NOVIA
UNIVERSITY OF APPLIED SCIENCES

Introducing a single UX score concept for Company X.

A pilot study on the Y app (internal application).

Mariana Antonaki

Thesis for a Master's Degree in Business Administration (UAS) – Service Design

Turku 2023

DEGREE THESIS

Author: Mariana Antonaki
Degree Programme: MBA
Specialization: Service Design
Supervisor(s): Elina Vartama, Reija Anckar

Title: Introducing a single UX score concept for Company X. A pilot study on the Y app (internal application).

Date 10.10.2023 Number of pages 51

Abstract

The aim of this thesis was to develop a generic metric which measures the user experience (UX) with digital tools. The thesis was conducted as a pilot study on behalf of Company X, a forestry corporation, and the scope focused on an internal application (Y app). By combining quantitative and qualitative data, the single UX score concept aims to quickly assess the user-friendliness of digital tools via a survey distributed to their users.

For this pilot study, one digital application used internally (only by employees) was evaluated with the single UX score concept. This digital application - the Y app - is a relatively new internal tool with multiple functionalities, used mainly by employees at the company's production sites. All 32 production sites were included in this study, and 154 responses to the survey were received in total. Additionally, several other research methods were used: desk research, mapping exercises, user archetypes, in-depth interviews, and a mini-workshop.

In June 2023, the pilot study was successfully completed. During the following months, a second pilot was launched by the company, measuring 15 internal digital applications with the single UX score concept. As of September 2023, the concept was officially included in the company's KPIs, as a reliable UX quality metric. By the end of 2024, all internal digital applications are expected to be measured by the single UX score concept continually, on an annual basis.

Language: English Key words: UX, user satisfaction, SUS survey, score, efficiency, effectiveness, learnability, satisfaction, universality, usability, user-friendliness

---

Acknowledgements

Thank you to the UX team who supported this thesis project in every possible way throughout this journey.

Special thanks to Walter Brenson, Data Analyst, whose input was incremental to the creation of this thesis's concept, and especially for the analysis and compilation of all quantitative data used in this thesis.

Table of contents

Vocabulary

BA	Business Areas
UX	User Experience
KPI	Key Performance Indicator
SUS	System Usability Scale
UI	User Interface
EX	Employee experience
EFQM	European Foundation of Quality Management

1 Introduction

The commissioner of this master's thesis project is the UX team of Company X, a Finnish corporation in the forestry industry. Company X includes five different sub-companies, each specializing in different end products, all of which use wood as their primary source material. Even though each of the companies (Business Areas) is independent, with its own production sites (factories) and separate organisations, the Company's ICT department provides its services to all Business Areas (BA) in the form of internal consultancy: dedicated professional roles work exclusively for one BA, while other teams within ICT work simultaneously with all BAs on a strategic level. One of these teams, the UX team of the ICT department, is actively working to improve the User and Employee Experience across the organisation, by conducting independent studies on specific digital tools, processes, and other projects in collaboration with local and centralized management.

The UX team's initial request was to create reliable UX metrics which could be applied to different digital solutions, internally. With this – quite vague – request, the exploration started from the User and Employee Experience approaches, and how to make these approaches measurable. While ideating about possible metrics, a mind map depicting the various attributes that could be measured, was created (figure 1). This mapping exercise helped in defining the usage stages of a digital tool and identifying the user's possible needs and wishes during each of said stages.

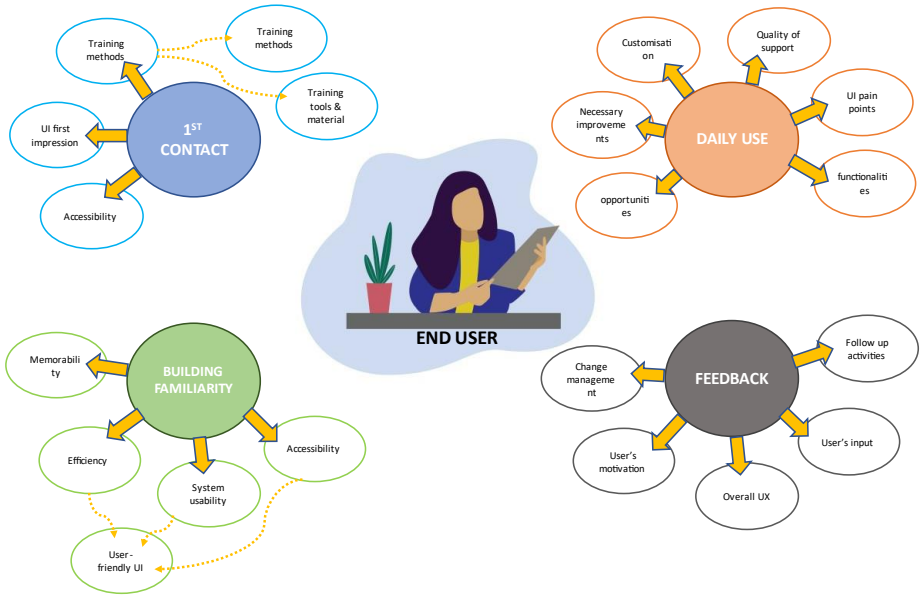


Figure 1. Mind map of the user's experience with a digital solution divided in stages

This mind map (figure 1), depicts how the user experience was divided into four stages: first contact with the digital tool, building familiarity, daily use, and feedback. Each phase represented a different point in time, and the emerging topics for each of them helped in identifying the main elements of user experience. Already at this point, attributes such as learnability and memorability of the tool, efficiency, accessibility, and training methods emerged as crucial to the overall user experience. At the same time, it became apparent that it would be impossible to dive further into the UX problem without focusing on one specific application or digital tool. By creating a pilot study focused on one application, it would be possible to understand its functionalities better, and investigate the users' expectations when working with it.

During the initial desk research, an earlier UX study performed in August 2022 on a relatively new internal application came up. Company X's external partner conducted 20 in-depth interviews with employees on-site and remotely, during which they assessed a mobile version prototype of the Y app. This application replaces earlier local digital solutions utilized at the company's production sites and aims to be a standard tool across the company's production sites.

1.1 Background information

Currently, Company X utilizes several digital applications at its production sites, some of which are only used locally or regionally; however, the newest digital solutions are designed for harmonized use across all mills and factories. Despite the fact that these digital applications are developed in a timely and costly manner by the company's partners (external providers), they often lack reliable UX quality standards.

The company's Production Management intended to implement a new, unified system measuring these applications' UX by applying the same metrics to all digital solutions across all production sites. In this way, the findings can be consistent and comparable with each other, demonstrating those applications with the highest UX according to their users, e.g. which aspects are better or more problematic.

Such a UX quality metric could help Company X perform more consistent quality checks during the development and usability testing of a new digital application, and revisit this metric continually, making sure that the UX quality is constantly improving or at least it

remains above acceptable levels. The single UX score concept was seen as a potentially useful UX quality metric by Production and ICT management, therefore this pilot study received abundant support.

1.2 Project's timeline

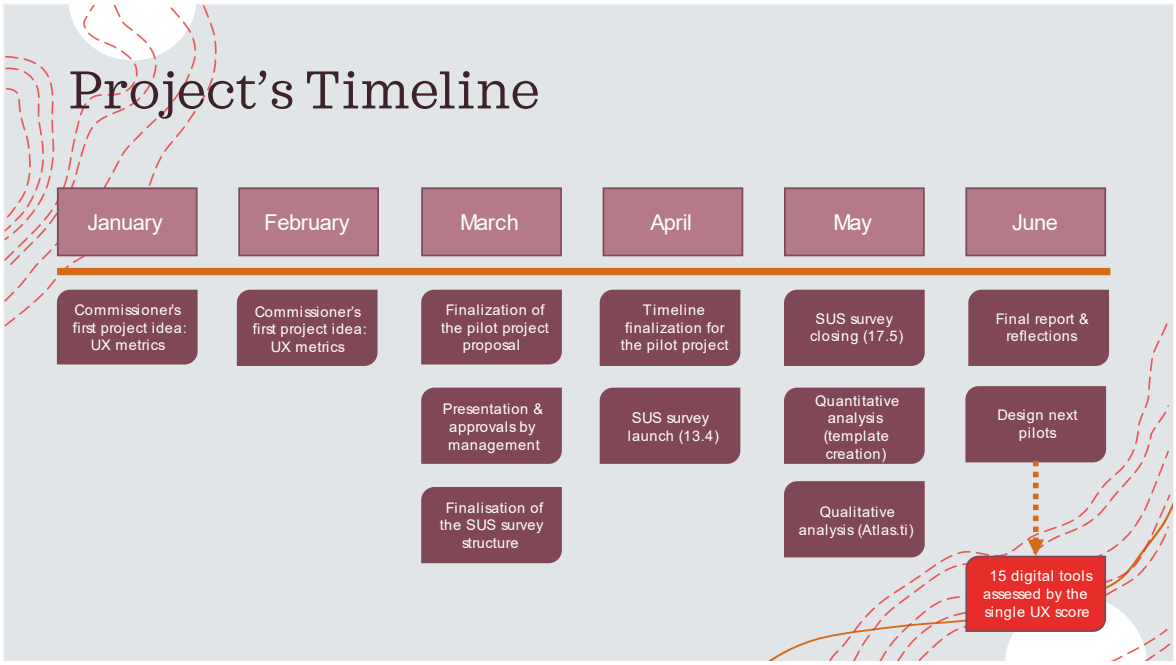


Figure 2. Timeline of the master's thesis project, including monthly sub-phases.

The UX team requested the development of a UX metric in early January 2023. There was no specific requirement or framework on which to base this metric; therefore extensive desk research was the only way to understand what the tools for measuring the employees' UX internally are, based on theory and other companies' practices.

In March 2023 the first draft of the pilot was presented to management, and the main tool, the System Usability Scale (SUS) questionnaire, was demonstrated. By April, the SUS customization for Y app users was ready and approved by all relevant partners (UX team, Production and ICT management teams), and the survey was launched on April 13, 2023. A variety of marketing tools were employed to advertise this survey to as many employees as possible, via intranet articles, QR codes to the survey displayed on info-screens locally, and in-app notifications to all Y app users. The survey closed on May 17, 2023.

A Power BI dashboard was created together with a data analyst, who calculated the quantitative data. The qualitative portion of the survey was analysed in Atlas.ti by the researcher.

The pilot's final report was submitted to management on May 30, 2023. In June, the main findings and future prospects for the single UX score concept were presented in multiple management meetings, and to various other internal stakeholders who wished to apply the concept to their digital solutions. Over the summer, the concept was successfully applied to 15 other internal applications, and more applications will follow during 2023.

2 The Y application

The Y application replaced the traditional 'shift diary' at some of the Company X's production sites in 2017, and until 2023 it has been implemented at 32 sites altogether. The application is mainly used by operators, production line workers, local managers, and equipment maintenance teams for communication and documentation purposes. More specifically, this application is the primary tool where all machine breaks, production issues, and any other equipment faults are logged (as texts or pictures) so different professionals are able to react accordingly. Each factory shift consults the Y app before starting their work on issues that might have occurred during the previous shift, and they maintain an overview of the machinery by exchanging messages in real time.

Additionally, the Y app contains all important documentation necessary for the work processes on the factory level: fire security certificates, work permits for operators, medical declarations for the workers' ability to perform their tasks, machine logs (production rates, faults, breaks) and many more types of documentation.

Each role uses the Y app in a different way. The production line workers and machine operators use it to log and read crucial information regarding the equipment they operate daily. Local managers and white-collar employees (e.g. finance professionals) use the app to retrieve reports and log data into other systems, especially regarding production figures and KPIs. Maintenance teams are alerted through the Y app on machinery breaks or faults, and they coordinate their reaction through the app together with their colleagues.

The Y app is only one of multiple applications utilized on the production line level; however, it is the most important tool for day-to-day production work. Therefore the Y app is a necessary and irreplaceable tool for its users.

The Y app initially worked solely as a desktop tool, meaning that the blue-collar users on the production line used it on shared computers with shared credentials. Individual accounts were more limited, and primarily used by white-collar employees and management. Later on, a mobile version was launched, which only worked with individual accounts, but until this study was completed (June 2023) the majority of the production line employees still used the Y app through shared accounts and computers.

