Tampere University

Pietari Pakarinen

CONSTRUCTING A CUSTOMIZABLE USER EXPERIENCE EVALUATION METHOD

To improve digital employee experience

Master of Science Thesis
Faculty of Information Technology and Communication Sciences (ITC)
Examiners: Prof. Jussi Okkonen
May 2024

ABSTRACT

Pietari Pakarinen: Constructing a customizable user experience evaluation method
Master of Science Thesis
Tampere University
Human-Technology Interaction
May 2024

Patria has been experiencing rapid growth, which has led to their software catalogue becoming bloated. To cut down on the bloat, Patria is working on several projects. For example, Patria started an application landscape rationalization process in the February of 2024. The goal of the process is to strategically align applications with the business goals of Patria, optimize capabilities, and ensure a more unified and cost-efficient application landscape.

Patria had recognized that the rapid pace of application landscape changes was both a threat and an opportunity. If not responded to appropriately, the changes would deteriorate the digital employee experience, which was already in a concerning state. On the other hand, if the changes were made with a focus on user experience, digital employee experience would likely improve rapidly as well.

The purpose of this thesis is to construct a tool for creating comparable evaluations of user experience of software. These evaluations can be used as part of decision making in projects like the application landscape rationalization, as well as any other software change Patria may do.

As a result of literature review, requirements engineering and collaboration with stakeholders, I constructed a customizable user experience evaluation method to be used in any kind of software evaluation situation Patria may face. Using it produces a score based on a selection of relevant elements of experience that can then be used to compare the software to the scores of other software that have been evaluated.

Patria is also seeking to improve their digital employee experience. In this thesis, based on a literature review and expert interviews, I define a road map for methodologically improving the digital employee experience from its current state. In the road map, I recommend that Patria continues to work on improving user experience of the software it uses by utilizing the user experience evaluation method constructed as part of this thesis. In addition, I recommend hiring or upskilling employees and creating a DEX special interest group.

This thesis consists of three parts. A literature review, case studies, and the user experience evaluation method. The literature review has the goals of gathering the requisite knowledge to create a road map for improving the digital employee experience of Patria, as well as design a user experience evaluation method appropriate for Patrias' needs. In the case studies section I discuss various cases that were researched as part of this thesis, as well as case Efecte, which serves as a proof-of-concept as well as a weak market test for the user experience evaluation method.

Keywords: Digital employee experience, User experience evaluation, Software consolidation, Method construction, Application landscape rationalization

The originality of this thesis has been checked using the Turnitin OriginalityCheck service.

TIIVISTELMÄ

Pietari Pakarinen: Mukautettavan käyttäjäkokemuksen arviointimetodin konstruktointi
Pro gradu -tutkielma
Tampereen yliopisto
Human-Technology Interaction
Toukokuu 2024

Patria on kokenut nopeaa kasvua, mikä on johtanut heidän ohjelmistokataloginsa paisumiseen. Patrialla on käynnissä ja käynnistymässä lukuisia projekteja, joilla pyritään karsimaan ohjelmistojen päällekkäisyyksiä. Esimerkiksi helmikuussa 2024 Patria aloitti sovelluskentän rationalisointiprosessin. Prosessin tavoitteena on strategisesti kohdistaa sovellukset Patrian liiketoiminnan tavoitteisiin, optimoida kyvykkyydet ja varmistaa yhtenäisempi ja kustannustehokkaampi sovelluskenttä.

Patria oli tunnistanut, että sovelluskentän nopeat muutokset olivat sekä uhka että mahdollisuus. Jos muutoksiin ei reagoitaisi asianmukaisesti, ne heikentäisivät digitaalista työntekijäkokemusta, joka oli jo huolestuttavassa tilassa. Toisaalta, jos muutokset tehtäisiin keskittyen käyttäjäkokemukseen, digitaalinen työntekijäkokemus todennäköisesti parantuisi nopeasti.

Tämän tutkielman tarkoituksena on rakentaa työkalu, jolla voidaan luoda vertailukelpoisia arviointeja ohjelmistojen käyttäjäkokemuksesta. Näitä arviointeja voi hyödyntää projekteissa, kuten sovelluskentän rationalisointi, osana päätöksentekoa. Arviointeja voi myös hyödyntää kaikissa muissakin ohjelmistomuutoksissa, joita Patria tulee tekemään.

Kirjallisuuskatsauksen, vaatimustenmäärittelyn ja sidosryhmien kanssa tehdyn yhteistyön tuloksena rakensin mukautettavan käyttäjäkokemuksen arviointimenetelmän, jota voidaan käyttää kaikissa ohjelmistoarviointitilanteissa, joita Patria tulee kohtaamaan. Menetelmä tuottaa tuloksena pistemäärän, joka perustuu valittuihin relevantteihin kokemuselementteihin. Pisteytyksiä voidaan hyödyntää ohjelmiston vertaamiseen muiden ohjelmistojen saamien pisteytysten kanssa.

Patria pyrkii myös parantamaan digitaalista työntekijäkokemustaan. Tässä tutkielmassa määritän kirjallisuuskatsauksen ja asiantuntijahaastattelujen perusteella tiekartan digitaalisen työntekijäkokemuksen metodologiseen parantamiseen sen nykyisestä tilasta. Tiekartassa suosittelen, että Patria jatkaa ohjelmistojensa käyttäjäkokemuksen parantamista käyttämällä tämän tutkielman osana rakennettua käyttäjäkokemuksen arviointimenetelmää. Lisäksi suosittelen palkkaamaan tai kouluttamaan työntekijöitä ja luomaan DEX-erityisryhmän.

Tämä tutkielma koostuu kolmesta osasta. Kirjallisuuskatsaus, tapaustutkimukset ja käyttäjäkokemuksen arviointimenetelmä. Kirjallisuuskatsauksen tavoitteena on kerätä tarvittava tieto DEX-tiekartan luomiseksi sekä käyttäjäkokemuksen arviointimenetelmän konstruktoimiseksi siten, että se vastaa Patrian tarpeisiin. Tapaustutkimusosiossa keskustelen eri tapauksista, joita tutkittiin osana tätä tutkielmaa, sekä tapauksesta Efecte, joka toimii sekä konseptin validointina että heikkona markkinatestinä käyttäjäkokemuksen arviointimenetelmälle.

Avainsanat: Digitaalinen työntekijäkokemus, Käyttäjäkokemuksen arviointi, Ohjelmistokonsolidaatio, Menetelmän konstruktointi, Sovelluskentän rationalisointi

Tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck -ohjelmalla.

PREFACE

This master's thesis wraps up a personal journey the destination of which admittedly, as often is the case with personal journeys, changed along the way. I was pushed onto this particular journey by Atte Haarni, a friend of mine who recommended that I try out programming. I ended up liking it so much, that I figured I'd get a degree in it.

However, while I did get my bachelor's degree in computer science, during it I discovered a fascination in human-technology interaction thanks to Don Norman's Design of Everyday Things and the wonderful professors at Tampere University. I ended up changing my master's degree to be in human-technology interaction, rather than the programming aspect of software development. I'd like to extend my thanks to all the people involved, as I wouldn't have written this thesis if not for them. Of course, my fellow students and friends were also indispensable for completing the various steps along the way before this final challenge. Thank you!

I'd also like to thank Patria for commissioning this thesis from me. It was a great show of faith and a fantastic opportunity for me. This topic was much more interesting than whatever I could have come up with on my own, and the support and resources provided were greatly appreciated. In particular I would like to thank Maria Murto, Janne Mäenpää, and Jarkko Marjasalo, who gave me feedback and guidance on the thesis every Monday. Who knows how off the mark this thesis would have been without your input. I'm also thankful to all my colleagues for their support and sharing their thoughts and ideas on the thesis with me.

Finally, I would like to thank my thesis supervisor, Jussi Okkonen, and seminar supervisor, Thomas Olsson, who provided guidance that shaped this thesis more profoundly than I could possibly have expected.

Tampereella, 1st May 2024

Pietari Pakarinen

CONTENTS

1. INTRODUCTION

This thesis has been done for Patria Group IT. It's one of the group functions of Patria Oyj, which is owned by the State of Finland (50.1%) and by Kongsberg Defence and Aerospace AS (49.9%). Patria Group consists of the parent company, Patria Oyj, and its subsidiaries. Patria also owns 61.8% of Millog Oy and 50% of Nammo AS, a Norwegian company, 60% of Milworks OÜ in Estonia and 49.9% of Kongsberg Aviation Maintenance Services AS in Norway. [1]

Patria is an international provider of defence, security and aviation life cycle support services, pilot training and technology solutions. Patria provides its aerospace and military customers with equipment availability, continuous performance development as well as selected intelligence, surveillance and management system products and services. Patria has several locations including Finland, Sweden, Norway, Belgium, the Netherlands, Estonia, Latvia and Spain. The net sales totaled EUR 627.1 million in 2022, and Patria employs over 3,000 professionals. [2]

Due to its size and a rapid pace of growth, Patria has seen new software added to its portfolio quite often for quite a long time. There is also always internal demand for new software to meet emerging needs. This has naturally led to a bloated application landscape, with a lot of overlapping applications. These applications require maintenance as well as license fees.

Motivated by these factors, in the February of 2024 Patria started an application landscape rationalization process. The goal of the process is to strategically align applications with the business goals of Patria, optimize capabilities, and ensure a more unified and cost-efficient application landscape.

The application landscape rationalization process began with the analysis of applications. The analysis consisted of interviews with the people using the applications and working to understand the requirements for operative work. Once the analysis was done, the target state and road map would then be defined by assessing the situation at the time, identifying problems, and addressing new demands.

For the application landscape rationalization, or software consolidation as I will henceforth refer to the process as, and other software evaluation projects, Patria required ways to evaluate and compare applications. When these projects are executed with the user ex-

perience of the software in mind, they can lead to an improvement in the digital employee experience.

This thesis aims to aid in that endeavor by creating a custom user experience evaluation method for software consolidation, which will then be one factor Patria takes into account when doing software consolidation. The method will also be useful for evaluating the user experience of applications for other purposes.

1.1 Background

This thesis was inspired by Patria Group IT's desire to get comparable evaluations of the user experiences of software. They hoped that by focusing on the user experience of software when doing software consolidation and procurement, they could improve their Digital Employee eXperience (DEX). Gartner has defined DEX as employee's overall experience with technology. They note that some call this the "end-user experience" or the "user experience." [3]

There was a concern that the current state of DEX at Patria wasn't adequate, and might further deteriorate if not managed in the face of many upcoming changes to the application landscape at the company. There was a fear of various kinds of bottlenecks, waste and poor quality software affecting the efficiency of work.

While they knew the issues they wanted to identify and analyze, the group function had no user experience design professionals, and thus the user experience maturity level of the organization was low. This led to some difficulties with requirements gathering for the project, as well as when defining the goals. It was only much later that I was told, that Patria has used Morville's honeycomb as some sort of a user experience guideline in the past. This indicated, that in some other part of the organisation UX maturity may be at a higher level. [4]

Rather than starting a major project to improve their User eXperience (UX) maturity, Patria decided to borrow existing expertise. Thus, Patria Group IT decided to hire a thesis worker to figure out how to do the evaluations they need and improve their DEX. I was hired and I set out to do my task. Despite the challenges, the personnel of Patria were supportive, approachable and encouraging, and the thesis project was a pleasure to do.

The reason Patria desired to find a way to improve their DEX was because they had many ongoing and upcoming projects relating to changing the software they have in use. They had determined, that while they were doing that, they should change the software they use to software that has a good user experience. This could even be extended to new application procurement. Patria, however, didn't have a method for systematically evaluating user experience.

At first, the goal of improving DEX seemed somewhat misguided. I had not heard of DEX, and to me it sounded more like what they really wanted to evaluate and improve was the user experience of software they were using or going to procure and improving their DEX would be some separate matter. I needed to understand why Patria had chosen to approach the challenge the way they had.

1.2 The Goals Of The Thesis

My task turned out to be two-fold. Firstly, to create the evaluations Patria wished for, I needed to create a user experience evaluation method that would meet their requirements. Secondly, I was assigned the task of researching what exactly is DEX, if implementing something called a DEX-model in Patria would be a good enough way to improve it and if so, how should such a model be implemented.

The first task I was well equipped to tackle. Patria even provided potential use case ideas on which I could test whatever user experience evaluation I construct. The first use case candidate was to do a user experience evaluation, using the method I constructed, to evaluate the user experience of a internal employee feedback tool. However, an even more relevant case was quickly pivoted to. The second case I worked on was about evaluating and comparing two different internal employee feedback tools. Once that would have been done, the results could have been used to compare the two tools to each other.

The second case seemed to have some potential synergy with the DEX-related secondary task I had. It was related to user experience, employee experience and software consolidation. According to research by Gartner, adapting new tools and features may overwhelm the service desk and other support channels, if communications aren't simultaneously improved. Assessing how the employees feel about such changes and gathering feedback, among other things, will help. [5] The internal employee feedback tools would clearly be a potential way to improve the communications between service desk and the rest of the organization. This will naturally become increasingly crucial to Patria due to the software consolidation project and other such software related change-projects.

