucial, and follow-up calls have a great impact on SRJ and, consequently on CX
	Number of times wrong information was provided to the customer	A new case category should be created so that cases that are open due to inaccuracy of the information provided can start being monitored
Customer Service	Average time to complete a return, per segmentation	If there is a service or a quality problem that makes it impossible to use the delivered products, the damaged material must be collected as soon as possible. The time this task takes has a great impact on customers but it is not being measured today
	Number of visits to customers	This helps to nurture the relationship with customers and should be carefully monitored. It should also be created a memo of the visits
	Number of action plans created, per segmentation	Action plans are created to assure customers problems do not repeat, and they are the result of cooperative work between BA and the customer. Although these plans are being created, they are not being monitored by BA. Thus, today, it is not possible to measure their impact
Planning and Production	Number of order changes and order cancellations due to production changes	Today, it is not measured the impact of production changes on orders. Knowing the impact and whether some customers are more impacted than others is very important for improving CX
Returnable Packaging	Average time to collect pallets, after customer's request	BA has KPIs to measure only this department's performance. Thus, in order to measure its impact on CX, it would be important to implement these 5 KPIs since customers are complaining about the time it takes for BA to collect the pallets and the fact they have to make multiple requests to be attended. They also complain about the fact they are not informed in advance of pallets' collection; and consider it takes too much time to issue the respective credit note
	Number of customers' requests until the pallets are collected	
	Average time to verify damaged pallets	
	Total average time from return request to credit note issue	
	Number of complaints due to lack of prior notice of collection	
Invoicing	Average time to issue a credit note of a service and quality complaint	Although it depends very much on the client, in general, money issues greatly affect CS. And the time BA takes to pay the customer may have a great impact, that today is not being measured. It is therefore important to start monitoring these 3 KPIs
	Average time to issue the credit note after pallets return to BA	
	Average time to correct a wrong invoice, after customers' communication	