2.1 Limitations

The Y application is used at most of the company's production sites – however not at all of them yet. Therefore, the users do not possess similar levels of familiarity with the tool. Company X is working closely with the vendor on building the application's features, and the current solution has a unified image for specific tasks or categories in the tool's interface, and fewer features customized for each production site accordingly.

Despite the tool's complexity, the current reporting capabilities remain low, and focus on the machinery performance rather than the users' interaction with the tool. There is an 'error' counter, but data such as completion times per task or incomplete tasks are not logged.

Additionally, most user accounts at the production sites are shared, meaning that each account and credentials are used by numerous employees who log their tasks and observations through the same shared device (laptop). The lack of individual accounts automatically makes the collection of data per user practically impossible. The company has decided to move away from the shared accounts' model in the near future; however, this implementation has not taken place yet.

A standard limitation regarding data gathering, has been the imbalance among production sites' locations and local languages. Most of the production sites are located in Finland, therefore the Finnish workforce is the clear majority. However, when it comes to user

experience and human-centric service design, we must take into account as many aspects as possible, and understand the different pain points of people.

2.2 Future plans

Prior to the Y app, the work was performed in various other ways. Some production sites had earlier applications, developed locally or regionally, with fewer functionalities than the Y app. Other sites logged their work manually, using Excel sheets or handwritten documents.

In order to improve this situation and make the information logs more reliable and unified, production management introduced the Y app initially in 2017 at a limited number of production sites. During the next few years, the app was deployed to the majority of the production sites, and now it is functioning at 32 sites either fully or in a pilot status.

In May 2023 the blue-collar staff started to receive company mobile phones and credentials for individual accounts. The goal is to eliminate shared accounts in all production applications (including the Y app) and each employee to have their own company mobile device and personal account for daily work use. One of the positive outcomes of this initiative is the ability to collect more precise data, and give more flexibility to the employees, allowing them to react faster during breaks and emergencies, as they can simply use their mobile device, instead of running to the closest shared computer and logging the relevant information there.

3 Research purpose

The hypothesis about the single UX score concept is that it can become a reliable universal tool for measuring the UX of internal users (employees) in a simple and standardized manner. The main argument is that by establishing this measurement as a company-wide quality standard, the overall employee experience will be improved. Additionally, if the single UX score is measured consistently, for instance annually, and if improvement steps are taken in-between measurements (new deployments of the tool, renewed training methods, usability testing, prototyping workshops), then not only will the UX score of any

given app gradually increase, but the overall employee experience will be improved as well, since people will recognize the company's efforts to listen to their needs and wishes and accommodate them to their best capabilities.

The purpose of this study was to understand the pain points and the opportunities offered, due to the Y application, according to its users and to demonstrate solutions created by the users themselves. On a second level, the aim was to spot the difference among various user groups' evaluation of the Y app, in order to better understand the priorities on a company level, for instance, the training needs or necessary User Interface (UI) changes. The study's furthest objective is to gradually implement the single UX score concept as a new quality metric for Company X, across its internal digital solutions, in the near future.

3.1 Research questions

For this study, we hypothesize that a digital system's user experience can be measured as a single score (%), which can also show the levels of user satisfaction across different pre-defined user groups. The single UX score will be measured via a quantitative survey of ten questions (SUS survey; Brooke, 1996), with the addition of a qualitative module: one open-ended field for ideas, suggestions, and free comments by the respondents. A few introductory questions would focus on the creation of user groups (user archetypes), helping with the interpretation of the results. Embarking from this hypothesis, the following research questions emerge:

1. How can a single UX score set the quality baseline for employee experience?
2. Why is the single UX score concept a reliable quality metric?

The introduction of the single UX score concept for the Y app is an attempt to quantify the tool's UX, while prompting users to give their own opinions, ideas, and recommendations for future improvements. Even though this approach is not 'traditionally' service design-oriented, the goal is to incorporate design thinking methods, by combining quantitative and qualitative data as crucial modules for building better UX of internal digital tools.

The single UX score is a numerical figure which outlines the overall users' satisfaction, but also dives into specific areas who define the tool's usability, such as its efficiency, learnability, and effectiveness. However, the score itself does not do much more than

reflect a static picture of the users' current opinions about it. Therefore, the addition of qualitative input is necessary as it sets the scene for emerging pain points, future opportunities, and new directions for improvement.

3.2 Research methods and Service Design Tools

The main challenge of this pilot was to fit a UX topic into the Service Design methodologies. Since these two approaches are different, a middle way was invented. Some of the standard research methods in Service Design were used: desk research, literature review, a survey, 20 in-depth interviews, and a mini-workshop on lessons learned. However, they were all adapted to the project's core: quantifying the Y app's UX.

For instance, a major part of the desk research – once the SUS questionnaire was selected – was performed on the company's intranet content, in search of similar projects done before, standard practices utilized currently for assessing the employee experience (EX), and the digital solutions' user-friendliness.

The System Usability Scale (SUS) survey was this project's primary data source, based both on quantitative and qualitative input. The addition of user archetypes' categorization to the original SUS survey structure served as a necessary tool for more accurate results, correlating the users' assessment of the Y app with their professional roles and technical background. The SUS survey is a straightforward tool for assessing a digital tool's usability and learnability, based on quantitative self-reported data by the tool's userbase.

Additionally, a mini workshop on lessons learned was arranged, with one training specialist, who reflected on the survey findings and focused on the improvement of learning practices. Specifically, the workshop participant produced a scenario-based trainer's journey map and a training service blueprint, focused on the information flow in training materials and during training sessions. This mini workshop, based on the survey respondents' input, as well as the trainer's own experience from successful training methods, serves as a good baseline for the creation of new learning tools and methods, focused on interactive participation and need-based training.

The 20 in-depth interviews used in this project as a secondary data source, were conducted by an external partner in 2022, and focused on the assessment of Y app's mobile version

prototype. Although that project's scope was mainly the new mobile version development, the users were asked to evaluate the Y app's user-friendliness and overall usability, and their insights corroborated this pilot's qualitative findings to a great extent. These semi-structured interviews followed a simple structure, and each participant was free to select the topic they wanted to focus on. The summaries of these interviews, as provided by the external partner, were embedded to the rest of the qualitative data, and analysed according to the same groupings and codes, as a whole, together with the survey's open comments.

Several service design mapping exercises were carried out during this thesis project, as assisting tools throughout the design process (Stickdorn et al. 2018). A mind map (figure 1) already presented in the introduction, was the first mapping exercise, which helped in the identification of potential UX metrics.

Later on, the single UX score concept was analysed through the framework of Innovation's lens (figure 4) to depict more precisely the activities during each phase of the design process.

An ecosystem map (Stickdorn et al., 2018) was created after the data analysis phase, to explain all affected parties after the single UX score concept's implementation at company X. Along the same lines, a business model canvas focused on the involved parties, the created value, as well as the cost and revenue streams and how they will be affected by the new UX metric.

4 Theoretical Background

This pilot study is based on the theoretical framework of Innovation and the EFQM model, through the lens of Employee Experience (EX). The framework of innovation is appropriate for explaining the single UX score concept's process, translated into the four fundamental phases of the double diamond, with the addition of guiding design principles. While diving into the framework of Innovation, the EFQM model discusses the crucial role of quality metrics in the workplace, and how their assessment can form a better Employee Experience in the future.

4.1 From the double diamond to the framework of Innovation

The Double Diamond is an established representation of the design and innovation process, used widely for the past 20 years. It is a simple structure that demonstrates the different

phases during the design process, from discovery all the way to delivery (Design Council, n.d.).

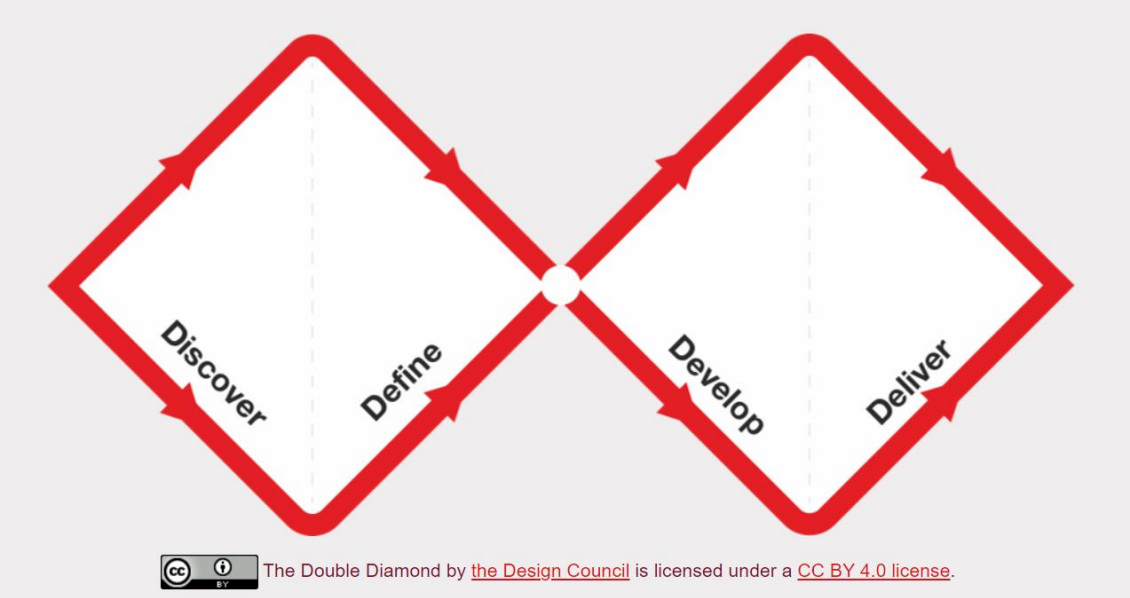


Figure 3. The Double Diamond structure.

The discovery phase is critical, as this is already when different stakeholders are involved, and the user is set in the middle of the process. Who do we design for? is the main question to be answered at the very beginning of the new design process. Additionally, during discovery deeper issues are investigated, always revolving around the user and their needs.

During definition, more concrete ideas are presented, which are the research pillars of the design process. In that stage, all important aspects are identified, using insights gathered during the discovery phase (Wikipedia, n.d.).

The development phase is crucial for the introduction of new ideas and concepts. This is the opportunity for creativity and good design to shine, revealing the main pain points, opportunities, and new ideas that have emerged from the users during iterative activities, such as workshops, interviews, and focus groups.

In the final phase, the delivery, all attention is concentrated on the presentation of tangible results. This is when new solutions are prototyped and tested, and only those with good potential are kept in the end.

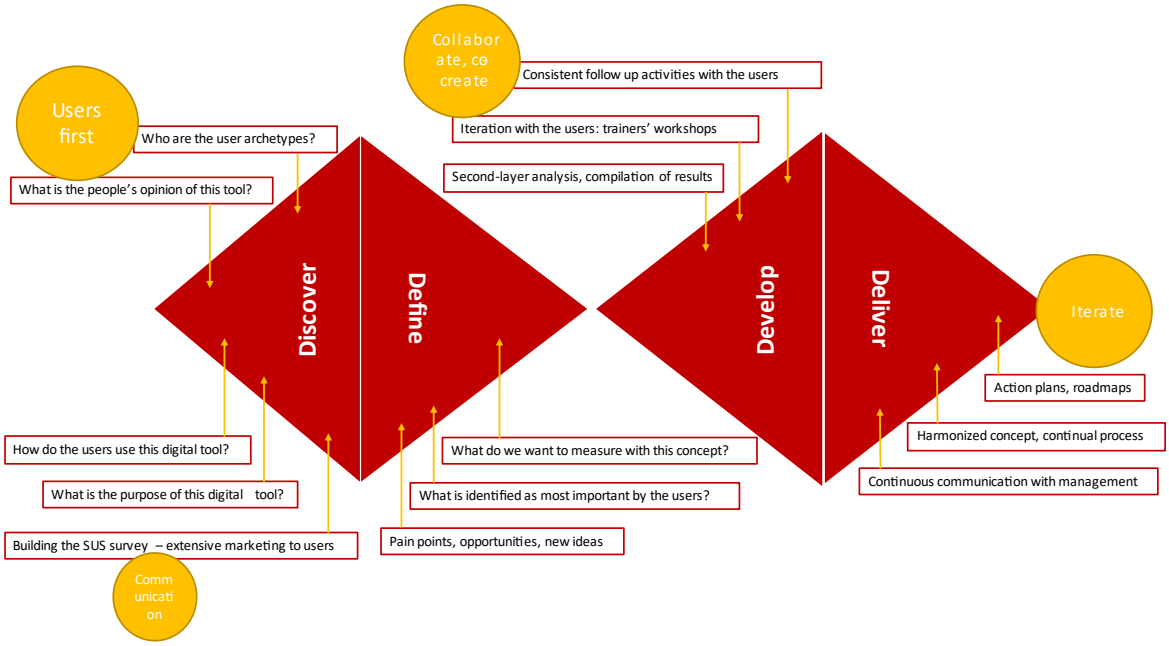


Figure 4. The single UX score concept through the framework of innovation.