However, neither of the tools were implemented yet, and aren't expected to be out of pilot testing use by the time I finish my thesis. In the state they are in, they aren't usable as a case study, but perhaps they could be evaluated using the method once it has been published alongside this thesis. I did work on the case for some time, before it became clear that the pace of my thesis was far too fast for the case to be usable.

Instead, the case study ended up being about ticket handling using Efecte ESM, which is an enterprise service management tool by the software company Efecte. Evaluating the user experience of Efecte and considering the options for which software to change to is

among the many upcoming projects related to changing the software they have in use at Patria that I mentioned.

While I felt well equipped to handle the second task, the second task required a fair bit of work to make sense of. The idea of improving DEX and adopting a DEX-model at Patria seems to have come from Gartner. The consulting company Gartner has been providing consulting services for Patria Group IT for some time now. While Gartner has provided Patria a fair bit of consulting and research materials, they do not have a DEX-model. Instead, I would have to figure out some other way to give Patria recommendations on how to improve the DEX at the company.

I decided to start by conducting a literature review, as I did have a lot of literature provided by Gartner to review. Even in the possible case that the literature review might not provide the answers Patria seeks, it should at least help ask the right questions from Gartner.

To reiterate and summarize the second task, Patria is looking to improve its digital employee experience. Patria is trying to understand where it currently stands on its' DEX-path, and what are the next steps forward on the path. Thus, Patria has asked me to develop a road map for improving DEX at Patria.

This thesis aims to gather the requisite knowledge via a literature review to create this DEX road map for Patria. For this purpose, I also have some access to Gartner's consultants as well as Patria personnel familiar with the subject. As Patria has yet to implement any method for improving DEX, it won't be too late to decide how to go about it once they have the road map to reference.

1.3 Research Questions

In order to fulfill the goals of this thesis, I need to conduct research on user experience evaluation methods, software consolidation and improving digital employee experience. Through various meetings, messages and other discussions with Patria representatives and my thesis supervisor, the research problem has been condensed into the following two research questions:

- How to design, construct and test a custom user experience evaluation method to aid in software consolidation and procurement?
- What should Patria do to methodologically improve their digital employee experience?

The user experience evaluation method will be used to give software a score that can be compared with other software's scores for various purposes. For example, for software consolidation. It should be noted, however, that there will be many other aspects that will be considered as well when the software consolidation is performed, with the score being

only one of them. Thus, this thesis will not be defining any other aspects of the software consolidation process, and how Patria chooses to use the score in their process is up to them.

As for the DEX road map, I have access to a vast pool of research on the subject by Gartner, as well as all the related scientific literature. DEX-tool testing will be limited by the budget Patria is willing to allocate to DEX-improvement. During this thesis I will have some access to Gartners' DEX-consultants, but no other budget.

1.4 Research Methods

The goals of this thesis are achieved by constructing a method to meet a particular business need of the customer company Patria. This method can be classified a construction. Kasanen et al. define constructions as entities that solve problems that emerge in running business organizations. According to their article, an important characteristic of constructions is that their usability can be demonstrated through implementation of the solution. [6]

To solve real world problems through research, Kasanen et al. introduced the constructive approach which means problem solving through the construction of organizational procedures or methods. This research procedure is particularly well suited for this thesis. The results of adopting the method created through my research, which could also be used to validate the method, may not become apparent until years later. Thus, demonstrating the usability of the method through its adoption is a more suitable validation method. This validation method is known as a weak market test. [6] Kasanen notes, that even if the weak market test fails, there is a need for theoretical analysis, for it may be possible that a similar failure might be avoidable in a different context, a different organization. In addition, such analysis will shed light on what are the changes that could potentially make the construct a working solution. [7]

Kasanen et al. note, however, that not all problem solving exercises pass as constructive research. An essential part of the constructive approach is to tie the problem and its solution with accumulated theoretical knowledge. The novelty and the actual working of the solution have to be demonstrated as well. [6] In this thesis I will ensure that these demands are met. Kasanen notes in his later writings, that even constructive research that fails in practice still has academic significance. [7]

The constructive approach may be characterized by dividing the research process into phases, the order of which may, of course. vary from case to case:

1. Find a practically relevant problem which also has research potential.
2. Obtain a general and comprehensive understanding of the topic.

3. Innovate, i.e., construct a solution idea.
4. Demonstrate that the solution works.
5. Show the theoretical connections and the research contribution of the solution concept.
6. Examine the scope of applicability of the solution.

[6]

Patria has provided a practically relevant problem which also has research potential, as defined in my research questions.

To tie the problem and its solution with accumulated theoretical knowledge, I will be conducting a **literature review**. Through the literature review I will also aim to:

- Gather the knowledge based on which I will design the method for Patria to evaluate user experience of software.
- Gather the requisite knowledge to create a DEX road map.

During this thesis, I will be meeting with a team of leaders from Patria group IT function once a week. In these meetings, I will update them on the progress I have made, explain why I have done what I have done as well as define the goals for the upcoming week. They will in turn help me align my research efforts with the needs of Patria group IT function. Based on the literature review and these discussion, I will innovate a solution idea.

Finally, I will be conducting **A case study**. The purpose of the case study is to serve as a demonstration of the solution working in practice, test and refine the solution as well as provide useful data to meet a practical need of Patria. After this, I will be able to examine the scope of the applicability of the solution.

1.5 The Structure Of The Thesis

The thesis is divided into literature review (chapter 2), case study (chapter 3), user experience evaluation method (chapter 4), as well as conclusions (chapter 5). Chapter two discusses gathering the theoretical background for the user experience evaluation method as well as the DEX road map. In the third chapter I talk about the case studies related to this thesis, mainly about the case Efecte. Chapter four is about the process of creating and evaluating the user experience evaluation method. Chapter five is a summary of the key findings of the research as well as a chapter for reflecting on the work done and its' implications for future work.

2. LITERATURE REVIEW

This literature review was conducted for the purposes of gathering the requisite knowledge to construct a user experience evaluation method for the needs of Patria and creating a DEX road map. Chapters 2.1 through 2.3 discuss the theoretical background on which I built the customizable user experience evaluation method. Chapter 2.3 discusses digital employee experience and includes a short road map for methodologically improving it at Patria.

2.1 User Experience

In the UX white paper, the authors wrote, that user experience (UX) needs to be accepted as a valued part of the overall design and development effort of an organisation. [8] Patria has also recognized this, as exemplified by them commissioning this thesis from me. The various software used by Patria group IT are changing rapidly, but there has thus far been a marked lack of user experience designer involvement, despite much of the changes being about software and due to user experience challenges.

To give an example of the importance of UX, I'll talk about how it affects data quality at Patria. It is known within Patria, that employees postpone, skip or work around phases of processes that have poor user experience, which directly leads to data quality deterioration. This kind of user experience induced deterioration in data quality had been found in, for example, the study done by Butler et al. In their study, they found that modifying the user interface of a Electronic Clinical Outcome Assessments -tool to better accommodate the clinical setting it was used in led to a statistically significant reduction in error rates. The reduction in error rates, as well as ensuring standardized administration and scoring significantly improved data quality. [9]

To solve this, managers at Patria ask employees to think of data quality, and a master data management team has been created for the purposes of data quality assurance. Meanwhile, the solution to the challenging user experience of some data handling processes is to change, rationalize and consolidate the software used to execute those processes. Patria group IT hopes that this thesis will shed some light on how to execute the change in a way that the software that is chosen actually improves the user experience, rather than just changes it.

In addition to causing issues with the quality of work, bad user experiences, like running into usability problems, can also lead to users putting off related tasks. This kind of procrastination can have multiplicative delays, if an ongoing process requires these tasks to be done in order to proceed. These sorts of delays may have serious effects on productivity. For a similar example, in their study Davids et al. found that improved usability of their e-learning resource translated into improved motivation and willingness to engage with the learning material. [10] This demonstrates the effect of improving usability on the motivation and willingness to engage with tasks. Though to be sure that this finding also applies to workers at Patria and not just students using e-learning resources, a study on the subject should be conducted.

Don Norman talks about the importance of the effect of user experience on how we behave in his book, The Design Of Everyday Things.

"Experience is critical, for it determines how fondly people remember their interactions. Was the overall experience positive, or was it frustrating and confusing? When our home technology behaves in an uninterpretable fashion we can become confused, frustrated, and even angry—all strong negative emotions. When there is understanding it can lead to a feeling of control, of mastery, and of satisfaction or even pride—all strong positive emotions. Cognition and emotion are tightly intertwined, which means that the designers must design with both in mind." [11]

This insight demonstrates a clear connection between a positive user experience and a positive digital employee experience. When, for example, an employee experiences mastery with a digital task, they also have a positive digital employee experience. I discuss this connection further in chapter 2.3 It can also be inferred, that in these cases the employee would be unlikely to procrastinate doing such tasks, and would pay special attention to doing those tasks properly, ensuring they participate in producing data of high quality.

Patria group IT makes extensive use of Efecte ESM. The product, however, isn't quite the perfect fit for their needs, so they have developed their own in-house tool to be used alongside it, particularly for data analytics. Efecte's data analytics user experience has been found to not be a great fit for Patrias' needs either. Norman's book has a relevant insight on this topic, which demonstrates what kind of effects poor design can have, which is also applicable to a data analysis process:

"Design is concerned with how things work, how they are controlled, and the nature of the interaction between people and technology. When done well, the results are brilliant, pleasurable products. When done badly, the products are unusable, leading to great frustration and irritation. Or they might be usable, but force us to behave the way the product wishes rather than as we wish." [11]

It should be emphasized, that there is a distinct difference between usability and user experience. Usability can be defined in a variety of ways, but according to ISO, usability is the *"extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use."* [12]

As with usability, there are many ways to define user experience. The way one chooses to define it also affects the way in which you view, analyse and evaluate it. The ISO defines user experience as "*a person's perceptions and responses that result from the use and/or anticipated use of a system, product or service.*"[12]

Patria requires that the constructed method is easy to use and can be used in a variety of use cases. Notable examples at the moment are software consolidation and procurement. These use cases will include situations in which user experience surveys aren't adequate as well as use cases where a more holistic user experience evaluation is excessive.

Thus, I had to construct the method to be very customizable. If I required that every part of the method must always be executed for the results to be valid or comparable, or alternatively limited the method to be a guideline for conducting usability evaluations, it would not serve all use cases.

In her dissertation on the evaluation of user experience in challenging circumstances, Keskinen chose a simple and flexible definition of user experience due to the practical nature of her user experience evaluation work. As my thesis involves constructing a user experience evaluation method for practical use, I will use the same definition of user experience as she did:

"*A user's subjective opinion about (or answer to) a certain statement (or question) about the system (or modality, interaction, or any other specified target) in a certain context at that time.*"

Her reasoning for choosing such a definition was to not restrict the kinds of users, opinions, statements, questions, systems, or contexts it can deal with. I also do not want to restrict the user experience evaluation method I construct in such a manner. [13]

2.2 Evaluating User Experience

As defined by Keskinen, user experience is subjective. Because of that, evaluating it is challenging. As is the nature of subjective experiences, they are different for everyone. Thus, comparing and analysing them is difficult. Despite that, a demand exists for evaluating the user experience and to answer that demand, many methods exist. Each are an attempt to evaluate the user experience in the context they are designed for. [14]

Patria, as a defense sector company with many different contexts in which its employ-

ees experience software, has its own needs. Patria requires a single user experience evaluation method that can be used to create evaluations for the purposes of:

1. application landscape rationalization
2. planning the improvement of other internal employee use of software
3. procuring new software

The evaluated software will have varying levels of required security clearance, usage environment, user role and experience level of users. These factors make the design of user experience evaluation challenging. While the needs of Patria are nothing unique, what's particularly unusual about them is the variety of user experience evaluation contexts in which Patria wishes to conduct evaluations. Thus, a very broad and customizable user experience evaluation method is required.

As Keskinen notes, amazingly little detailed information still is available on how to actually evaluate user experience in practice. Thus, researchers are forced to create new methods or at least variations of existing ones. [13] This is true in this case as well. To meet the needs of Patria, I designed a new user experience evaluation method based on existing methods. In the following subsections, I will briefly introduce them.

2.2.1 SUXES

SUXES is an evaluation method for collecting subjective metrics with user experiments. [15] As mentioned earlier, user experience is subjective and evaluating it is challenging. Methods like SUXES are the best attempts at doing so.