However, the most important aspect of the double diamond process, is its iterative nature. The four phases are not linear, and the whole design process should be approached as a continuous circle, inside which the phases are revisited and the steps repeated, and the delivery stage is not the end, as in design there is no final end (Design Council, n.d.).

The framework of Innovation, based on the double diamond structure, goes one step further: it defines design principles, and includes leadership and user engagement as necessary aspects of a successful design process.

When applying the framework of innovation to the single UX score concept, the focus starts with the users' needs and expectations from a digital tool. The users participate in a survey built on their archetypes, in order to better understand each user group's needs by taking their special characteristics into account.

The collected input is analysed in multiple ways: different quantitative calculations, qualitative coding and grouping. This process aims at creating rich outcomes of the users' responses to the survey, and taking advantage of the whole spectrum of people's answers, whether it is numerical input, open comments, or interviews. This deep analysis provides a

solid basis for the development phase, when iteration with the users is central. Workshops, focus groups, discussions with multiple stakeholders on the findings, are the main methods compiling the development phase. The goal is to corroborate persistent shortcomings and pain points flagged by the users, while also identifying new ideas, solutions, and opportunities.

Finally, the delivery of well-designed solutions, action plans, and roadmaps with details on the necessary changes and improvements, ensures the fruitful conversation between leadership (decision-making) and users (engagement) (Design Council, n.d.).

4.1.1 The SUS survey

The single UX score concept was based on the SUS (System Usability Scale) survey, developed by John Brooke (Brooke, 1996) and provides a fast and reliable way to assess a digital system's usability. The SUS survey is a reliable tool during the empathizing stage because:

- it reaches large numbers of people fast
- its form is compact, comprehensible, and easy for the respondents
- it shows to the respondents that their user experience is important for their organisation, and that there is a concrete willingness to improve it

For these reasons, it was selected over other polling methods that have been popular in corporate environments such as the NPS (Net Promoter Scale) or CFF (Customer Feedback Form).

The SUS survey is an accurate tool for defining the users' wants and needs, spotting pain points, and giving a quick view of a system's usability, depending on its users' input. The definition phase in a design thinking process is really important, as it sets the tone for the next steps, and whether the overall path is appropriate for a certain audience. By relying problematic with any given digital tool.

Reaching the ideation phase, the data analysis received by the SUS survey's respondents was the core activity of this project. The quantitative and qualitative data were analysed separately, and alongside the UX score, a number of pain points and opportunities were identified.

In order to dig deeper into the survey's findings, a mini workshop on lessons learned was organised, as an example of how ideation should be approached. During this session, the SUS and interviews' findings were compiled in a concise report which focused only on one angle: learnability. Since the workshop participant was an experienced Y app trainer, the ideation was on what people had said about the training methods and tools, the ability to learn to use the Y app independently, and what could be improved in that area.

Starting with the lessons learned on learnability, new solutions were discovered, especially during the creation of the training service blueprint. The trainer gave insights on the usual training methods, while highlighting shortcomings and potential solutions to gaps already pointed out by the users. Even though additional workshops covering the other attributes could be beneficial, learnability was the key to interpret the study's findings: it was the most commented-on topic and consistently the lowest marked attribute of the survey, across user groups.

During the prototyping phase, the Y app's roadmap to better UX was created by grouping and prioritizing the users' qualitative input (SUS and interviews). The main topics mentioned by the users formed the pain points' list. For the roadmap recommendations, two aspects were taken into account: critical or non-critical changes according to the users' views, and low urgency/effort/cost – high urgency/effort/cost changes. In the roadmap, the priorities were listed from most easy to accomplish to most complicated and time consuming changes.

4.2 Frame of reference: The EFQM model

The EFQM (European Foundation Quality Management) model is a frame of reference that Company X heavily utilizes. Its focus is on the quality management of internal tools and processes, but also stakeholders' networks, projects, and external relationships between Company X and its external partners. In a nutshell, the EFQM model focuses on three questions (EFQM model, 2020):

Why does an organisation exist? (DIRECTION)

How does it intend to deliver on its Purpose and its Strategy? (EXECUTION)

What has it actually achieved to date? (RESULTS)

Figure 5. The guiding principles of the EFQM model.

According to the model, these areas are interconnected and viewed as a whole, as purpose cannot be separated from strategy, and the delivered value is defined by both of these aspects.

The EFQM model was used as a frame of reference for the single UX score concept, specifically regarding the 'Execution' setting and all of its points (figure 4). The model utilizes design thinking and relies on the stakeholders' input, focusing on change management, quality improvement, and helping organisations in their journey to innovation. The EFQM core is the achievement of a sustainable future by creating and maintaining high quality products, services, and processes.

According to this frame of reference, there is a stakeholder-centric view, and the organisation must harmonize their approaches and processes to it. The EFQM is founded upon EU values (EU, n.d.), the 10 principles for sustainable and socially responsible business by the UN Global compact (UN, 2000), and the 17 UN sustainable development goals (UN

Sustainable Development Goals, n.d.), making it very relevant to European organisations and their stakeholders, in line with future challenges and megatrends.

Approaching the organisation as an ecosystem, where all actors are interrelated, there are no linear goals for change and innovation, but rather a more holistic approach of dynamic growth based on anticipated future challenges and set sustainability targets.

The single UX score concept fits the EFQM framework by measuring the UX quality of digital tools, thus contributing to more sustainable change and user-friendly tools. Additionally, it reinforces the stakeholders' impact on the organisation's strategic goals and policies, by driving the tools' performance, and creating sustainable value. Essentially, the single UX score concept is built on sustainability principles: it contributes to the stakeholders' vision of an innovative organisation by assessing their user experience, and it assists in the establishment of an inclusive culture and integral organisational change by focusing on performance and meaningful innovation (EFQM, 2020).

A crucial criterion of the Execution setting is the stakeholders' engagement. Identifying the stakeholders and mapping their position in the organisational ecosystem, is critical for understanding their needs and expectations, and aligning them with the organisational strategies. For this pilot study, a stakeholders' map was created, and users' archetypes were utilized as a method to identify the main user groups of the Y app. The purpose of these tools was to focus on the stakeholders in more detail, and in order to define the most important aspects of their user experience with corporate digital tools.

Creating sustainable value, another Execution criterion, focuses on the impact of organisational change over the stakeholders' viewpoints. This is why collecting their feedback continually, within a harmonized process, guarantees that the organisation will be able to spot shifts fast, and timely plan corrective moves. The single UX score concept relies a lot on continuity and carefully planned follow-up activities, in order to reinforce the perpetual creation of value for the stakeholders and the organisation as a whole (EFQM, 2020). Follow-up workshops with experts are very important for creating action plans and change management roadmaps. These activities are also crucial for confirming the organisation's pledge for sustainable innovation through a human-centric lens, which has great impact on the stakeholders' morale and ensures their intention to actively support positive change.

Since September 2023, the single UX score concept is adopted by one of the Business Areas of Company X as a new EFQM metric, and it is included in all relevant reports. This development means that the concept was recognized for its efficiency and positive impact on sustainable innovation, according to the EFQM standards, and that the company has decided to apply it to all internal digital solutions on annual basis.

4.3 Design Thinking meets Employee Experience

This thesis's main objective was to approach and assess certain aspects of the employee experience (EX). What is the EX? It can be defined as the employee's expectations and needs throughout their employment relationship with their employer, starting from the hiring process. By seeing this process as a holistic experience, a sense of belonging and performance co-creation emerges (Whitter, 2019). Decades ago, when the EX appeared as a term, an incremental part of the employee experience was the use of internal digital tools, necessary for their daily tasks. The focus was on productivity, later on engagement became the EX cornerstone. Today, when talking about EX evolution, the focus is on designing experiences by following the employees' ideas and suggestions (Morgan, 2017). Therefore, assessing the user satisfaction and overall experience of these tools, is crucial for smooth business continuity.

Employee Experience as a concept goes beyond the concept of payment for a pre-defined task, as employment has been traditionally seen. It touches upon meaningful work and motivation of the employee, who seeks professional development through their performance. Even the relationship between employees and managers are redefined, with the manager assuming the role of a coach, who wants to help their employees thrive and be the best versions of themselves, not only by improving their work performance but also by seeking more challenging roles (Bridger & Gannaway, 2021). The latest development in the EX landscape has been the approach of the employee with customer standards. This means that the employees' wishes and needs, as well as their motivation on the job, their hopes, and advancement plans are very important for their employer as they are the core elements of the employee's engagement.

In design thinking terms, the process of building a more iterative EX starts with empathizing. It is crucial for the EX process designers to understand the employees, and

their deeper motivations in the workplace. Contrary to traditional business methods, jumping to solutions is not the goal in this case; approaching and understanding the employee should not be biased or considered as unnecessary (Bridger & Gannaway, 2021). In the same path, the single UX score concept focuses on the future EX and how it could be closer to the employees' actual needs, by looking for solutions and ideas directly from the source: the employees themselves. The iteration starts from a blank page, without pre-assumptions of the users' ideas on their digital UX.

Figure 6. The three EX design lenses (Bridger, Gannaway. 2021).

Since the user-satisfaction focus at Company X has been solely on the external customers so far, the user-friendliness of internal digital solutions was somewhat neglected. One of the main goals of this pilot was to provide a unified idea of what is meant by 'employee experience'(Kallio, 2021 November 17), by creating a commonly-used feedback method that every user can understand, such as a survey. Therefore, this pilot study had to utilize different tools, more fitting to internal assessments. For instance, when selecting the input form for the data, methods such as the NPS (Net Promoter Scale) or the CFF (Customer Feedback Form) that are usually preferred for external surveys, were not applicable. The overall goal of this study was to assess the internal users' satisfaction with the Y app, taking into account that currently there is no alternative to this tool, and that some mandatory work tasks are solely done via this application.

4.4 User Archetypes

Before processing the data received by the SUS survey, a user group categorization was necessary. The Y app is a mandatory tool for a variety of professional roles, therefore users are of various technical and professional backgrounds as well. User archetypes were preferred to personas for this study, as the main objective was to determine the application's user-friendliness across each user group, focusing only on a few common characteristics among each group's members and not a user's full persona (Laubheimer, 2022, May 15). The archetypes made data processing easier, since the pilot's sample size was quite large, and the most essential information for the users were these: what's their job function and what's their tech-savviness level. By this categorization, the universality and learnability attributes would be most relevant, showing if the Y app can be used only by certain groups or if it is a universally easy tool.

Nine user groups (archetypes) were identified, and the survey was modified accordingly, so the respondent first would have to self-assess their job function and technical skills' level, before proceeding to the standard SUS questionnaire.

The categories were classified as follows. Blue-collar employees, the production sites' workers whose job is directly on the production line; most of them do not have company credentials and use the Y app on shared laptops on the factory level. White-collar employees, are not necessarily located at the production sites, but their job requires Y app use; they have personal credentials and always use the Y app from their personal devices (mobile, laptop). The key users, are employees who have participated in the development and deployments of the Y app, as well as the users' training; they have personal credentials and can use the Y app both on personal and shared devices, and they possess the most expertise around the app's features and functionalities.

The individual technical skills' level was self-assessed by the respondents, and there was no standard scale against which the skills would be measured. This poses a risk, as there is no corroboration of the users' assessment, but it remains low as employees are required to have a certain level of technical background in order to perform their daily tasks in general.

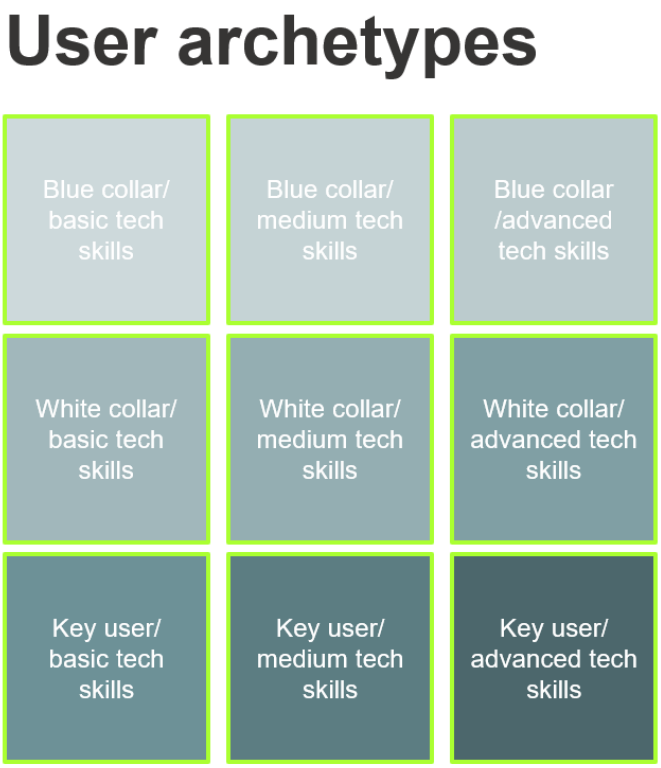


Figure 7. The 9 user archetypes created for the pilot study.

5 The single UX score concept's structure

A first draft of the single UX score concept for Company X was introduced in March 2023 to the ICT and Production management teams. Even though initially this plan was designed solely as a master's thesis project, after management's attention was drawn to it, the scope expanded. It was soon decided that all production sites should participate to the pilot, because such a comprehensive UX study was never done before in the company.

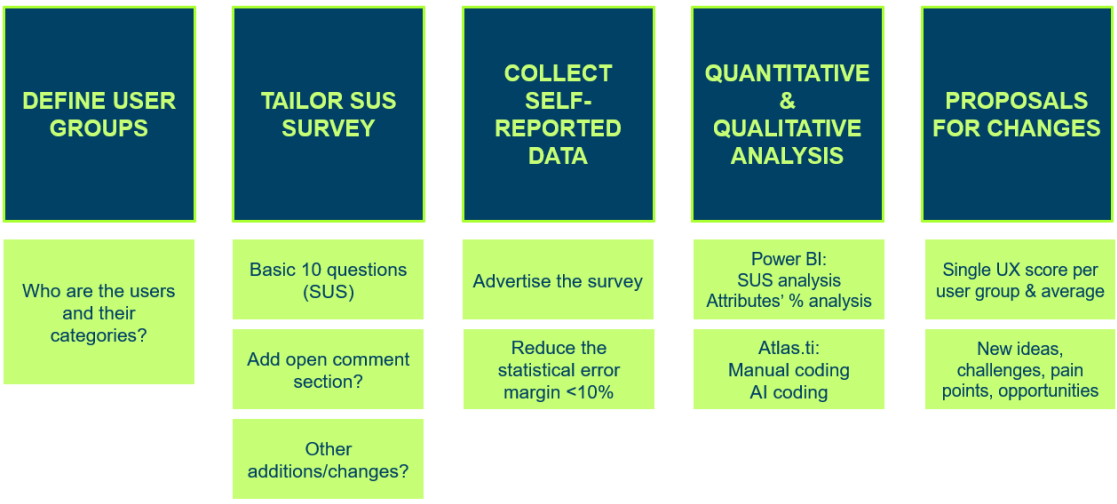


Figure 8. The single UX score process.

The concept is designed as a customisable tool, incorporating numerical and qualitative input. For each project, the process starts with the definition of user groups, continues with the collection of data through the SUS survey, and the data analysis of the received input. A crucial part of the process is conducted after the data analysis, and includes co-creation sessions with experts, focusing on the observed pain points, challenges, opportunities, and new ideas. In a nutshell, the single UX score concept provides deeper insight to the users' pain points, by observing 5 attributes:

a. ease of use, appealing interface (satisfaction)
b. system's learnability (learnability)
c. understandable content with logical menus and paths (effectiveness)
d. effective specifications and functionalities (efficiency)
e. independent use regardless of individual background (universality)

The final attribute, universality, assesses if a digital tool is user-friendly across all user groups. By comparing the UX single score of each user group, it is visible which groups mark highest and lowest among the rest of the groups. If the distance between the highest and lowest group is too wide, it means that the application is not user-friendly across all the population, therefore, some particular user groups require more support and longer time to gain confidence with the tool.

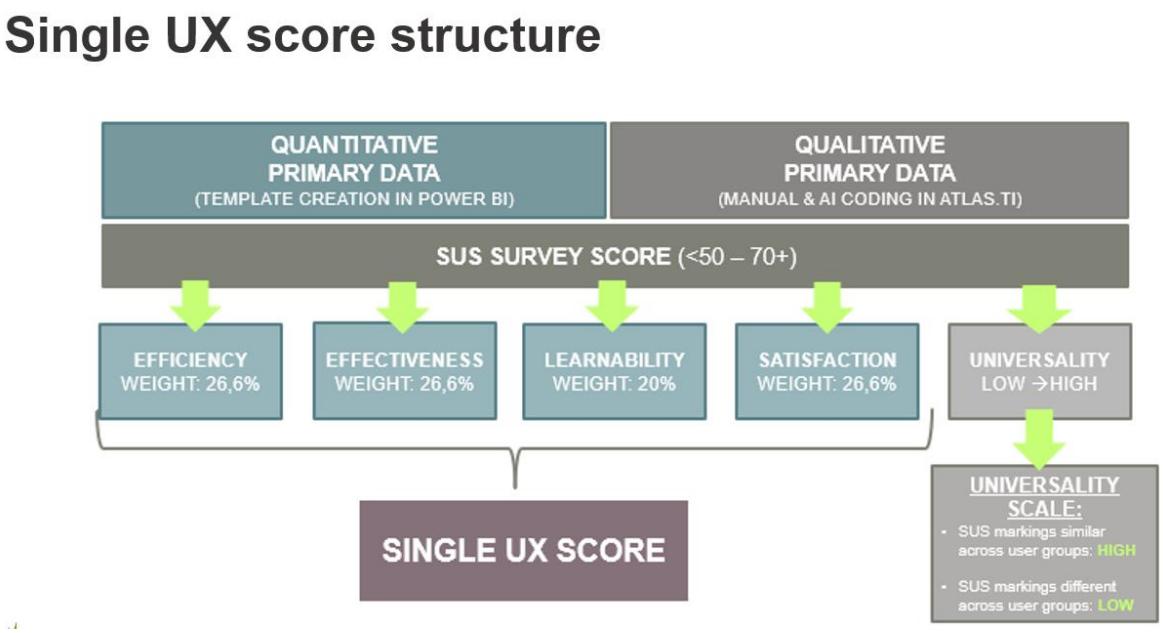


Figure 9. The Single UX score structure, including the 5 main attributes

The quantitative portion of the concept, presents a static image of how the users think of the digital application they are assessing. The overall score (%) gives a glimpse of the application's usability, before diving into deeper explanations. However, a tool's usability scale is determined by multiple factors: usability and learnability (Lewis & Sauro, 2009). As already established, the usability factor is quite diverse, including different attributes, such as the user's satisfaction from the tool, and the tool's efficiency and effectiveness. Since the SUS questionnaire assesses all of these attributes, it makes sense to place them hierarchically, based on their representation in the SUS questions, corresponding to each attribute's definition.

Usability is compiled by three attributes, according to ISO 9241-11 standards (LeSuer, 2022 March 3):

> **Satisfaction:** the user's overall impression of the tool; if they are comfortable using it or frustrated. The user's general feelings.
>
> **Efficiency:** the tool's ability to deliver the tasks in the quickest possible way, without wasting the user's time and efforts.
>
> **Effectiveness:** the tool's general ability to perform and deliver the necessary tasks it is meant to deliver
>
> **Learnability**: the user's ability to understand and memorize the tool, either with training materials or other learning methods.

By assigning specific questions and weight to each of the four attributes, it is possible to extract which attribute scored highest or lowest for each user group. The three usability attributes are weighed by 26,6% each, as they correspond to eight out of ten SUS questions. Learnability corresponds to the remaining two SUS questions; therefore, it is weighed by 20%.

The 'weighing' calculation method creates a deeper understanding on which attributes matter most for the users, and gives an insight on what should be the next steps. All in all, the single UX score concept is designed to be something easily adaptable to different purposes, generic enough so it could produce comparable results when utilized in different case studies.

QUESTIONS	ATTRIBUTES
1. Which is your work site?	USER GROUP CATEGORIZATION
2. What's your role? White collar/Blue collar/Key user	
3. You use & on mobile, desktop, or both?	
4. What's your technical skills level (basic/medium/advanced)	
5. I enjoy using the Y app	SATISFACTION
6. I find the Y app unnecessarily complex	EFFICIENCY
7. I find Y app simple to use	SATISFACTION
8. I often need support when using Y app	SATISFACTION
9. I think the Y app functionalities are good	EFFECTIVENESS
10. The Y app is not very consistent	EFFECTIVENESS
11. It is easy to learn how to use the Y app	LEARNABILITY
12. The Y app is inconvenient to use	EFFICIENCY
13. I am confident while using the Y app	SATISFACTION
14. Training is necessary before using the Y app	LEARNABILITY
15. Open comments (obligatory)	

Figure 10. The SUS survey for Y app, including attributes per question.

5.1 The concept's business value for Company X

With the implementation of a new quality metric, there is a direct impact to the employees' experience (EX) in the workplace, and more specifically the digital employee experience (DEX) (Gheidar & ShamiZanjani, 2020). This concept is designed to focus on the respondents' views and ideas, giving priority to their own knowledge and expertise on a digital tool. It is meant to facilitate a smooth digital transformation in the workplace, with a bottom-up approach: starting from the employees' wants and needs, and moving upwards to management's decision making processes.

However, the concept has an overall impact to the company as well. For this reason, a business canvas model was created, in order to demonstrate principle key partners, activities, structures, and propositions, during the implementation of the single UX score by Company X.