SUXES is based on a service quality metric called SERVQUAL. It was developed by marketing academics for evaluating the service quality of real world services. SUXES is adapted from it for evaluation of novel interactive applications and different modalities in multimodal applications. The purpose of SUXES is to capture both user expectations and user experience of different interaction techniques, or modalities, and the whole application.[15]

The method I constructed focuses on evaluating the user experience in the various evaluation contexts found in Patria. SUXES does contain the interesting idea of evaluating the user expectations and comparing them to the actualized user experience. This is something I also recommend for a future iteration of this construction, as this could be particularly useful for Patria for evaluating the user experience of software they are considering procuring. Comparing the expectations Patria employees have of software that is being procured with the actual user experience they have when trying it out or taking it into use would provide a key data point that would speak volumes of the user experience of that software.

2.2.2 Experiential User Experience Evaluation Method

Experiential user experience evaluation method is the main inspiration for the user experience evaluation method that I constructed for Patria. Experiential user experience evaluation method is based on SUXES and the Experience Pyramid model. The Experience Pyramid model is a theoretical model for tourist products. It is designed to to be utilized for designing, analyzing, and developing particularly tourism products emphasizing the experiential aspects. [13]

Keskinen, the author of the dissertation in which she introduces her Experiential user experience evaluation method, explained that for measuring the user experience of an interactive system on a more general level, the SUXES method described above was chosen. To address the experiential aspects, she chose the Experience Pyramid. The "experientality" aspect of the Experience Pyramid served as a foundation for the Experiential user experience evaluation method. The Experience Pyramid is also the inspiration for the statement pairs that each correspond to a element of experience, which is also a key feature of the user experience evaluation method that I constructed. [13]

Based on the elements of experience and statement pairs, I created the user experience evaluation questionnaire to be used as a tool in the user experience evaluations conducted with the user experience evaluation method I constructed. However, I adapted the seven-step rating scale in between the negative-positive statement pairs to a five-step rating scale on the recommendation of my thesis supervisor, Jussi Okkonen. This change was also welcomed by the Patria group IT leaders, who felt that keeping the user experience evaluation questionnaire as simple as possible was wise. The pool of respondents to these questionnaires at Patria would be wide, deep and diverse. If the questionnaire was any more complicated than it had to be, it could lead to different interpretations of what the questions mean as well as what the different steps of the rating scale are.

Keskinen noted, that her core list of elements of experience, based on the Experience Pyramid, *"cover only some aspects of user experience"* and that *"additional inquiries may be needed."* Thus, she introduced the possibility of *optional measures* that can be included or excluded as necessary or desired. These measures can concern roughly any aspect of the system or a specific interaction technique, for instance. [13]

Due to the broad spectrum of evaluation contexts that Patria has for the user experience evaluation method, it had to be constructed to be highly customizable. Thus, I also included these optional measures as part of the user experience evaluation method that I constructed. I deemed it necessary to allow further customization of the method and thus ensure viability in the various evaluation contexts. In addition to adding the possibility of optional measures for elements of experience, I also made every part of the method optional. An exception to this, the only mandatory part, is the user experience evaluation

questionnaire, which forms the basis for the score that is given to evaluated software.

2.3 Digital Employee Experience

When I set out to conduct this part of the literature review, my goals were to understand what is a DEX-model, and how should DEX be measured and then create a DEX road map. After going through a fair bit of literature on the subject, it became clear that there is no DEX-model.

2.3.1 What Is Digital Employee Experience?

In the research their consultants have conducted, Gartner declares the following:

"The adoption of digital technology and migration to distributed enterprise, where employees work from anywhere, has been unprecedented during 2020/2021. For the hybrid or remote digital worker, technology is the primary means by which they interact with colleagues, managers and customers. The spotlight is now on infrastructure and operations leaders to drive digital employee experience (DEX) improvements to help the company attract and retain talent. Rising to the challenge demands a strategy that focuses on the employee's overall experience with technology. Although some call this the "end-user experience" or the "user experience," Gartner calls this DEX, and this strategy is supported by DEX tools." [3]

In other words, Gartner has designed a employee-focused strategy that can be utilized to methodologically improve the digital employee experience. Defining user experience as "overall experience with technology" could be seen as an over-simplified version of the ISO definition. The ISO defines user experience as "*a person's perceptions and responses that result from the use and/or anticipated use of a system, product or service*." [12] Renaming it DEX is somewhat confusing, and made understanding Patria's requirements for the thesis more complicated. While Gartner did not have a DEX-model, they seem to be able to help companies like Patria devise a DEX-focused strategy.

2.3.2 Why Does Digital Employee Experience Matter?

According to research by Gartners' analysts, digital employee experience is now considered a major component of overall employee experience. This necessitates a greater focus on continually measuring and improving employee sentiment, technology adoption and solution performance. [5]

Gartner also found, that *"As work becomes more digital, employees increasingly depend on technology, and organizations become more distributed enterprises, the digital experience has become a larger component of the overall employee experience."* Particularly

relevant for this thesis is the finding, that "*Most organizations are trying to rationalize and consolidate the number of tools used to manage digital workplace technology, making it difficult to justify additional investment."* [3] This is something that Patria is also currently doing, and Patria is making an effort to do it in a way that also improves DEX, hence they commissioned this thesis.

According to other research from Gartner, *"The pandemic-induced disruption of work on every front elevated digital workplace to an executive-level priority."* My thesis is motivated in part by these findings, as Patria leadership also consider DEX an executive-level priority. The key tenet of Gartners' digital workplace program is *"centering the human experience and applying technology to improve business outcomes."* [16] A user experience designer is the ideal background for a researcher aiming to improve Patria's digital employee experience.

2.3.3 How Should Digital Employee Experience Be Measured?

"Greater dependence on technology means the employee's primary connection to the organization is digital workplace technology. If this technology is not reliable, intuitive or engaging to use, employees face frustration-inducing digital friction that inhibits productivity. DEX tools can identify and reduce digital friction by aggregating the quantitative aspects of endpoint and application performance and adoption, and qualitative aspects of employee sentiment and organizational context. This data is processed by advanced analytics and machine learning to generate actionable insights that can be resolved through self-healing automation and contextual nudges to influence employee behavior." [5]

Gartner's DEX blueprint document discusses several metrics that can be used to measure digital employee experience.

- Digital friction - the quantitative aspects of endpoint and application performance and adoption, and qualitative aspects of employee sentiment and organizational context
- Digital workplace maturity - a model defined by Gartner
- IT support demand
- Employee engagement - The extent to which a worker is willing to apply discretionary effort in order to achieve individual, team and organizational goals and feels the organization enables them to do their best work.
- Employee enablement - An IT-driven strategy focused on enabling employees with the tools, support and skills required to increase digital dexterity and self-sufficiency.
- Employee empowerment - The collective policies, practices and technologies that encourage employees to actively participate in co-creating digital capabilities to

improve business outcomes and instilling a sense of purpose, as influencers and champions.

[17]

Gartner's DEX Market guide gives recommendations regarding DEX measurement and scores:

"Automating measurement to understand the impact that digital workplace technologies, services and processes have on the workforce. This includes collecting, quantifying, evaluating and removing digital friction and defining human-centric success metrics." [18] I took these recommendations to heart when constructing the user experience evaluation method, as its purpose is to improve the digital employee experience at Patria. The method was constructed to have human-centric success metrics, the elements of experience, and to be as automated and automateable as possible.

2.3.4 The DEX Road Map For Patria

Gartner's DEX Blueprint highlights, that *"high-quality UX is integral to DEX. Employees with a high-quality UX are more likely to have high levels of work effectiveness, productivity, intent to stay at their current employer and ability to deliver a good customer experience. However, UX is often misunderstood and underappreciated by technology-centric IT leadership. UX is not just about an application's user interface (UI) and performance. Great UX reduces cognitive load, context switching, accessibility barriers and other distractions, inhibitors and causes of digital friction."* [17]

In their article on how to enable a hybrid future of work, Gartner states that *"Seismic shifts in the post-pandemic workplace and workforce require new and evolving expertise in driving adoption of digital workplace applications and supporting emerging business technologists."* [16]

It's clear from these findings, that Patria should focus on improving the user experience of the software it uses in order to improve their digital employee experience. This work has started at Patria with the commissioning of this thesis. Next steps are to utilize the constructed user experience evaluation method in the various projects related to changing the application landscape at Patria. This can mean hiring UX personnel and upskilling resources beyond traditional IT support and engineering roles. This will enable an environment for successful DEX deployment within the organization. Creating a special interest group for DEX is also recommended. [19]

3. CASE STUDIES

As the topic of the thesis is about creating a construction, it must be validated through practical application of the construction. In this case, a case study for the purpose of weak market testing was chosen as the validation method.

During the thesis process, three different case studies were worked on. The initial idea was to simply find a relevant case to work on and go through with it, but it didn't quite work out that way. The first case study was dropped in favor of a case that seemed more relevant and would have had much more impact in the Patria group IT organisation. That second case study was then dropped due to the related software being stuck in pre-pilot phase for too long. The third case was fortunately also relevant and of importance to the organisation, and it was finished in time without further complications.

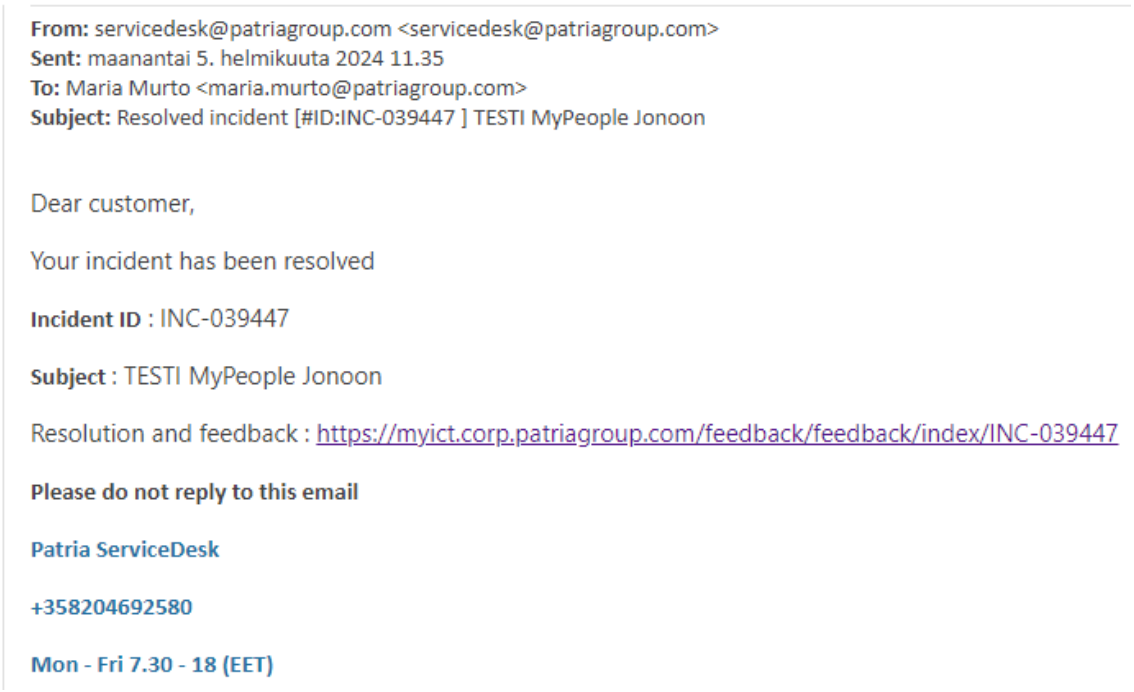
In the following sections I go through the work I did on the two abandoned cases, and explain how they shaped this thesis and their influence on the construction. I also go through the successfully completed case.

3.1 Case Patria Helpdesk Feedback

The first potential case study I was presented with was the improvement of the user experience of the Patria Helpdesk Feedback system. Patria Helpdesk Feedback is a system that customers of Patria Helpdesk can use to give feedback on tickets that they are customers to and have been resolved.

Figure 3.1 shows an example of the first phase of the user experience of using the Patria Helpdesk Feedback system. Once a ticket, in the case of figure 3.1 an incident ticket, has been resolved, the customer of the ticket is sent the email in the figure with the link to the feedback form attached.

The main issue with the Patria Helpdesk Feedback system was, that people weren't clicking on the link and responding to the feedback survey. Patria hoped, that by conducting a user experience evaluation of the Patria Helpdesk Feedback system using the method I construct as part of my thesis, I would be able to give Patria recommendations on how to improve the user experience of the system. That would also hopefully make the process of giving feedback simple enough, that customers would do it more often.