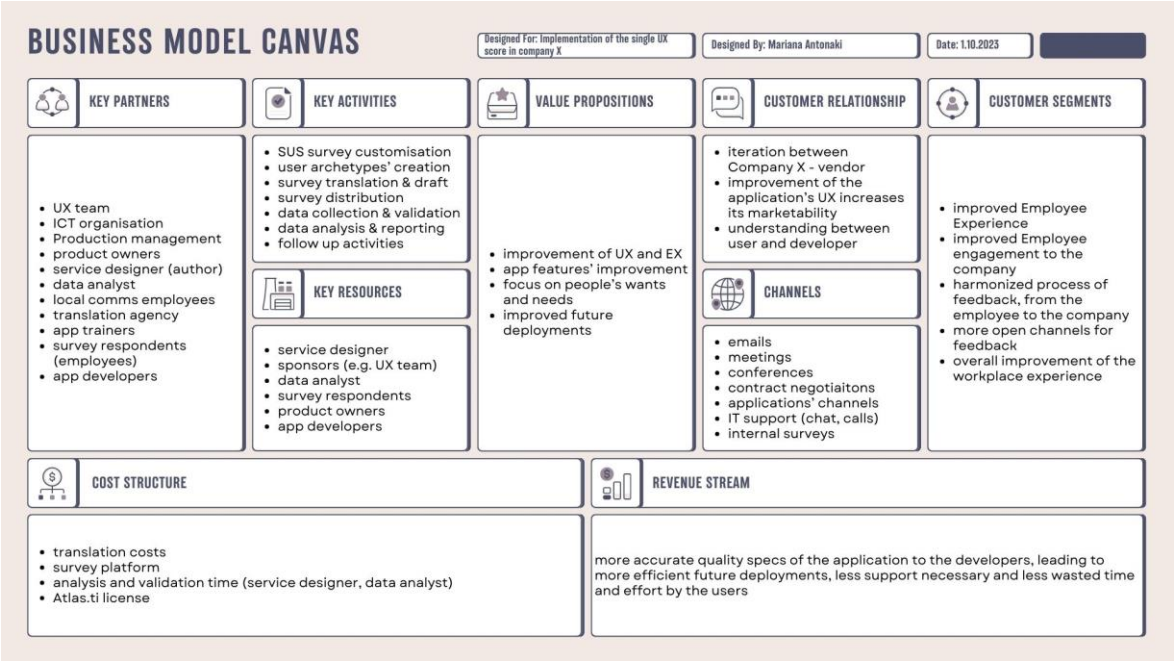


Figure 11. Business model canvas for company X.

In this canvas it is evident that there are several actors involved to the implementation process. All these stakeholders are working together to create new value, as seen in the value propositions' area. The activities are carried out mainly by the service designer and the data analyst during the preliminary phases (design – development – data analysis – reporting), but the product owners, vendor developers, and managers are crucial for the follow up activities, when actual planning and decision-making processes take over. The idea behind the single UX score concept is continual iteration between the application's users with the developers, the trainers, the managers, and all other groups involved with the development and smooth operation of the digital tool. Another important aspect of the concept concerns the cost structure and revenue stream. By implementing this method, the company dedicates minimal capital to the data collection (translation costs, service designer-data analyst hours, software licenses), but this minor investment generates great revenue, as it produces large amounts of high quality data through the survey, and a harmonized quality control method for the future, only by repeating the method consistently.

5.2 The Y app case

By following the ISO 9241-11 standards on software product usability (Usability partners, n.d.), initially the focus remained on the measurement of these attributes: effectiveness, efficiency, satisfaction. While these attributes showcase a system's usability, more aspects

had to be taken into account in this particular case. Since the Y application is an obligatory tool for employees working at the production sites, it was necessary to measure their ability to learn and memorize it, therefore the learnability (Lewis & Sauro, 2009) was the fourth attribute added to the study. The fifth attribute, universality, was added when it became apparent that the users should be somehow categorized, based on their work roles and technical skills. This general categorization helped in understanding the users' needs better, while taking into account their capabilities.

The SUS was redesigned to fit better the Company X's landscape, and an open comments' section was added, in an effort to acquire the users' ideas and recommendations in qualitative form. Thus, the end result consisted of the following:

- 4 introductory questions, necessary for defining the users' groups according to their professional roles and technical skills
- 10 questions identical to the SUS template with the same phrasing and order as the original SUS (Brooke, 1996)
- an obligatory open-comments section at the end of the survey, necessary for the respondents' direct input

Additionally, earlier interviews conducted by an external partner on a selected number of Y app users (20) in 2022 were utilized as secondary data, focusing on their assessments of a Y app mobile prototype. These interviews were summarised and their main findings regarding pain points, new opportunities, and general comments about the Y app were included in the SUS qualitative data and analysed as a whole in Atlas.ti. The reasons why the interviews were not separately analysed, was that there was no access to the primary data (recordings, transcriptions), therefore their findings could only serve as secondary data. Nevertheless, their findings were evaluated as particularly relevant to this pilot study, as these interviews were the only available source regarding the Y app's UX assessment in Company X.

In order to reinforce the holistic approach, a mini workshop on lessons learned was organised with an experienced Y app trainer, focusing on the SUS findings while also drawing new conclusions via a trainer's journey map and a training service's blueprint. This workshop is an indicative example of necessary follow up activities to the single UX score

process, in order to ensure that a changes roadmap and concrete action plan will be created, once the users have submitted their input via the SUS survey.

6 Findings

The single UX score concept proved to be a reliable quality metric in various ways. Firstly, the SUS questionnaire was a fast and efficient method to collect large amounts of data, with minimal costs. The survey received 154 responses in total, which – even though this was a large number – remained below the statistical margin error market standard, at 8% (Graglia, n.d.). Despite the relatively low participation to the survey, the results were insightful.

6.1 Numerical data

The single UX score for the Y app across user groups was found to be **40.37%.** This score clearly followed the SUS point system's findings, and confirmed that the Y application's user-friendliness is unacceptably low for the majority of its users (figure 12).

3.15	4.76	3.78	4.24	40.37 %
Learnability	Efficiency	Effectiveness	Satisfaction	UX single score

Professional role and Tech skills	Users	Q5Average	Q6Average	Q7Average	Q8Average	Q9Average	Q10Average	Q11Average	Q12Average	Q13Average	Q14Average	TotalAvg	Average score
Key User, Advanced	6	5.83	6.83	5.33	9.83	3.67	5.83	5.00	4.67	6.50	4.17	5.77	57.67
Key User, Intermediate	5	5.00	6.40	4.40	7.60	5.20	5.80	5.00	7.00	4.60	2.40	5.34	53.40
Key User, Basic	2	4.50	6.00	4.00	5.00	3.50	7.50	4.50	8.00	5.00	4.00	5.20	52.00
Blue collar employee, Advanced	15	3.07	4.33	3.13	8.33	3.13	3.73	4.60	5.40	4.20	3.67	4.36	43.60
Blue collar employee, Basic	7	4.86	4.29	3.14	6.29	3.14	4.57	3.86	4.71	3.29	3.00	4.11	41.14
Blue collar employee, Intermediate	25	3.04	3.80	2.88	7.96	2.68	4.36	3.80	4.08	4.40	2.44	3.94	39.44
White collar employee, Basic	34	3.06	5.35	2.94	6.12	2.62	5.00	2.97	4.91	3.24	2.15	3.84	38.35
White collar employee, Intermediate	38	3.05	4.61	2.61	7.26	2.68	4.66	3.24	4.45	3.05	2.47	3.81	38.08
White collar employee, Advanced	22	2.64	5.00	2.45	6.50	2.14	5.05	3.27	4.23	2.59	3.18	3.70	37.05
Total	**154**	**3.27**	**4.82**	**2.96**	**7.16**	**2.79**	**4.77**	**3.58**	**4.70**	**3.58**	**2.72**	**4.03**	**40.33**

Figure 12. The Y app single UX score by user group category.

The SUS survey contained two data sets: numerical and textual. The numerical part was calculated individually in two different ways, and the results were displayed in a Power BI dashboard. This method was selected because it is more efficient for comparisons among different digital tools' findings; therefore, it is an easy template to utilize in the future, when more digital tools will go through the single UX score assessment.

The first calculation of the SUS numerical data (10 questions) defines if a digital tool's user experience is below average (<50 points), passable (50-70 points), or good (>70 points). Then, each of the four attributes is calculated as an average (%) for each user group. Each question is represented by one of the four main attributes: satisfaction, efficiency, effectiveness, and learnability. Each attribute is weighed according to its representation in the SUS survey. As already mentioned, 2/10 questions measure learnability (20% weight) and 8/10 questions measure usability (effectiveness, efficiency, satisfaction) therefore these 3 attributes are weighed with 26.6% each.

Five out of ten questions have a positive phrasing, more specifically, questions 5,7,9,11,13. These are calculated by each answer's scale position minus 2 (x-2).

Five out of ten questions have a negative phrasing, questions 6,8,10,12,14. These are calculated by 10 minus each answer's scale position (10-x).

For the final score, all the points are summed up, divided by 2, and then multiplied by 2.5.

Figure 13. The single UX score's calculation formula.

According to J. Brooke, this point system classifies a digital solution in three categories:

<50 points: the tool's usability is unacceptable

50-70 points: the tool's usability is passable

70+ points: the tool's usability is good

By following this point system, a score for each user group was created. These scores varied from 57.67 points (the highest scored user group) to 37.05 points (the lowest scored user

group). Once the points for each user group were calculated, the Y app universality scale was ready. It demonstrated how each group's user satisfaction correlated with the others. Since the highest and lowest user group were more than 20 points apart, it was evident that the Y app's universality scale would be low. Additionally, since even the highest user group did not assess the Y app within the highest rating scale (70+ points), it meant that the tool's usability was placed below the passable threshold (Pearson, 2021).

Figure 14. The Y app's universality scale across all users' groups.

The universality scale, which was a new invention included in the single UX score concept, compares the UX score results among the previously identified user groups. When looking at the study's universality scale (Figure 12), it is obvious that for the majority of the respondents, the application is not user-friendly, regardless of the user's professional role and tech-savviness. As expected, key users with advanced technical skills gave the highest score to the Y app. However, even this group's mark was relatively low, despite their advanced expertise and engagement with the app in comparison with all the other user groups.

According to J. Brooke (1996) the passable threshold for a digital tool's usability must be at 50-70%, while the target should be at 70% or more. As seen from this universality scale (Figure 12), only three groups marked Y app's UX above 50%, and all of them compile the key users' groups. This result demonstrates beyond doubt that the average user does not see Y app as a user-friendly tool, and has significant difficulties while using it. Even the key

users do not give particularly high marks to its user-friendliness, even though they have received lengthy trainings on the tool, and many of them have even participated in the tool's development. The key users' assessment confirms the general view that the Y app has vast shortcomings regarding its usability and the overall user experience it provides is poor.

The attributes' calculation, which was the second layer of numerical data analysis, did not show a better picture (figure 15):

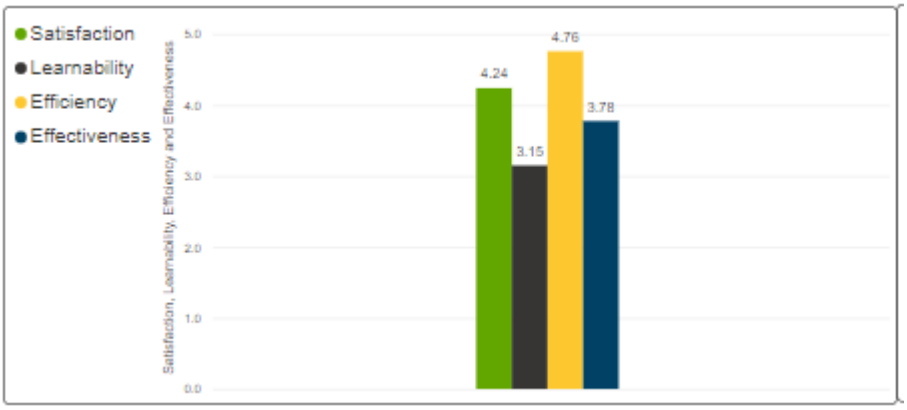


Figure 15. The average scores for all four attributes, Y app.

It is evident that none of the four attributes reached a high score, and all of them remained below the passable threshold (50%). An interesting observation was that learnability scored much lower than everything else, which indicates that there are significant problems with the training tools and methods, and the limited ability of users to work on the Y app independently. Efficiency, the highest scoring attribute, showed that the users somewhat confirm the tool's ability to deliver its tasks in a direct manner, without wasting time and efforts. However, the Y app's effectiveness received low marks, pointing the tool's overall architecture and whether its features were well-thought-out for the purposes they were designed to deliver. User satisfaction was mediocre – still below the passable threshold – and implied the poor User Interface (UI) design, as well as the overall image of the tool to its users.

All in all, the attributes' marks confirmed the first calculations. The most problematic area is definitely the ability to use the Y app independently, mainly due to poor training. What was observed from the universality scale (Figure 12) was corroborated by the attributes' calculations: users with more thorough training (key users) gave higher marks to the Y app, compared to the average users, who rely only on the existing training tools and methods.