From: servicedesk@patriagroup.com <servicedesk@patriagroup.com>
Sent: maanantai 5. helmikuuta 2024 11.35
To: Maria Murto <maria.murto@patriagroup.com>
Subject: Resolved incident [#ID:INC-039447] TESTI MyPeople Jonoon

Dear customer,

Your incident has been resolved

Incident ID : INC-039447

Subject : TESTI MyPeople Jonoon

Resolution and feedback : https://myict.corp.patriagroup.com/feedback/feedback/index/INC-039447

Please do not reply to this email

Patria ServiceDesk

+358204692580

Mon - Fri 7.30 - 18 (EET)

Figure 3.1. *Notification when incident has been resolved, with the link to the feedback tool included.*

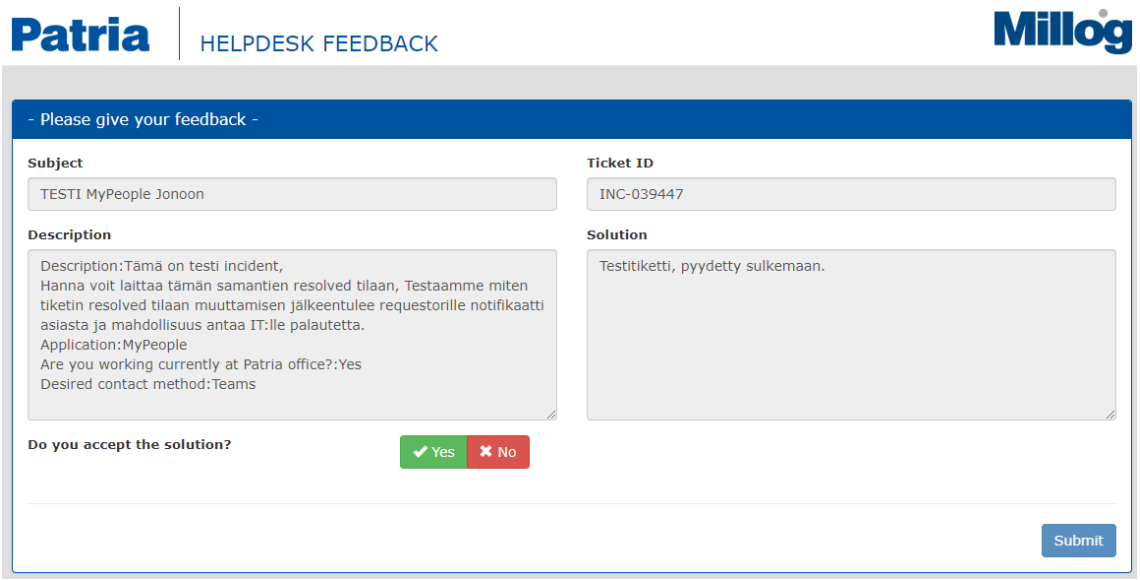


Figure 3.2. *The initial view of the Patria Helpdesk Feedback system.*

For the purposes of the case study, a test incident ticket was created and resolved. I was sent the link as shown in figure 3.1, and it led me to the view shown in figure 3.2. The initial view includes information about the ticket that was resolved and a simple yes/no question: "Do you accept the solution?"

If the customer of the ticket does not accept the solution, they are asked to provide their comment, as shown in figure 3.3. This situation was the most common use case of the Patria Helpdesk Feedback system, as it seemed that usually people only felt the need

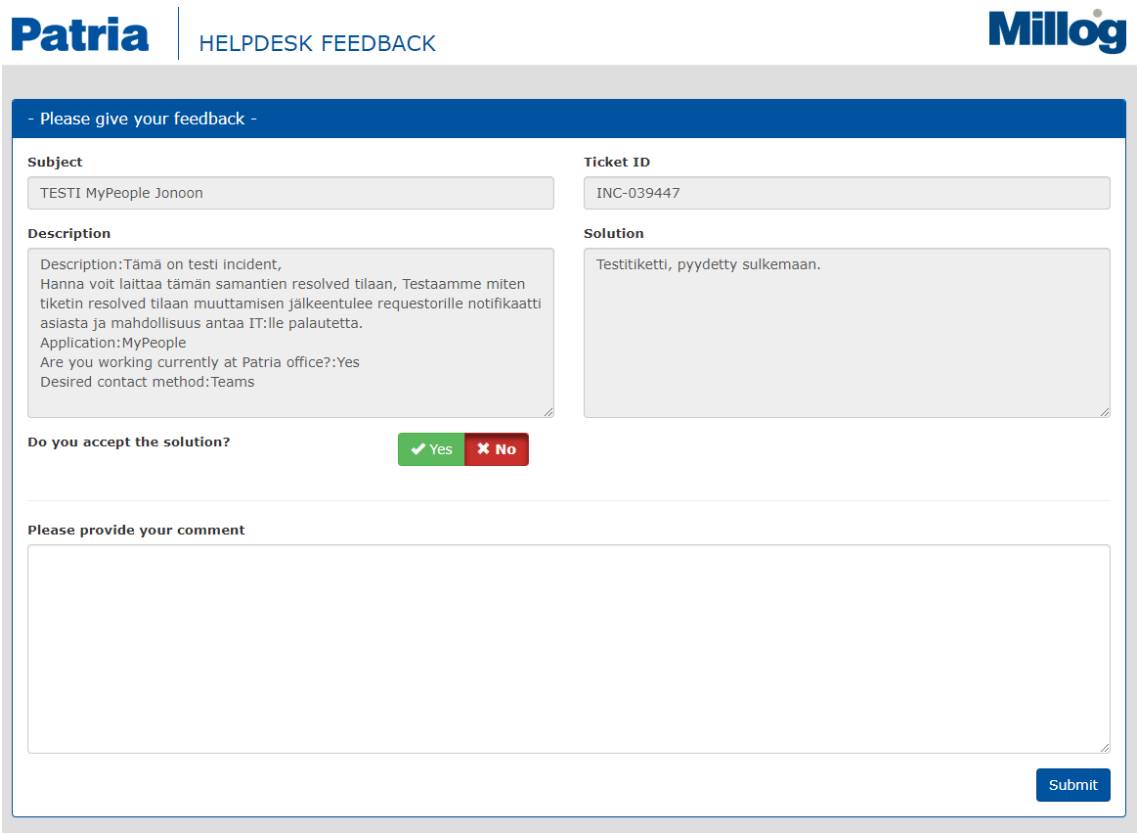


Figure 3.3. *The view the user of Patria Helpdesk Feedback system sees, when they select the option that they do not accept the solution of the ticket.*

to give feedback when they had a problem with the solution of their ticket and wanted to express it.

If the customer did accept the solution, they were asked to rate how well the service met their expectations and how satisfied they were with the given service. The two very similar questions used entirely different rating scales. This option also included a text input option, but it was titled "Free comment", instead of a request for verbal feedback, as shown in figure 3.4.

The Patria Helpdesk Feedback system was in need of a user experience evaluation. The evaluation would be able to uncover the reasons that customers aren't using it. Based on those findings, it would be possible to give recommendations on how to remove those barriers. This would elevate the communication between Patria Helpdesk and its' customers to a new level, allowing for targeted improvement of the service the helpdesk provides. The improvements to the quality of this service would also lead to an improvement in the digital employee experience.

However, this was not as relevant of a case as case Anselmo vs. Patria Group IT Survey, which was proposed to me in the very early stages of this case. Thus, this case was abandoned for now, though I do recommend that Patria group IT returns to this case and conducts the usability evaluation when time and resources allow.

Figure 3.4. *The view the user of Patria Helpdesk Feedback system sees, when they select the option that they accept the solution of the ticket.*

3.2 Case Anselmo vs. Patria Group IT Survey

Case Anselmo vs. Patria Group IT Survey would have been a case study involving introducing two competing employee feedback tools to Patria group IT function. Once we had defined that I was to create a method for evaluating user experiences for software consolidation, this proved to be a particularly relevant and interesting use case of my construction. Despite being left unfinished, it provided a useful starting point for the practical portion of my thesis.

The goal of the case study was to create comparable evaluations of both of the employee feedback tools to see the differences between the two. That knowledge could then have been utilized as part of the decision making process of which employee feedback tool to adopt. This would have been an example of software consolidation done using the construction I created. It should be noted that the Anselmo product had much more resources directed to it than the Patria Group IT Survey. Thus, the situation was biased towards adopting Anselmo from the get go.

The purpose of adopting and improving a employee feedback tool such as these at Patria

group IT function is to gather feedback on the perceived user experience of the software that is in use at Patria. This feedback will be used as part of a later iteration of the user experience evaluation method constructed here, but the data will also have other internal uses in various kinds of software-related decision making processes. Gathering this data and acting upon the insights it provides will lead to improvements in the digital employee experience of Patria employees via improved user experiences of software they use.

In this case, the knowledge gained from the evaluations could also have been used to improve the employee feedback tool that received a better user experience evaluation by integrating into it the things that the worse performing employee feedback tool did better. Additionally, the case study could have produced some suggestions on how to further develop the two employee feedback tools in a way that would have produced more relevant data in greater variety for the purposes of user experience work.

Discussions were held on the topic of the current capabilities of the two applications and if they could be adjusted to better serve the purpose of aiding in the process of improving digital employee experience at Patria group IT. Based on those discussions, we decided that the most appropriate way to achieve this would be to design new questions to be asked by the tools.

3.2.1 Anselmo

Anselmo, or Patria-questionnaire as it's called currently, is a tool that sends interactive desktop notifications. It was initially developed by A-insinöörit, and it's currently being modified for use at Patria by Elmo, which is a close partner of Patria group IT.

The Anselmo software tool in its current state is manually opened and closed using the icon shown in figure 3.5. The tool can send out two kinds of notifications. Figure 3.6 displays the user interface into which both kinds of notifications appear and where they are interacted with. Figure 3.7 is the only other view of the application and contains the privacy statement, although only in Finnish.

The first type, which is also the only type it will have on-release, gathers Net Promoter Score (NPS) type data in response to the notifications. In practice, the Anselmo software sends users of Patria devices a desktop notification which asks you to give a rating from 1-10, which relates to a specific service request the recipient of the notification will have asked service desk to complete.

In a later phase of the Anselmo project, the software will start also sending out a second type of notification, which gathers data more akin to Customer satisfaction score (CSAT) type data in response to the notifications.

Before taking on the case, the expectation was, that the tool would enter pilot-testing

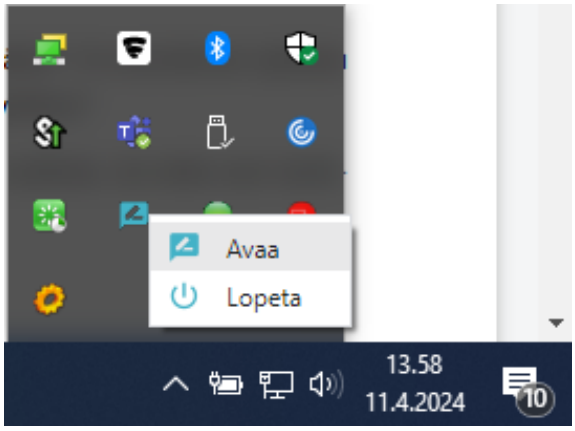


Figure 3.5. *The blue icon is the Anselmo questionnaire icon. Displayed in the figure is how to open or quit the application manually.*