6.2 Textual data findings (SUS questionnaire and interviews)

The textual data of the SUS survey was analysed in Atlas.ti. This software creates groups and codes of textual sources that can be then displayed as graphs, showing which are the most mentioned topics in the text. The creation of the codes and groups is manual, by the researcher, but it is possible to automate some parts of it, for instance by selecting a specific group and adding codes to it, or by creating new titles for segments, making it easier to spot similar references in the original data. All qualitative data were analysed with the software Atlas.ti, after they were translated from all source languages to English.

Recently, Atlas.ti implemented an AI coding functionality, which automatically creates groups and codes from the source text. This feature was also used during this pilot study, in order to assess its efficiency. Surprisingly, the AI codes matched most of the manual codes, and especially the main categories: problem areas, positive/negative/neutral feedback, challenges, and new ideas. Of course the tool was not entirely reliable, as it could not spot problems in the text's translation, and it did not recognize undertone such as irony. Nevertheless, the AI codes which matched the manual codes in at least 80% of the cases, were created in 4 minutes, whereas the manual codes took approximately 2 days to create, which showcases the potential that this tool opens for qualitative research in the future.

The SUS survey open comments were translated to English and compiled in one document, together with the 20 in-depth interviews. In order to divide the two sources, each interview was transcribed in such a way that it fits the main codes discovered during the textual analysis of the SUS open comments.

More specifically, since the only available data from the interviews was a summarized Miro board provided by the external partner who conducted them, all the content was categorized under major topics/codes:

- daily use
- overall Y app feedback
- problematic areas
- challenges
- new ideas and suggestions

The qualitative findings of the SUS survey and the 20 in-depth interviews conducted earlier (August 2022), had a double goal:

a. to confirm the numerical findings and enhance them with details and viewpoints expressed by the users

b. to draft a roadmap for future actions, scaled according to their urgency, required effort, and cost-effectiveness.

The most prominent finding was compiled in a list of pain points, according to the number of times mentioned by the users.

The main pain points found were:

1. Redundant features
2. UI design and features
3. Slowness and limited reliability
4. Lack of training

These four categories compiled hundreds of comments and interview answers, showing what matters the most for the Y app users. Evidently, the tool's efficiency and effectiveness scored the highest, by populating three out of four pain point categories. The users have problems with the Y app's structure, the paths followed in order to complete an action in the tool, and the overall UI which is very technical with too much information cramped in a small area, making it very difficult to navigate through and search for something specific.

Additionally, learnability made it to the list confirming what the numbers had shown: the provided training is of low quality, with poor material and not sufficient time allocated. The current training methods do not allow the user to familiarize with the tool sufficiently, and as a result many users either create workarounds to avoid using the Y app when possible, or memorize the bare minimum of the tool's features and only work within this framework, as often there is no time or resources dedicated to the improvement of the user's ability to operate the Y app independently to its full capacity.

An additional feature of Atlas.ti was to create visualizations of the qualitative input to the SUS survey and the interviews. The Sankey chart below depicts the feedback flow, from problems to opportunities. More specifically, the users' comments were divided into identifying pain points and suggesting new ideas. As visible in the chart, the negative and positive feedback are interconnected. A problematic area can be interpreted as a new opportunity, as challenges can be transformed into new ideas.

More specifically, reported problems regarding the app's connectivity issues, complex menus and work paths, as well as difficult UI with a lot of unnecessary information on display, directly link to new ideas on how the UI can be improved by becoming customizable, and how the menus and work paths can be re-designed to fit the specific tasks per work role instead of maintaining an unsuccessful one-fits-all system.

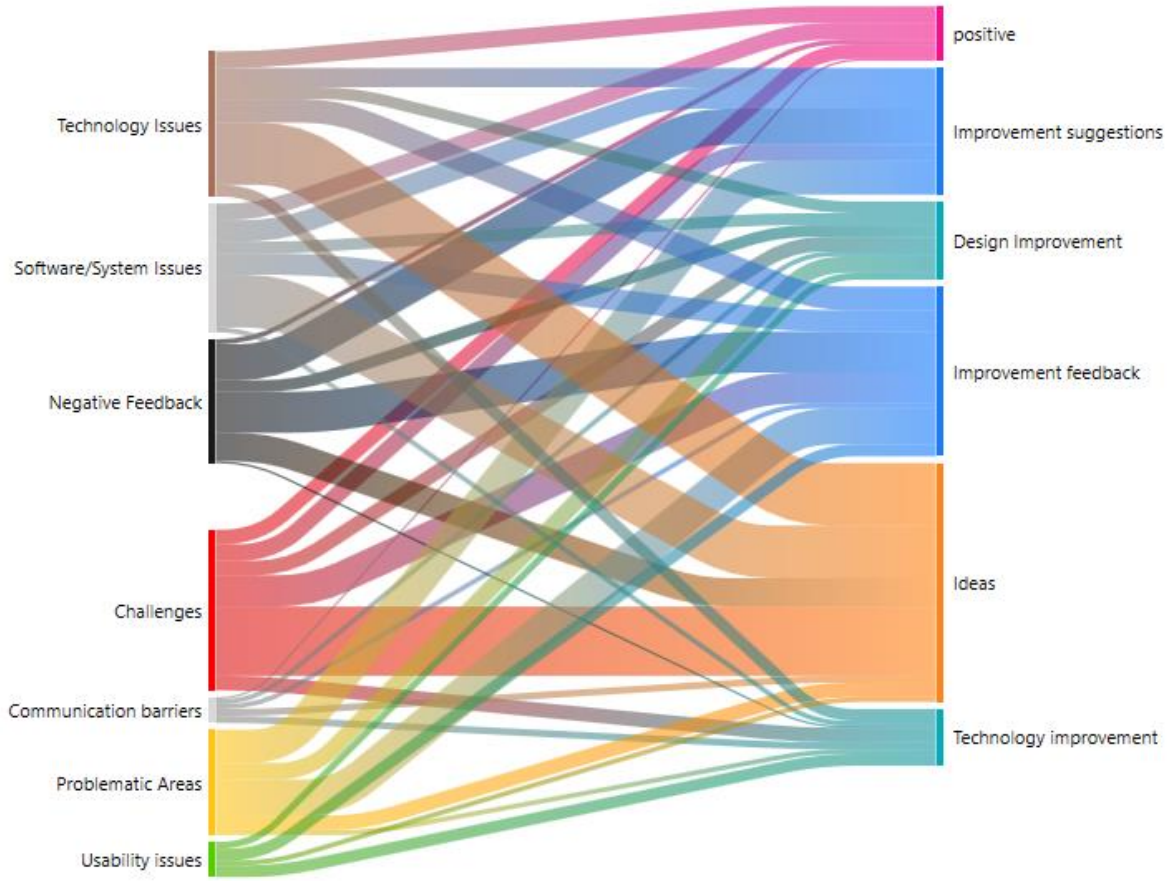


Figure 16. Sankey chart of comments' codes: from problems to opportunities.

The second goal of the qualitative data analysis, to build a future roadmap recommendation, was primarily based on the SUS and interview findings. However, additional aspects had to be taken into account: urgency, cost-efficiency, and required effort.

Urgency was defined by the users' comments, and constructed the points that should be included in the roadmap at all costs. Its structure relied heavily on the pain points' list mentioned above.

Cost-efficiency was a more complicated topic, which took in account different aspects: the Company X's plans for new Y app deployments, the budgets allocated to such activities per organisation and department, the commitments promised by the Y app vendor. Since the scope of this pilot study was not to create a detailed action plan for future changes in the Y app, but only to produce a recommended roadmap as realistically as possible, cost-efficiency was linked to required effort.

In order to define **required effort**, all stakeholders involved in the Y app change roadmap were mapped.

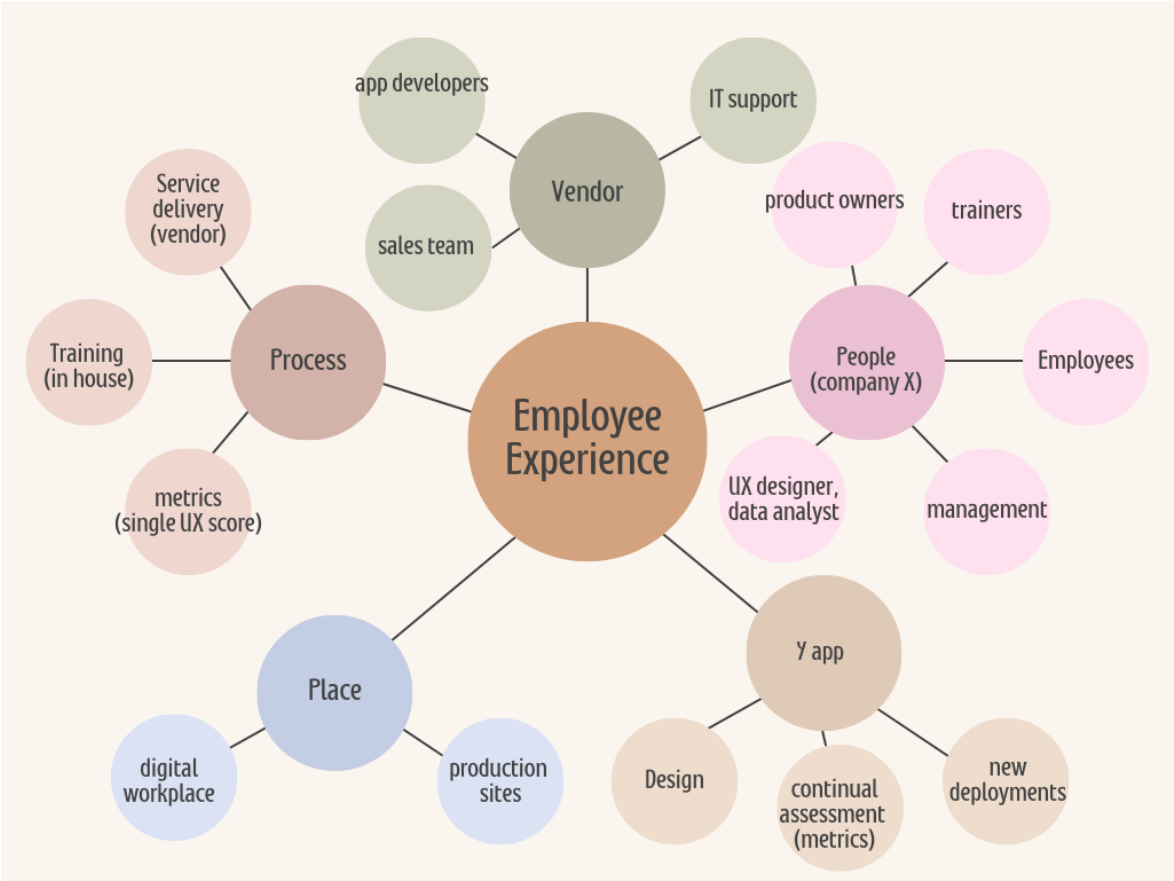


Figure 17. Employee experience ecosystem map.

The ecosystem map offers an overall idea about the involved people and teams in both parties (Company X and vendor company), and divides them in different circles. Starting with the central point, Employee Experience and its improvement is identified as the main

focus point for this new concept. The stakeholders involved belong either to company X or the vendor company, depending on their professional roles.

However, aspects such as the place of activity, the process itself, as well as the Y application were also depicted on the map, as they are important elements to focus on. Basically, this map visualizes the different actors affecting the employee experience, meaning that each of these bubbles has its own importance in improving the whole situation. Obviously, not all stakeholders possess the same influence, but each aspect has to be individually visited and seen as a unique component in the process of improving the employees' experience in company X.

The ecosystem map helped in identifying the required effort scale, which was critical for the creation of a realistic roadmap. For this thesis project, the required effort aspect is defined as the parties involved, time, and processes requirements in order to implement a change. For instance, lower effort changes involve the least parties, and deliver results in minimum time, without interrupting existing processes. Such changes would be simple fixes inside the tool related to connectivity capacity, or the improvement of training materials.

When all three aspects were identified (urgency – cost efficiency – required effort) the desired changes' plan started from those changes requiring the least resources across all of these aspects, and proceeded with the more impactful changes. The most pressing change as depicted across all data, requiring limited resources, and low effort, was to improve learnability. This change involved only the Y app trainers, and focused on the assessment and improvement of their materials and methods. Departing from this realization, a mini workshop was scheduled with one of the most experienced Y app trainers in house (chapter 6.3). The rest of the changes included in the roadmap were listed hierarchically:

1. Training improvements
2. Reliable performance
3. Menus & features of the Y app
4. Overall UI facelift

Points 2,3, and 4 require the involvement of both parties (Company X, Y app provider) and significant resources must be allocated for these changes. Some of the more specific changes listed under points 2 and 3 are scheduled to happen during the upcoming new deployments of the Y app (2023-2024), while others were observed by the users but not known to the app developers, therefore it is more costly and time consuming to proceed with those. These changes were summarized under point 4, which is crucial for the users as shown in the study, but not an easy or quick goal to achieve.

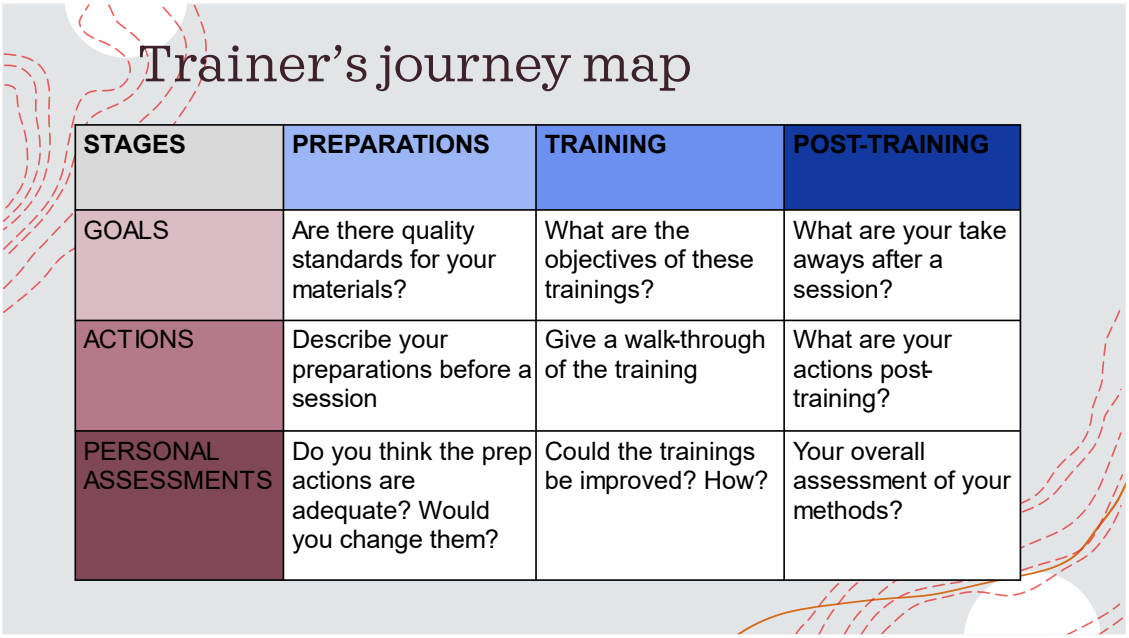
6.3 Mini workshop on lessons learned

Since learnability was continuously depicted as a crucial pain point for the Y app users across all user groups, it was interesting to explore how this attribute could be improved towards the right direction. Unfortunately, time and resources were very limited, therefore the solution of a single workshop session on lessons learned was organised with one of the most experienced Y app trainers in the company.

The trainer was provided with a summary of the pilot study findings focused on learnability, in advance. He saw how this attribute scored across all 32 production sites, and which were the main comments regarding training from the SUS survey and the earlier interviews from his own work location. The trainer was requested to complete two assignments before the workshop:

1. to prepare a scenario-based trainer's journey map based on a template, with the following scenario:

You have to introduce the Y app to a group of newcomers. They all belong to similar professional roles (you may choose them). Please describe the training process for them. It can be one or more training methods/sessions'.



STAGES	PREPARATIONS	TRAINING	POST-TRAINING
GOALS	Are there quality standards for your materials?	What are the objectives of these trainings?	What are your take aways after a session?
ACTIONS	Describe your preparations before a session	Give a walk-through of the training	What are your actions post-training?
PERSONAL ASSESSMENTS	Do you think the prep actions are adequate? Would you change them?	Could the trainings be improved? How?	Your overall assessment of your methods?

Figure 18. Template of the trainer's journey map

2. to complete a training service blueprint focused on information flow:

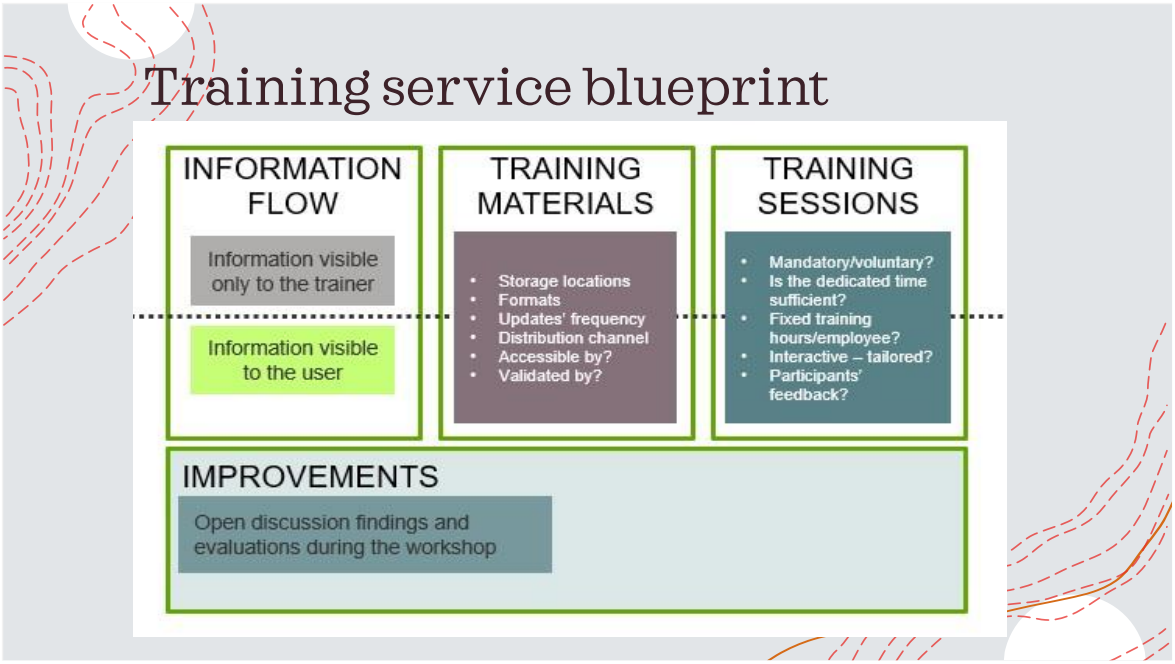


Figure 19. Template of the training service's blueprint

The purpose of this journey map exercise was to understand all factors influencing the user's experience while attending the training sessions, by identifying crucial touchpoints (Stickdorn et. al., 2011).

Additionally, the service blueprint exercise aimed to visualize both the user's and the trainer's perspectives during the training process (Stickdorn et. al., 2011). More specifically, the focus was on how the information flow between these two groups and how the resources could be optimized.

Based on this pre-work, during the workshop the discussion focused on the low scores of learnability, indicating serious issues with the internal training methods and processes. The trainer gave valuable insights on earlier training flaws and how he managed to tackle them. For instance, he has adopted a more 'custom-made' approach for his training sessions, dividing the participants into groups according to their roles and what Y app features they are expected to use the most. He delivers interactive trainings to said groups where the users work on given scenarios, replicating testing sessions, and resolving emerging issues on the spot. Additionally, he ensures that all necessary training materials are always accessible to the users in multiple locations and documentation systems.

STAGES	PREPARATIONS	TRAINING	POST-TRAINING
GOALS	***Are there quality standards for your material?*** Standard, self-created training material is used for newcomers in an introductory training session of 4 hours	***What are the objectives of these trainings?*** Explain to the new users how the digital tool works and what is expected from them as users. Provide basic competence to newcomers, and increase competence through new training rounds to more experienced users	***What do you learn from each training? Improvement points?*** Every topic should be explained slowly and in the simplest way, both by presenting and by having the users work on their computers
ACTIONS	***Describe your preparations before a session*** For new features/deployments, the first training is for managers. All existing material is checked and updated. Depending on the group of users, exercises are created for certain areas/features	***Describe what you do during a training session*** • Presentation of who I am and what's my role • Go through the training agenda • Group introduces themselves and their work areas • Explain the training's purpose • Encourage people to interrupt for questions • Training is interactive: everyone is on their computer following the instructions • After each area/part, time for questions	***What do you do after the training?*** Review questions/opinions and fix whatever is possible to do on my own. I make sure to document all pending questions, and answer to the user as fast as possible via their manager
PERSONAL EVALUATION / OPINIONS	***Would you change something to your preparatory activities?*** Yes and no. Sometimes you need to think outside the box and train users without relying on material, but if the group remains silent, without questions, it's good to have more material at your disposal	***Could the training sessions be improved?*** I pay attention to the participants' feedback and always improve the sessions/materials accordingly, but there is always room for improvements	***What's your overall assessment of your methods?*** It works well, it's efficient, but there's always room for improvement

Figure 20. The trainer's journey map filled out by the workshop participant.

In his view, the main problem with internal training – and thus learnability – is that very limited time is allocated to it during working hours and that the trainers do not receive standardized tools and materials to help them with their work. All in all, there is no unified approach to the employees' learning needs and paths, across all production sites therefore it is a usual occurrence to have very highly trained employees in one location but completely neglected training in another, which shows that the process mostly depends on the trainer's availability, experience, and access to good quality tools and materials.

While discussing the information flow and the potential improvements to the materials and the training methods, the workshop participant worked on a learning service blueprint provided by the researcher. This exercise helped the trainer outline the information flow between his role and the average Y app users, while ideating on potential improvements and new ideas (figure 21).

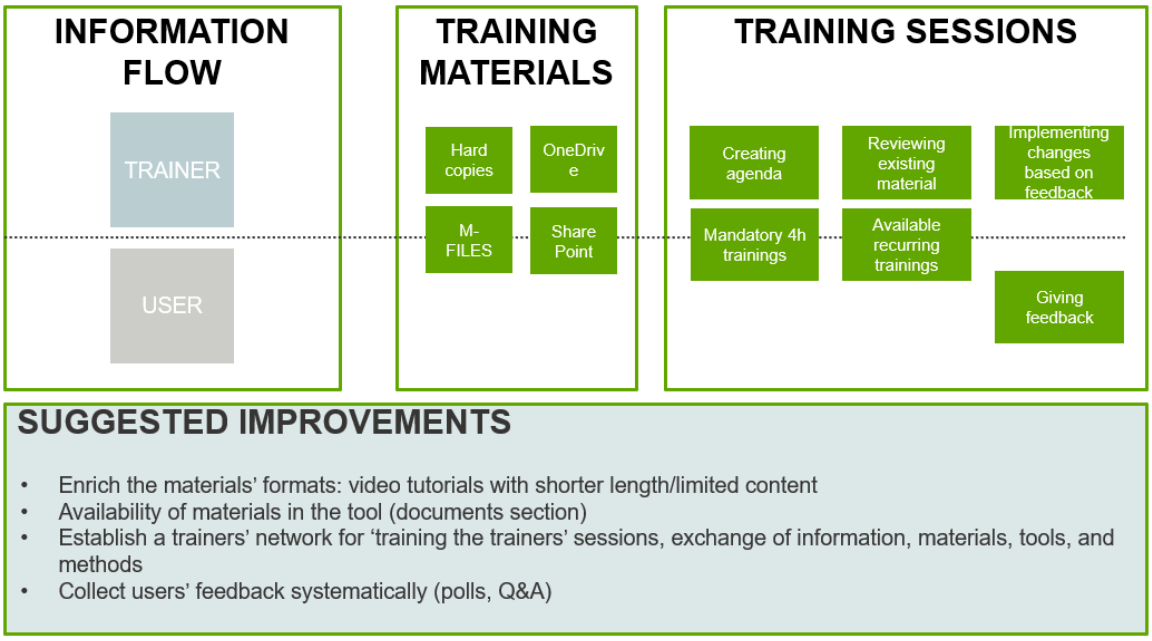


Figure 21. The training service blueprint, filled out by the workshop participant.