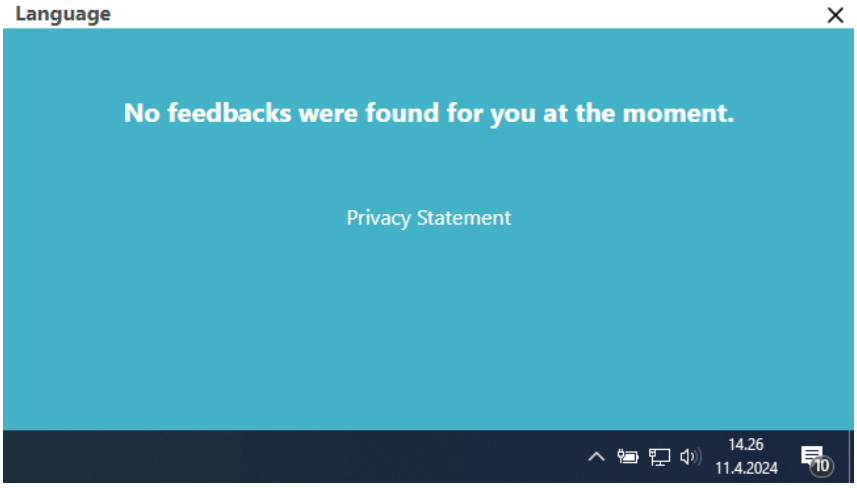


Figure 3.6. *Anselmo questionnaire open with no pending questions.*

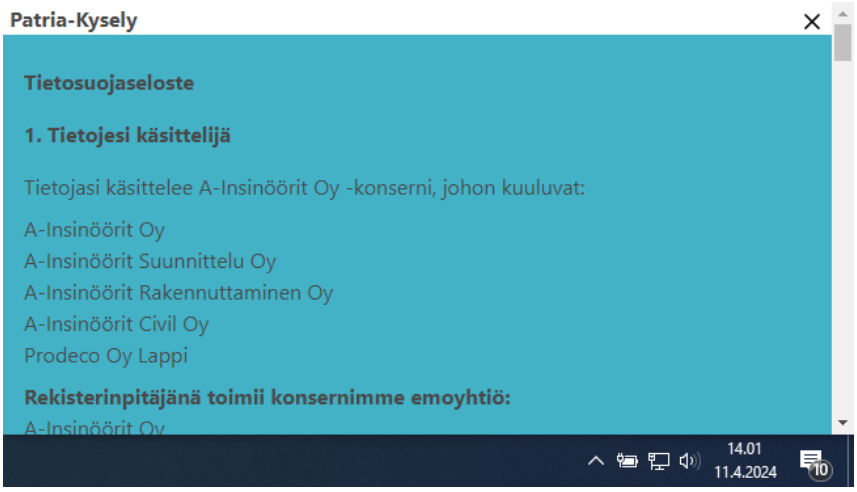


Figure 3.7. *Anselmo questionnaire privacy statement. Not available in English.*

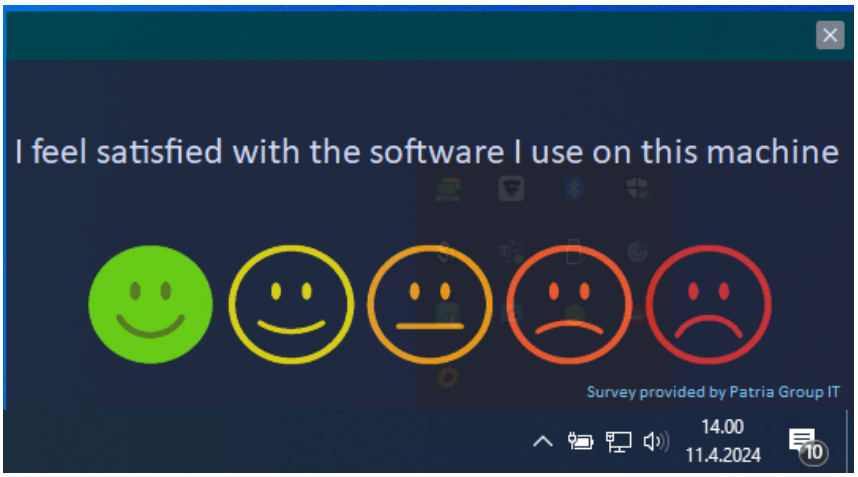


Figure 3.8. *Patria group IT survey open with the most positive smiley selected.*

phase by early March. It however ended up being delayed well into April. It's not expected to reach the phase of sending out the second type of notification before May, and that is too late for this thesis.

3.2.2 Patria Group IT Survey

During the planning phase of the case study, I was informed that Patria group IT function also had an existing, very similar tool which was at the time unnamed. That enabled expanding this case study from just a user experience evaluation of a employee feedback tool for the purposes of improving digital employee experience, to also include software consolidation. That's what led to the decision to abandon case Patria Helpdesk Feedback, as this case seemed to fit the purposes of this thesis better.

While Anselmo was developed by a team of software developers, this tool had been developed by one Patria group IT employee. For the purposes of this case and to potentially aid in the Anselmo pilot, the employee worked more on his tool, and named it Patria group IT survey.

Before the case began, the tool was just a way to send users an interactive notification with five smiley faces which represented scores in a range of 1 to 5. An example of this is displayed in figure 3.8 Data gathered via this tool is sent to a database in internal use, from where it can be retrieved for analysis. That integration with data-analysis systems that were already in use, Efecte and eLat, was also one of the main selling points of this tool. However, Anselmo also eventually got to the point where it was well integrated with Efecte. The data-analysis done based on the data gathered with these tools could help identify which software had poor user experience, and perhaps even what could be done about it, depending on what's asked from the users.

For the purposes of this case study, the software was modified to launch on startup and

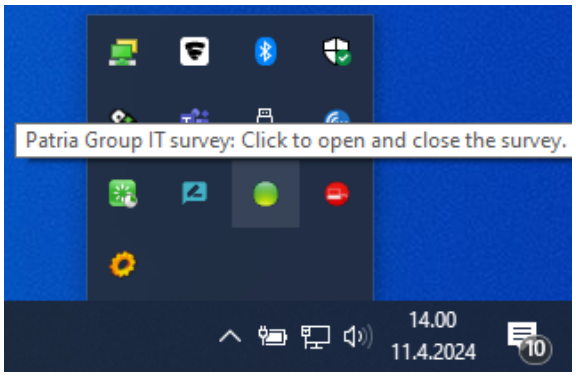


Figure 3.9. *Patria group IT survey icon, which can be used to manually open and close the survey, as explained by the tooltip.*

open to ask the specified questions from the specified users at specified intervals. The possibility to ask specific questions was added as well. Figure 3.9 displays the icon of the tool, which can be used to manually respond to whichever question is currently active.

3.2.3 Developing Questions For The Two Tools

While the ability to ask specific questions was added to the Patria group IT survey, I needed to actually research which questions would be most suitable for gathering data using these tools for the purposes of the user experience evaluation. They would be useful for Anselmo later as well.

First the decision had to be made about the direction and the purpose of the questions. These were discussed in one of our Monday meetings with the Patria group IT leaders. We debated if the questions should aim to:

- Gather an understanding of the tasks done on various software for the purposes of future user experience evaluations.
- Recognize software that had problems with user experience.
- Understand the current state of the usability of the various software in current use.
- Gather users feedback and wishes regarding various software in current use.
- Gather data for creating the user experience scores of software.

We ended up selecting the final option, though perhaps the other options could also be included as parts of the tools in future iterations. The final option was the most straightforward and beneficial direction for the purposes of this thesis.

My starting point for designing a set of useful questions began by referencing existing sets of questions. One set I refer to is designed for measuring the perceived usability of artefacts, the system usability scale (SUS). [20] The other is for measuring the post-task experience, NASA-Task Load Index (NASA-TLX). [21] Post-test questionnaires like the

SUS measure perceived usability of an entire system; post-task scales [like the NASA-TLX] identify problematic parts of a design. [22]

The set of questions SUS provides, listed below, gave good direction for the format of the questions, but none of the questions made it onto the proposed list as-is. Due to the state of the tools at the time, these questions would not have been appropriate. The tools weren't able to ask users these questions during or shortly after the use of a system. They required further development to have that capability.

1. I think I would like to use this system frequently.
2. I found the system unnecessarily complex.
3. I thought the system was easy to use.
4. I think that I would need the support of a technical person to use this system.
5. I found the various functions in this system well integrated.
6. I thought there was too much inconsistency in this system.
7. I would imagine that most people would learn to use this system very quickly.
8. I found the system very cumbersome to use.
9. I felt very confident in using the system.
10. I needed to learn a lot of things before I could get going with the system.

As the direction of the construction of the user experience evaluation method was chosen in collaboration with the leadership of the Patria group IT function, my approach to curating the selection of questions was to prepare a list of potential question candidates for one of the Monday meetings I had with Patria group IT leadership.

I had the goal of preparing questions for each use case of both applications. This means questions for use of Patria Group IT Survey and both types of questions for Anselmo: NPS and CSAT.

I created a shortlist of questions that were task-specific and system-specific based on the NASA-TLX, SUS and NPS. I also included a generic question of "Why did you give a score of (score)?", in case we could, or wanted to, only ask free-form questions attached to one type of rating question.

Questions to ask the user of a software:

Task-specific:

1. How successful were you in accomplishing what you were trying to do?
2. How hard did you have to work to accomplish your level of performance?
3. How mentally demanding was the task?
4. How frustrated were you?

5. Overall, how did you feel about the task?

System-specific:

1. How likely would you recommend this system to a colleague?
2. I would imagine that most people would learn to use this system very quickly.
3. I found the system unnecessarily complex.
4. I felt very confident in using the system.
5. I think I would like to use this system frequently.

The initial list of questions I proposed to the IT leaders, listed below, was a curated from the shortlist. It was still somewhat unclear, if questions could be ask during specific tasks or if they would need to be more system level. Thus, the list included a variety of example questions suitable for different situations. These were also meant to gather data related to some of the elements of experience, that were being designed at this time. Notably cognitive load and utility are highlighted in these questions.

1. Why did you give a score of [X]? (General open feedback)
2. How successful were you in accomplishing what you were trying to do? (NASA-TLX)
3. How hard did you have to work to accomplish your level of performance? (NASA-TLX)
4. How mentally demanding was the task? (NASA-TLX)
5. How likely would you recommend this software to a colleague? (modified NPS)
6. I feel satisfied with the software I use on this machine. (modified SUS)

The initial list of questions I proposed that we could have the feedback tools gather were not well received. The Net promoter score (NPS) -like question drew particular ire, as the experience among Patria group IT leadership was, that the negatives outweigh the positives of the metric. The benefits of the metric are that it's simple, comparable, works well with large sample sizes and puts customer satisfaction at the forefront. On the other hand, some of the benefits can also be seen as negatives. For example, simplicity can also be a negative attribute. It doesn't identify problems or offer solutions, it's just one simple number. [23]

The stakeholders also raised the issue of complexity of the questions. They felt that employees would be reluctant to respond to the questions at all, but they would be particularly disinclined to do so, if the questions asked were complicated. Complex questions were also expected to increase the chance, that different employees would interpret the questions differently, which would become a major confounding factor in analysing the results.

After receiving feedback on the first iteration of the list of questions, I set about creating a new set of questions that took the stakeholders concerns into account. I worked on this task for a while, but around this time, it became apparent that neither software related to this case would be finished during the time I had to complete my thesis. Thus, this case would also be abandoned, and we set out to find a new case study, one that I could surely finish and would be just as relevant and important as this one would have been.

Despite this, Patria will continue to work on the Anselmo tool, and perhaps Patria Group IT Survey as well, as it was recognized that these tools could be of great help in the process of improving the digital employee experience of Patria employees.

Despite this case study being cut short, I did create the next iteration of the set of questions to ask, listed below, which did end up getting validated by the Patria group IT leaders during one of our Monday meetings. The validation was easy to get, as it took into account their concerns. It also included more of the elements of experience, as they had also been further refined by this time.

I recommend that these questions are integrated into the tools and used to gather data about these elements of experience. This data could serve as a basis for building the user experience scores, and that would make these tools powerful additions to the improvement of digital employee experience of Patria employees. The creation of these questions and this potential synergy with the construction means, that this case was certainly not a waste of time.

1. I feel very confident in using the system. (SUS, learnability)
2. I feel there was a lot to learn about the system. (modified SUS, learnability)
3. I feel the use of the system is often exhausting. (Modified TLX, cognitive load)
4. I feel the system is well-suited for the tasks I need to use it for. (Utility)
5. I feel the system is confusing to use. (Learnability, cognitive load, error prevention and recovery)

3.3 Case Efecte

Efecte ESM is an Enterprise Service Management tool by the software company Efecte. In their marketing materials, they promise to deliver "A system for incident processing by design, a No-Code Workflow Engine, and a process visualization application [which] enable the user to implement any process flow." [24] Evaluating the user experience of Efecte and considering the options for which software to change to is among the many upcoming projects related to changing the software they have in use at Patria.

The case started as soon as it was decided that Case Efecte is the one I will use to test my User Experience Evaluation Method. In addition to testing it, it also served as

a weak market validation of the construction, the user experience evaluation method. I immediately began informally interviewing an expert user of Efecte.

In the interview, I discovered that the most important as well as most commonly done task in Efecte is ticket management. Analytics is done relatively infrequently and is of comparatively low importance. The interviewee felt that analytics in Efecte was lacking in some aspects, and better done using other tools. Other functions of the software are of tertiary importance.

It had also been noticed by many, according to the interviewee and other discussions I had during the case, that people would often find shortcuts and skip parts of the ticket handling process. It was thought that this was due to some usability problem related to the ticket handling process in Efecte. These skips and shortcuts led to a deterioration in the quality of data Efecte produces about the tickets.

This deterioration was highlighted by the fact that Patria had recently created a Master Data Management team, which was busy spreading knowledge about the importance of data quality, and how everyone could play their role in improving data quality. Thus, the link between usability problems and data quality was another important reason that this particular case ended up being the one I took on.

Efecte has been in use at Patria for, according to informal interviews of Patria employees with long tenures at the company, as long as two decades. Efecte is in heavy use at the company, and quite divisive. It's not uncommon to hear complaints about the software, and any user can immediately point out a usability problem or two. Evaluating the user experience of Efecte and considering the options for which software to change to is among the many upcoming projects related to changing the software they have in use at Patria that I mentioned.

As at the time of the case Digital Experience Monitoring tools, internal questionnaires or other similar tools were not in use, the user experience evaluation was designed based on manually gathered information. The first data gathering method was the Efecte User Experience questionnaire.

3.3.1 Efecte User Experience Questionnaire

The questionnaire was designed to gather information about how Efecte ESM scores on the various elements of experience. The questionnaire was sent to all 264 Efecte ESM license holders. It received 74 responses, which equals a 28% response rate. The questionnaire was made to be as fast and easy as possible to respond to, with only optional text input. It took on average 03:46 to fill out the questionnaire. Total time spent on responding to the questionnaire was 4 hours and 42 minutes.

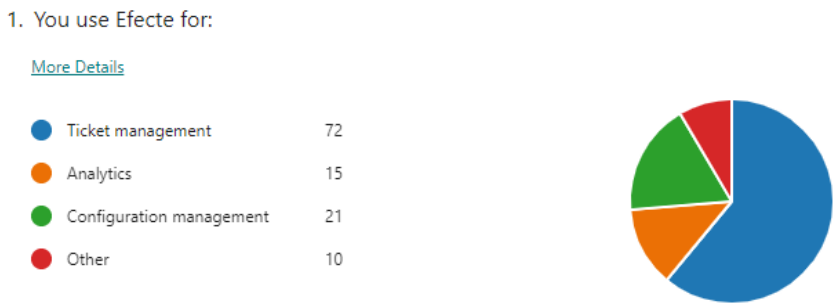


Figure 3.10. *The responses to what the respondents use Efecte for.*

The questionnaire started with an introduction to the questionnaire and its purpose:

"This is a questionnaire related to a master's thesis that Patria has ordered on the topic of user experience evaluation. All Efecte license holders are invited to participate in this pilot of the user experience questionnaire.

As part of this questionnaire, I will collect identifiable information. It will only be used to analyse how your Efecte use has affected your reported user experience with Efecte."

As mentioned earlier, the focus of this user experience evaluation was on the ticket handling aspect of the Efecte ESM. Because of this, we decided to ask the respondents what they use Efecte for, see figure 3.10 Most only used it for ticket management and only three respondents didn't use it for ticket management at all.

We were also interested in how active and experienced users of Efecte, or applications in the same domain, the respondents were, see figure 3.11 This was hypothesized to have an effect on the respondents experience with the various elements of experience of Efecte. In fact, Segniel and Dittberner have found that the activity of users relates more to the users' performance of technology use than duration of experience, but we were curious about both. [25] Most of the respondents were very active and experienced users, with a large majority using it daily and nearly half of the respondents having more than three years of experience.

I designed a set of eight questions for the questionnaire, each corresponding to a element of experience. The questions and the format of negative-positive statement pairs were based on the Experiential User Experience Evaluation Method and the SUXES - user experience evaluation method. See table 3.1. for the questions and which element of experience they associate to.

Notably, two of the elements of experience were left out of the questionnaire, accessibility and user role. Accessibility can be seen to be composed in part by how well the software scores on the other elements of experience. This score can also of course be supported by a accessibility evaluation, but that is not strictly required and was not conducted in this case. As for the user role, we felt it was covered well enough by the earlier questions,

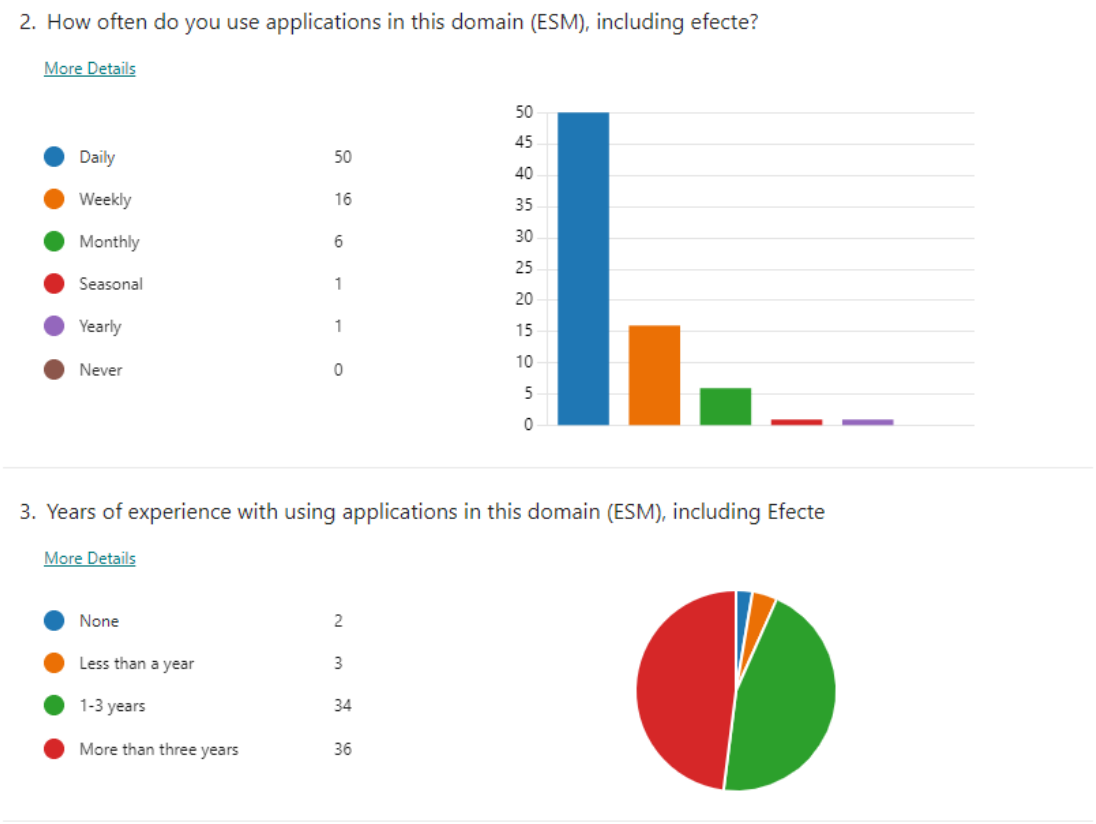


Figure 3.11. *How active and experienced users of Efecte, or applications in the same domain, the respondents were.*

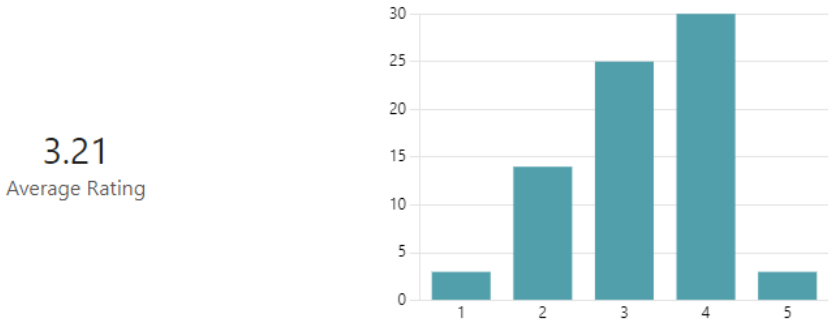


Figure 3.12. *Completing my tasks with Efecte is inefficient (1) - Completing my tasks with Efecte is efficient (5) (screen capture of Microsoft Forms questionnaire results)*

shown in figures 3.10 and 3.11 It also seemed quite difficult to ask the respondents about their user role in a way that would capture the spectrum of roles that the respondents had and in a way that would satisfy all the stakeholders and result in usable data.

3.3.2 Analysis Of The User Experience Questionnaire Results

The questionnaire gathered interval data about eight different dependent variables, which are also elements of experience:

Element of experience	Negative statement	Positive statement
Efficiency	Completing my tasks with Efecte is inefficient.	Completing my tasks with Efecte is efficient.
Utility	The functions of Efecte are poorly suited for completing my tasks.	The functions of Efecte are well-suited for completing my tasks.
Learnability	Learning to use Efecte is difficult.	Learning to use Efecte is easy.
Cognitive load	Using Efecte is complicated.	Using Efecte is simple.
Error-free use	It's easy to make mistakes when using Efecte.	Efecte is easy to use without making mistakes.
Robustness	Efecte has errors often.	Efecte does not have errors.
Aesthetics	Efecte is aesthetically unpleasant.	Efecte is aesthetically pleasant.
Usage environment	Efecte is difficult to use without a distraction-free environment.	Efecte use is not affected by distractions.

Table 3.1. *The statement pairs (negative-positive) used in the user experience evaluation of Efecte ESM as part of the user experience questionnaire.*

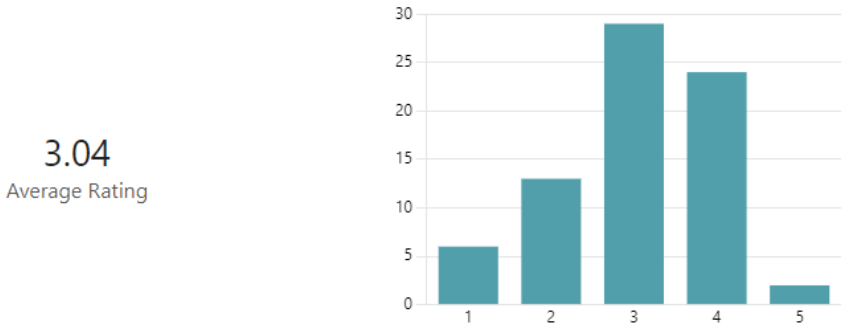


Figure 3.13. *The functions of Efecte are poorly suited for completing my tasks (1) - The functions of Efecte are well-suited for completing my tasks (5) (screen capture of Microsoft Forms questionnaire results)*

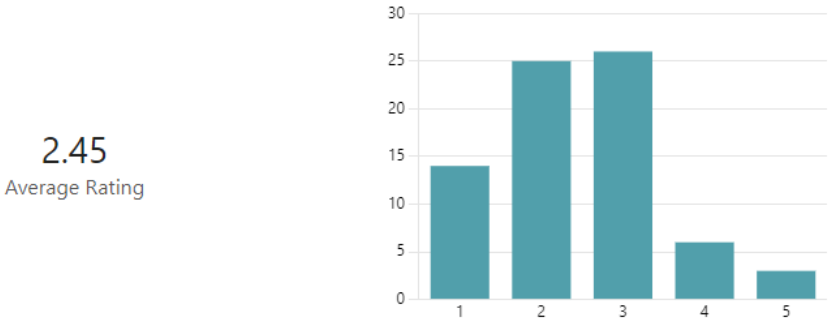


Figure 3.14. *Efecte is aesthetically unpleasant (1) - Efecte is aesthetically pleasing to me (5) (screen capture of Microsoft Forms questionnaire results)*

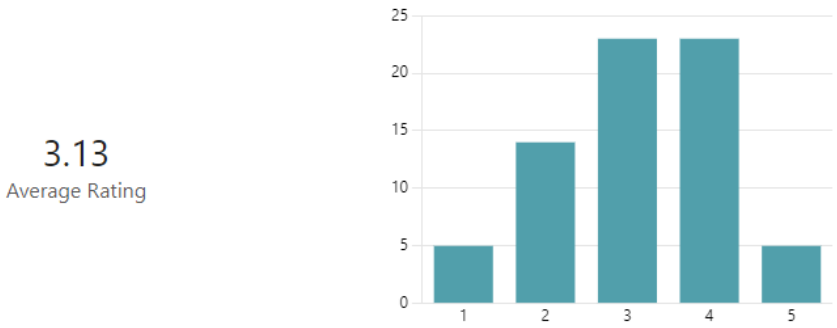


Figure 3.15. *Efecte is difficult to use without a distraction-free environment (1) - Efecte use is not affected by distractions (5) (screen capture of Microsoft Forms questionnaire results)*

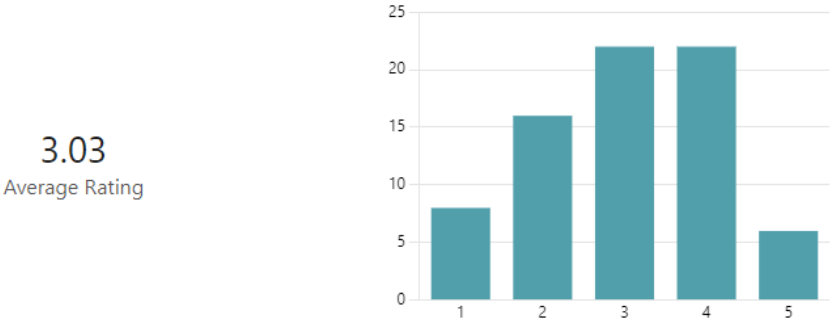


Figure 3.16. *It's easy to make mistakes when using Efecte (1) - Efecte is easy to use without making mistakes (5) (screen capture of Microsoft Forms questionnaire results)*

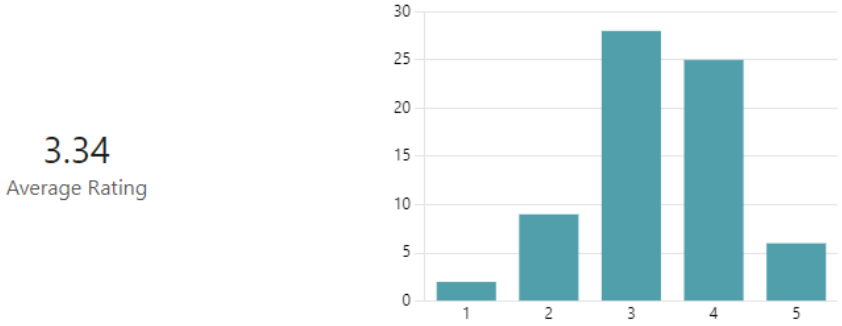


Figure 3.17. *Efecte has errors often (1) - Efecte does not have errors (5) (screen capture of Microsoft Forms questionnaire results)*