7 Discussion and future steps

This pilot study was a very interesting project, as it explored a relatively unknown area for Company X. Assessing the UX of internal users on specific digital tools quickly became a tangible goal, once the single UX score concept was completed. The commissioner allowed access to all of the Y app's userbase, which received positive feedback by the users and other stakeholders, since it was a very promising first step for improving the EX of internal digital tools in the future.

Throughout this project, the terms UX and EX have been used interchangeably. This is the result of equalizing the user with the employee in this project's context, which allows deeper exploration of the overall EX standards and directions together with UX principles.

Even though the main focus of the single UX concept has been the UX of the Y app during this pilot study, the general EX concept has emerged in multiple occasions.

7.1 Research questions and answers

Two main research questions were defined in the beginning of this study:

1. **How can a single UX score set the quality baseline for employee experience?**

The single UX score concept is generic therefore it can be customized – to an extent – according to the subject matter's needs. It provides solid data on a digital tool's user-friendliness, and can be applied to all digital tools without restrictions. This characteristic is what makes it a reliable quality baseline: all tools can be assessed with the same method so the results are comparable, and each tool can be assessed by this method continually, in order to record progress over time, when changes are implemented.

The concept is easy to comprehend, fast, and cost-efficient; this means that it can be deployed on very short notice, and that its findings can be immediately utilized for reporting to management, future road mapping exercises, and other activities.

Specifically for Company X, the single UX score concept has been a proven UX assessment method since September 2023. Soon after this pilot study's completion, 15 other internal digital tools ran the same survey on their own userbase. The SUS survey remained the same, and the only changes affected the user groups' categorization and subsequently each tool's universality scale. The scores produced for each of these applications were directly comparable with the Y app scores, and the qualitative input confirmed similar shortcomings; for instance, learnability remained consistently low in newer applications.

Additionally, this method adds value to the employee experience, as it shows the company's intentions to listen to its employees and take their input seriously. Of course, the next steps are critical in maintaining this image, because if the company does not implement any of the suggested changes and does not follow up consistently, then it is very easy for the employees to lose trust in the company, which can have a direct impact on the company's employee engagement figures.

2. **Why is the single UX score concept a reliable quality metric?**

As already established in this thesis, the concept relies on solid theoretical basis (SUS, ISO standards, user archetypes, double diamond, EFQM and innovation frameworks and employee experience) therefore its methodology can be directly traced to existing UX concepts.

Additionally, the single UX score relies on anonymous statistical and qualitative data, received through a survey. This method guarantees transparency, and equality, as no employee group or location is favored in comparison to others. For this pilot study, the survey was translated into eight languages and was distributed at all 32 production sites through the same marketing channels. This approach guaranteed objectivity and reliability on the results. The data was analysed in Power BI and Atlas.ti, which minimized personal biases by the data analyst and the researcher, as well as human error.

Also, the concept was used 15 more times successfully, replicating precisely the same process regarding data collection tools and analysis methods. The results produced by these later studies confirmed the concept's purpose and methodology. Therefore, it earned a spot on the EFQM metrics' list as a successfully tested metric.

7.2 Evaluation and future steps

The single UX score concept is designed in a generic and standardized manner, so it can adjust to different needs. The main concept remains simple and straightforward: how user-friendly is a digital tool according to its users?

7.2.1 Evaluation of the concept

The ideal application for this concept is the assessment of digital tools with pre-defined userbases. As large numbers of respondents are required for statistical validation, it is important to distribute the SUS survey across all the userbase, and to eliminate biases and statistical margin errors as much as possible. Corporate digital solutions are ideal candidates for the single UX score concept, as their user groups and userbase can be easily defined. On the contrary, when a tool has multiple functionalities and each user group uses it in a different way, then this concept cannot be equally reliable, unless undergoing significant customization.

An additional point requiring attention is the users' categorization. This process should be well-thought-out, as much of the method's findings rely on its universality scale, which is impossible to create if user groups are not properly identified in the beginning of the project. The recommended approach to this riddle is to create user archetypes, which can divide the different users into wider groups without compromising their opinions on the tool's UX.

Lastly, the concept relies solely on self-reported data by the users. This can be seen as a shortcoming for UX designers and UI developers who base their assessments on system usability studies and reports directly from the tool by measuring errors per action, action completion times, and more (Sauro, 2012, May 30). Service designers have a different perspective though.

In Service Design, qualitative data is a valuable source of information, and all of its methods rely on people's views. As a service designer, I opted to understand the users' needs and wishes in their own words, by adjusting a purely quantitative method to service design standards. Including an obligatory open comments' section at the end of the SUS survey was the only way to receive the respondents' opinions fast and in bulk. Since the pilot's requirement was to deliver this project quickly, and the single UX score concept was marketed as a method to collect reliable data quickly and efficiently, there were no resources for following more traditional Service Design methods, such as workshops and focus groups. However, even this limited source of qualitative data was not only able to corroborate the numerical findings but also to produce a roadmap for future changes and a map of pain points, both of which are very useful for the tool's future improvements.

7.2.2 Future steps

As already mentioned, the single UX score is a new EFQM metric for one of the Company X's Business Areas, since September 2023. The plan is to apply the concept to several more internal digital tools in different Business Areas, by the end of 2024, and continue with annual repetitions of the SUS survey in order to assess the UX continuity across internal digital solutions, renegotiate deliverables with vendors, and establish a high-quality UX assessment process for all employees.

Outside company X, the single UX score concept can be utilized in similar setups: digital tools with pre-defined userbases can be assessed with minor customizations to the survey – solely focused on identifying the tool's main user groups.

The concept is promising, as it utilizes UX tools with a Service Design twist. Usability studies usually rely on automatic reports and scenario testing sessions, but this method adds the user's viewpoint to the mix. When qualitative data is added, the users feel heard. Ideally, workshops can follow, with specialized stakeholders discussing only certain aspects of the findings, thus providing further input to the necessary changes and improvements. These sessions guarantee that there are measures taken towards the improvement of the UX, and can act as mediators between the digital solutions' UI architects and the average user, which is a very frequent pain point vastly affecting the user's experience with digital solutions.

7.3 Ideas for further research

The Employee Experience terrain is very interesting and relatively unexplored so far. As observed during this pilot project, the employees' input on several occasions can prove to be valuable, providing straightforward answers to persistent problems. The renewed bottom-up approach offered by EX in comparison with earlier approaches such as engagement, allows direct iteration between the employee and the employer, and several models have been already developed in this direction, such as the holistic employee experience (HEX) and the digital employee experience (DEX) (Bridger & Fannaway, 2021).

Based on these EX models and principles, further research could focus on the overall employee experience with digital tools during their development and implementation. An interesting approach by Jon Yablonski (2021) attempts to link users' behavioral patterns with established psychological models, and outline optimal design tools and methods for improved UX.

For instance, a new metric concept could assess a digital tool against Jakob's law, which states that users will be positively predisposed to familiar sites and applications, and could suggest improvements accordingly. In a similar direction, a new training process for digital tools could be based on Miller's Law, which states that the average person has room for 7

+/- 2 items in their working memory, and design the learning process in a user-friendlier way.

The possibilities for further research combining UX, EX, and service design in the workplace are endless, nevertheless the core remains identical: human centricity.

8 References

Albert B., Tullis T., (2013); *Measuring the User Experience : Collecting, Analyzing, and Presenting Usability Metrics*, Elsevier Science & Technology. Accessed on 30.3.2023

Bridger E., Gannaway B. (2021) *Employee Experience by Design. How to create an effective EX for competitive advantage.* Kogan Page Limited: GB, USA

Brooke, J., (1996): SUS: a quick and dirty usability scale. *In* Usability evaluation in industry (pp. 189-194) . London: Taylor & Francis.

Design Council (n.d.) Framework of Innovation. Retrieved on 1.10.2023 from https://www.designcouncil.org.uk/our-resources/framework-for-innovation/

EFQM model (2020), European Foundation of Quality Management. Accessed on 13.9.2023

Emmett, J., Komm, A., Moritz, S., Schultz, F. (2021, September 30); This time it's personal: Shaping the 'new possible through employee experience. *In* McKinsley & Company. Retrieved 10.8.2023 from https://www.mckinsey.com/capabilities/people-and-organizational-performance/our-insights/this-time-its-personal-shaping-the-new-possible-through-employee-experience .

European Union (n.d.) Aims and Values. Retrieved 19.9.2023 from: https://european-union.europa.eu/principles-countries-history/principles-and-values/aims-and-values_en

Gheidar Y., ShamiZanjani M. (2020) Conceptualising the digital employee experience, *Strategic HR Review*, 19 (3) pp. 131-135

Graglia, D., (n.d.) *How many survey responses do I need to be statistically valid? Find your sample size* [Blog post]. Retrieved 4.4.2023 from https://www.surveymonkey.com/curiosity/how-many-people-do-i-need-to-take-my-survey .

Kallio, S. (2021, November 17), *Employee Experience and Enterprise Service Management* [Blog post]. Retrieved 1.9.2023 from

https://www.servicedeskinstitute.com/employee-experience-and-enterprise-service-management/ .

Laubheimer, P. (2022, May 15), *Personas vs Archetypes* [Blog post]. Retrieved 1.9.2023 from https://www.nngroup.com/articles/personas-archetypes/ .

Lewis, J. R. & Sauro, J., (2009); The factor structure of the system usability scale. Proceedings of the human computer interaction international conference (HCII 2009) , San Diego CA, USA.

LeSuer, S. (2022, March 3) *Measuring usability metrics: 6 test metrics & free scorecard* [Blog post]. Retrieved 3.3.2023 from https://slickplan.com/blog/measuring-usability-metrics .

Morgan, J. (2015, December 15) The Three Environments That Create Every Employee Experience. *In* Forbes. Retrieved 1.9.2023 from https://www.forbes.com/sites/jacobmorgan/2015/12/15/the-three-environments-that-create-every-employee-experience/?sh=2139db6e66c6 .

Nair Sanjay; (2020). *The Service Desk handbook. A guide to service desk implementation, management and support.* IT Governance Publishing.

Pearson, C. (2021, January 31) *The Single Usability Metric (SUM) – A completion rate conundrum.* [Blog post] Retrieved 2.3.2023 from https://uxdesign.cc/the-single-usability-metric-a-completion-rate-conundrum-c6fd0e5a54ca .

Sauro, J. (2012, May 30) *10 things to know about the single usability metric (SUM)* [blog post]. Retrieved 15.2.2023 from https://measuringu.com/SUM/ .

Stickdorn M, Schneider J, et al. (2011) This is Service Design Thinking. BIS publishers: The Netherlands.

Stickdorn M, Hormess M, Lawrence A, Schneider J, (2018) This is Service Design Doing. O'Reilly Media: Canada.

Stickdorn M, Hormess M, Lawrence A, Schneider J, (2018) This is Service Design Methods. O'Reilly Media: Canada.

United Nations, Sustainable Development Goals (n.d.). Retrieved 13.9.2023 from https://www.un.org/sustainabledevelopment/sustainable-development-goals/ .

United Nations, Global Compact (n.d.). Retrieved 13.9.2023 from https://unglobalcompact.org/ .

Usability Partners (n.d.), ISO 9241-11 Quality standards for usability *In* Usability Standards. Retrieved 1.4.2023 from https://www.usabilitypartners.se/about-usability/iso-standards.php

Whitter David, (2022); Employee Experience, 2nd edition. Kogan Page Limited: USA.

Yablonski, J (2020) Laws of UX. Using Psychology to Design Better Products & Services. O'Reilly Media.

List of figures