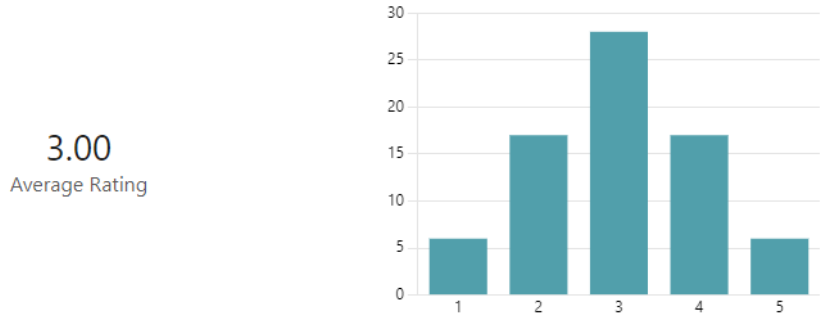


Figure 3.18. *Learning to use Efecte is difficult (1) - Learning to use Efecte is easy (5) (screen capture of Microsoft Forms questionnaire results)*

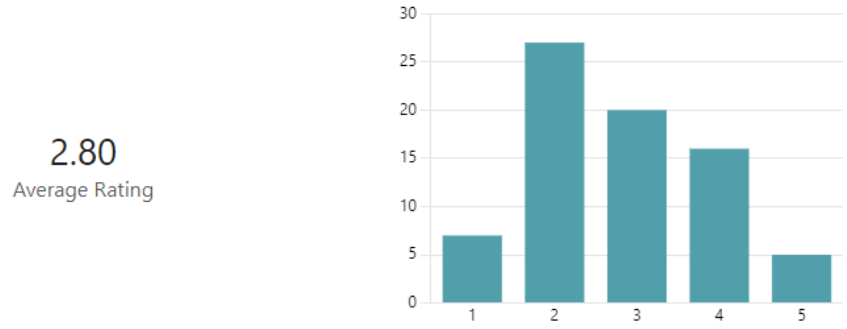


Figure 3.19. *Using Efecte is complicated (1) - Using Efecte is simple (5) (screen capture of Microsoft Forms questionnaire results)*

- Efficiency
- Utility
- Learnability
- Cognitive load
- Error-free use
- Robustness
- Aesthetics
- Usage environment

The internal and external validity of the questions asked in the questionnaire vary. For example, *Efecte has errors often (1) - Efecte does not have errors (5)* is a question with a rather narrow breadth, and thus it has low external validity. Additionally, the accuracy of the answer is high, and thus the internal validity is high. The opposite is true for the question: *Using Efecte is complicated (1) - Using Efecte is simple (5)*

Interval data have equal distances between adjacent values. However, the zero point is arbitrary. According to MacKenzie, it's also commonly used in questionnaires where a

response on a linear scale is solicited, which is also the case here. In these cases, verbal responses are symmetric about a neutral, central value with the gradations between responses more-or-less equal. According to MacKenzie, "it is this last quality – equal gradations between responses – that validates calculating the mean of the responses across multiple respondents." [26]

MacKenzie, as well as my seminar supervisor, note, that here is some disagreement among researchers on the assumption of equal gradations. Do respondents perceive the difference between 1 and 2 (strongly disagree and mildly disagree) the same as the difference between 2 and 3 (mildly disagree and neutral)? [26]

There is evidence that respondents perceive items at the extremes of the scale as farther apart than items in the centre. In their research, Kaptein and others show the errors in the parametric treatment of non-parametric data in experiments of the size typically reported in HCI research. [27]

I analysed, if the usage frequency of Efecte affects the user experience the respondent has with Efecte. Because a majority of the respondents use Efecte daily, fifty out of 74, while the rest use it less frequently, the independent variable is whether the respondent uses applications in this domain (ESM), including Efecte

- Daily?
- More infrequently than daily?

The null hypothesis (H0) for this analysis was:

- The usage frequency of Efecte does not affect the user experience the respondent has with Efecte.

On the other hand, the alternative hypothesis (H1) is:

- The usage frequency of Efecte affects the user experience the respondent has with Efecte.

This independent variable was chosen due to the research by Segniel and Dittberner that shows, that activity of users relates more to the users' performance of technology use than duration of experience. [27] Through this analysis, we will be able to see if the activity of the respondents affects how they perceive the user experience of Efecte.

I've opted to use a non-parametric test, Wilcoxon Rank Sum Test [28], to analyse the results of the user experience questionnaire. Alternatively, Fisher's exact numbers could be used in this case, but for simplicity's sake, I opted for Wilcoxon. This decision was made based on the findings by Kaptein and MacKenzie introduced earlier in this section. [27] [26]

To see if the results of the Wilcoxon Rank Sum Test, also known as the Mann-Whitney U

Test, are statistically significant, I set the p-value: p = .05. In other words, the probability that usage frequency affects the user experience the respondent has with Efecte will need to be 95% or higher to be considered statistically significant. The participants each only selected one level of the independent variable. Either they used Efecte daily, or less frequently. Thus, this dataset was within subjects data.

Based on visual analysis of the frequency distribution of the results of the questionnaire, see figures 3.12 through 3.19, the data is assumed to not have a normal distribution. We can see that 3.18 follows a normal distribution, but it is the only one to do so.

Mann-Whitney U Test assumes, that the variable being compared between the two groups is continuous. In her literature review, Grace-Martin found that five-step scales, like the one used in the user experience evaluation questionnaire, measure *the amount of agreement* the respondent feels towards the negative-positive statements. Based on this insight we can consider the data used in this case to be continuous. [29]

The respondents were categorized into two groups:

- Respondents who use Efecte daily - group 1 - VAR00001 value 1.00.
- Respondents who use Efecte less frequently than daily - group 12 - VAR00001 value 2.00.

Using IBM's SPSS, I ran the Wilcoxon Rank Sum Test, which SPSS calls the Mann-Whitney U Test. The resulting ranks of the test can be seen in figure 3.20 From the Mean rank column in figure 3.20 we can see that group 1 rated Efecte's user experience better in all but questions nine and eleven, which measured Robustness 3.17 and Cognitive load. 3.19

The only statistically significant differences between the two groups, where $p<.05$, was question 4 3.12, see figure 3.21. From the figure we can see that the Asymp. Sig. (2-tailed) column's value, the p value, is only less than .05 for question 4, .044 to be specific. Question 4 is in the first column from the left in the figure 3.21. Question 4 measured the element of experience of efficiency. The respondents who use Efecte less frequently than daily found Efecte to be significantly more inefficient to use than the group who use Efecte daily.

From these results we can see, that while the group that used Efecte more often did find the user experience to be better, the difference was only statistically significant regarding one elements of experience, efficiency. These results do imply that the H1 hypothesis is true, the usage frequency of Efecte affects the user experience the respondent has with Efecte. However, this would require a greater sample size to be sure, as currently the results aren't statistically significant across all elements of experience measured.

Ranks

	VAR00001	N	Mean Rank	Sum of Ranks
Completing my tasks with Efecte is inefficient (1) - Completing my tasks with Efecte is efficient (5)	1.00	50	40.78	2039.00
	2.00	24	30.67	736.00
	Total	74		
The functions of Efecte are poorly suited for completing my tasks (1) - The functions of Efecte are well-suited for completing my tasks (5)	1.00	49	38.22	1873.00
	2.00	24	34.50	828.00
	Total	73		
Efecte is aesthetically unpleasant (1) - Efecte is aesthetically pleasing to me (5)	1.00	49	37.08	1817.00
	2.00	24	36.83	884.00
	Total	73		
Efecte is difficult to use without a distraction-free environment (1) - Efecte use is not affected by distractions (5)	1.00	46	36.70	1688.00
	2.00	23	31.61	727.00
	Total	69		
It's easy to make mistakes when using Efecte (1) - Efecte is easy to use without making mistakes (5)	1.00	49	37.36	1830.50
	2.00	24	36.27	870.50
	Total	73		
Efecte has errors often (1) - Efecte does not have errors (5)	1.00	48	34.58	1660.00
	2.00	21	35.95	755.00
	Total	69		
Learning to use Efecte is difficult (1) - Learning to use Efecte is easy (5)	1.00	50	37.44	1872.00
	2.00	23	36.04	829.00
	Total	73		
Using Efecte is complicated (1) - Using Efecte is simple (5)	1.00	50	37.43	1871.50
	2.00	24	37.65	903.50
	Total	74		

Figure 3.20. *The ranks of the Wilcoxon Rank Sum Test. (screen capture from IBM SPSS output)*

Test Statistics[a]

	Completing my tasks with Efecte is inefficient (1) - Completing my tasks with Efecte is efficient (5)	The functions of Efecte are poorly suited for completing my tasks (1) - The functions of Efecte are well-suited for completing my tasks (5)	Efecte is aesthetically unpleasant (1) - Efecte is aesthetically pleasing to me (5)	Efecte is difficult to use without a distraction-free environment (1) - Efecte use is not affected by distractions (5)	It's easy to make mistakes when using Efecte (1) - Efecte is easy to use without making mistakes (5)	Efecte has errors often (1) - Efecte does not have errors (5)	Learning to use Efecte is difficult (1) - Learning to use Efecte is easy (5)	Using Efecte is complicated (1) - Using Efecte is simple (5)
Mann-Whitney U	436.000	528.000	584.000	451.000	570.500	484.000	553.000	596.500
Wilcoxon W	736.000	828.000	884.000	727.000	870.500	1660.000	829.000	1871.500
Z	-2.009	-.744	-.049	-1.037	-.213	-.278	-.273	-.042
Asymp. Sig. (2-tailed)	.044	.457	.961	.300	.832	.781	.785	.966

a. Grouping Variable: VAR00001

Figure 3.21. *The test statistics of the Wilcoxon Rank Sum Test. (screen capture from IBM SPSS output)*

Element of Experience	Rating	Score
Efficiency	3.21	55.25
Utility	3.04	51
Learnability	3	50
Cognitive load	2.80	45
Error-free use	3.03	50.75
Robustness	3.34	58.5
Aesthetics	2.45	36.25
Usage environment	3.13	53.25
Averages:	3	50

Table 3.2. *The ratings and scores Efecte received from the users in the user experience questionnaire.*

3.3.3 Calculating The Score In The Case Of Efecte ESM

As described in section 4.4., the user experience score is calculated based on the scoring of each element of experience rated in the user experience questionnaire. This score can then analysed. Supporting data can be gathered to help understand the score. In the case of Efecte ESM, supporting data was gathered via three usability evaluations.

The elements of experience evaluated in the user experience questionnaire in this case and the corresponding positive-negative statement pairs can be seen in table 3.1 Table 3.2, displays the ratings and scores Efecte received from the users in the user experience questionnaire. The final user experience score is the average of the scores.

3.3.4 Usability Testing

Usability testing can be included as part of the user experience evaluation method. The data it provides, the subjective feedback and comments, observation data, and so on, can

be used to understand the experiences per se and possible reasons for them. [13]

As part of this case, I conducted three usability tests for the purpose of validating usability testing as a part of this method and to demonstrate, that it can provide supporting data of value. In this case, usability testing could, for example, be done to better understand why the users of Efecte experienced the elements of experience the way they reported to have experienced them on the user experience questionnaire.

Nearly all of the respondents to the user experience questionnaire were experienced and active users of ESM software. To better understand the effects of experience on these elements of experience, I invited inexperienced users to the usability tests.

During the usability tests the Hawthorne effect was experienced by the participants. One of the participants mentioned, that they could tell their behavior during the tests were influenced by my observation. They mentioned not reading the text on the user interface as carefully as they normally would, for example. This should be taken into account when running these usability tests in the future. I should note, that the existence of any Hawthorne effect is under scientific debate, but the participants did feel that my observation of the usability test had an effect on their behavior. [30]

The results of the three usability tests showed, that completely new users may find Efecte confusing, particularly when working in a hurry. This finding is in line with the result that users of Efecte find it to cause somewhat high cognitive load.

3.3.5 Discussing The Case

From the user experience questionnaire, we can find two other independent variables, the effects of which on the dependent variables that we could analyse. One of the independent variables would be the various use cases the respondents have for Efecte:

- Using Efecte exclusively for ticket management
- Using Efecte for ticket management and other tasks
- Using Efecte only for tasks other than ticket management

Analysing this independent variable would reveal if the way the respondents use Efecte affect their perceived user experience of Efecte. I hypothesize this to be the case, as different use cases result in interacting with different features of Efecte, which would likely have variety in their user experience.

The final independent variable would be years of experience using applications in this domain (ESM), including Efecte. Analysing this variable would reveal if years of experience with applications in this domain would affect the perceived user experience of Efecte, but as mentioned, it has been found that the activity of users relates more to the performance

of technology use. Thus, this would likely be less relevant of an analysis than the activity of use. Making these analyses is important for understanding why the score the evaluated software, in this case Efecte ESM, receives from its users.

In conclusion, I found that the user experience of Efecte was experienced to be a 50 on a scale of 0-100. As this is the first user experience evaluation done using this method, it is hard to say how this compares to the average ESM software, but a 50 out of a 100 seems to be a result that is neither bad or great. The study found, that users who use Efecte daily found it to be significantly more efficient to use than the group who use Efecte less frequently. In other words, people get better with using Efecte with frequent use. To see if this is caused by a steep learning curve or if users have difficulty remembering how to use Efecte, further analysis of the results would be required. One final note is that Efecte was found to be particularly displeasing to the eye. As it happens, Efecte will have a user interface revamp in the near future, which will hopefully help.

4. USER EXPERIENCE EVALUATION METHOD

In this chapter I describe the method developed for Patria to evaluate the user experience of software for various purposes, like application landscape rationalization and new application procurement with the overarching goal of improving digital employee experience at Patria.

Based on the SUXES and the Experiential User Experience Evaluation Methods, I constructed a customizable user experience evaluation method to fit the specific needs of Patria.

4.1 Requirements For The User Experience Evaluation Method

Based on our weekly discussions with the leadership of Patria group IT, I have gathered a list of requirements for the user experience evaluation method:

1. The method must be usable for evaluating the user experience of any software that Patria has or may procure.
2. Can be used to score and create otherwise comparable results for evaluated software based on the chosen elements of experience.
3. Can be used to evaluate the user experience of software.
4. Can be used to create recommendations for improving the user experience of software.
5. The method must be as simple to use as possible.
6. The method must not be labor intensive to use.

4.2 The Elements Of Experience

The method measures how the user experience of the evaluated software scores on the elements of experience that are relevant to it. This list of ten elements of experience serves as the core elements that can be used in most evaluation contexts. These elements can be adjusted by removing and adding any of them to fit any evaluation context.

In her Experiential User Experience Evaluation method, Keskinen introduced *optional*

measures, as her measures only covered some aspects of user experience. The same is true for the user experience evaluation method I have constructed. Including and excluding elements of experience is fine, and the optional measures can concern any aspect of the system or a specific interaction technique, for instance. Keskinen provides an example of a self-created optional measure: *Excitement: The application is boring—The application is exciting.* [13]

The Elements of Experience:

1. Efficiency (how long it takes the user to complete tasks using the system)
2. Utility (how well the functions of the system fit the purpose of the user)
3. Learnability (how long it takes the user to learn to use the system)
4. Cognitive load (how demanding of the working memory it is to use the system)
5. Error prevention (how easy it is for the user to make slips or mistakes) [31]
6. Recovery from errors (how well the users notice and understand errors and recover from them) [31]
7. Accessibility (how well the system fulfills accessibility requirements) [32]
8. Aesthetics (how visually attractive the users find the system)
9. Usage environment (how the usage environment impacts the user experience)
10. User role (how the role of the user (management, HR, . . .) affects the user experience)

These elements of experience were chosen based on existing user experience evaluation methods and the requirements of Patria. Elements that affect digital employee experience, data quality and reliability of the software were emphasized when selecting these.

About the importance of reliability and robustness, Norman writes: *"Designers need to focus their attention on the cases where things go wrong, not just on when things work as planned. Actually, this is where the most satisfaction can arise: when something goes wrong but the machine highlights the problems, then the person understands the issue, takes the proper actions, and the problem is solved. When this happens smoothly, the collaboration of person and device feels wonderful."* [11]

Because the user experience evaluation method builds on ideas of SUXES and the experimental user experience evaluation, I also created statement pairs that were fit for the needs of this user experience evaluation and the evaluation context it was expected to be used in by Patria. See table 4.1 These statement pairs are the operationalization of the elements of experience, and are to be used in the user experience evaluation questionnaire, which forms the basis of the user experience evaluation score of the software that is evaluated. Though, these elements of experience can be modified, as mentioned, to fit

Element of experience	Negative statement	Positive statement
Efficiency	Completing my tasks on the application is inefficient.	Completing my tasks on the application is efficient.
Utility	The functions of the application are useless for completing my tasks on the application.	The functions of the application are useful for completing my tasks on the application.
Learnability	Using the application is hard to learn.	Using the application is easy to learn.
Cognitive load	Using the application is exhausting.	Using the application is not exhausting.
Error-free use	Using the application is not error-free.	Using the application is error-free.
Robustness	The application does not function error-free.	The application functions error-free.
Accessibility	The application is not accessible.	The application is accessible.
Aesthetics	The application is aesthetically unpleasant.	The application is aesthetically pleasant.
Usage environment	Using the application requires a distraction-free environment.	Using the application does not require a distraction-free environment.
User role	The application is ill suited for completing my tasks.	The application is well suited for completing my tasks.

Table 4.1. *The statement pairs (negative-positive) of the customizable user experience evaluation method, based on SUXES measures.*

whatever the evaluation context demands, but they should always aim to measure these things in some way, as these are fairly fundamental aspects of user experience.

4.3 The User Experience Evaluation Method

In the following subsections I will describe the evaluation process, how the user experience score is calculated, explain the user experience questionnaire and how to supplement it with qualitative data.

It should be noted, that the user experience evaluation method is modifiable. If certain tools, like Digital Experience Monitoring -tools or internal employee feedback tools, are available, they can be used, but they aren't mandatory.

4.3.1 The Evaluation Process

Before evaluation:

- Identify and define the purposes and goals of the evaluation case.
- Recognize the challenges, limitations and opportunities presented by the evaluation case.
- Recognize the use cases and purposes of the software to be evaluated.
- Recognize the spectrum of roles of the users of the software to be evaluated.
- Identify the usage environments of the software to be evaluated.
- Recruit the participants to the evaluation based on the purposes and goals of the evaluation case.
- Decide which parts of the user experience evaluation method to use in the evaluation case.
- Decide how to use the chosen parts of the method based on the purposes and goals of the evaluation case.
- Customize the method based on the decisions and choices made.

During evaluation:

- Distribute the user experience questionnaire to the participants.
- Introduce the evaluation case to the participants.
- Distribute the background questionnaires, send invitations to test sessions and reserve times for them, if applicable.
- Introduce the software to be evaluated to the participants, particularly if they aren't experienced users of the software to be evaluated.
- Instruct on how to use the software during the test session. One extreme is free use of the software to be evaluated under observation of the UX researcher. The other extreme is more akin to an usability test with pre-defined tasks to be completed.
- Collect recordings, logs, observations and other supporting data.
- Interview the test participants.

After evaluation:

- Conduct an accessibility evaluation if applicable.
- Analyse data of interest, qualitative and/or quantitative, using applicable data analysis methods.
- Calculate a score for the software that was evaluated based on the user experience questionnaire.
- Make conclusions from the score based on the analyses conducted.
- Draft improvement ideas for the software that was evaluated, if applicable.

Rating	Score
1	0
2	25
3	50
4	75
5	100

Table 4.2. *How different ratings of statement pairs that correspond to an element of experience on the user experience questionnaire are scored.*

- Identify what an ideal software would look like for this use case. What are the requirements it should meet? This can be used as a reference for software procurement and when evaluating other applications in the same domain.

4.4 Calculating The User Experience Score

The basis of the score is the user experience questionnaire. It uses a five-step scale to measure how respondents rate the user experience of the evaluated software. Each element of experience has equal weighting. This data is used to calculate a score, which can be adjusted and analysed based on other data to create comparable results between different software.

Table 4.2 illustrates how different ratings of statement pairs that correspond to an element of experience on the user experience questionnaire are scored. These scores are summed together to form the user experience score for the software. The user experience score is the average of the scores each element of experience receives in the evaluation, so it's always between 0-100 even if less or more than ten elements of experience are evaluated. If not all elements of experience are evaluated in the user experience questionnaire, the scoring should still be done in a similar manner using whatever other evaluation method was chosen for that element of experience. An example scoring can be found in subsection 3.3.3

4.4.1 The User Experience Questionnaire

The user experience questionnaire is the basis for the user experience score of the evaluated software. An example application of this questionnaire can be found in subsection 3.3.1

The user experience questionnaire consists of negative-positive statement pairs that correspond to the elements of experience that are chosen for the evaluation based on the demands set by the evaluation context. These statements are judged on a five-step scale

by the respondents, who are also users of the software to be evaluated. The background information questionnaire can be combined with the user experience questionnaire, if considered convenient in the evaluation context. Other questions can also be asked as deemed necessary or useful. Keskinen notes, that *"Open questions are very worthwhile to include in the experiences questionnaire to collect qualitative data in case there will not be any kind of interview, for instance."* [13] However, you should be considerate of the time it takes to respond to the questionnaire.

Speaking of the time it takes to respond, sample size should also be carefully considered. *"No matter which sampling method is used, it is important to carefully determine the target sample size for the survey, i.e., the number of survey responses needed. If the sample size is too small, findings from the survey cannot be accurately generalized to the population and may fail to detect generalizable differences between groups. If the sample is larger than necessary, too many individuals are burdened with taking the survey, analysis time for the researcher may increase, or the sampling frame is used up too quickly."* [33] This wisdom should also be taken into account when considering what sample size to use with the user experience evaluation questionnaire.

Due to this user experience questionnaire being mandatory, the user experience evaluation method results at least in quantitative data consisting of comparable user experience ratings of specific statements relating to elements of experience, similar to SUXES and the The Experiential User Experience Evaluation Method. The major differences from these two methods are that a score is formed and the method is highly customizable for the needs of the evaluation context. In a future iteration, the method will also be more automatic, with the integration of digital experience monitoring tools and automated internal employee feedback tools.

4.4.2 Usability Evaluation

While the user experience evaluation questionnaire gathers quantitative data on the experienced user experiences of the users with the evaluated software, the usability evaluations gather subjective feedback and comments, observation data, and so on, which can be used to understand the experiences per se and possible reasons for them. [13] An example usability test can be found in subsection 3.3.4

1. Determine the most commonly done tasks, the most important tasks and the most time consuming tasks done on a software. In the absence of monitoring tools or other easy ways to get this knowledge, interviews of expert users as well as observation of the software in daily use work.
2. Recruit a variety of users. Users with different but relevant roles as well as both expert users and people unfamiliar with the software to participate in the usability

tests.

3. Instruct on how to use the software during the test session. One extreme is free use of the software to be evaluated under observation of the UX researcher. The other extreme is more akin to an usability test with pre-defined tasks to be completed. Ignore next two steps if not defining tasks.
4. Design tasks that test how efficiently the user can accomplish the common, important and time consuming tasks. Measure the cognitive load of these tasks, noting the usage environment if possible.
5. Design tasks that can or will cause errors to understand how the software handles such situations.
6. Compare the users with different levels of experience to evaluate the learnability of the software.
7. Conduct a post-test interview to measure the utility and aesthetics of the software.
8. Analyze the results and use them to understand the score the software received.

4.4.3 Accessibility Evaluation

Patria employees use software in a wide variety of contexts. When gathering the requirements for the user experience evaluation method, I found that Patria has employees working in large assembly halls, where they are periodically required to travel to a PC endpoint of some sort to interact with a software in order to follow process and protocol. In addition, they would struggle with the software that was designed for office environments.

While process and protocol are important, it would be practical if walking long distances in a large assembly hall just to interact with some software could be avoided, and employees could do more productive work instead. This case illustrates why usage environment and accessibility are important elements of experience to include in this user experience evaluation method.

It is recommended to conduct accessibility evaluations, as instructed in the Web Content Accessibility Guidelines (WCAG), [34] to score the accessibility element of experience as part of this user experience evaluation method. However, it's not strictly necessary as part of this user experience evaluation method. The accessibility of the evaluated software can also be scored based on the ratings other elements of experience received, as well as by referencing other supporting data, like usability evaluations and observations.

There's also the option of further iterating on this user experience evaluation method to include automated accessibility evaluation software, particularly those for authenticated environments. Automated evaluation tools are software programs that examine the code of Web pages to determine if they conform to a set of accessibility guidelines that are often

based on the Web Content Accessibility Guidelines Version 2.0. A study in 2014 studied ASES, one such tool. The study found ASES to have problems of efficiency, interaction, validity, and reliability in the results presented. [35] Despite this, I recommend Patria look into this method of automating accessibility evaluations to further save resources while improving digital employee experience, as things may have since improved.

5. DISCUSSION

The purpose of this thesis was to produce something that Patria could use to evaluate the user experience of software. Patria also hoped that this thesis would help them improve their digital employee experience. Based on these needs I defined two research questions, which I answered in this thesis:

- How to design, construct and test a custom user experience evaluation method to aid in software consolidation and procurement?
- What should Patria do to methodologically improve their digital employee experience?

The answer to how to design a custom user experience evaluation method was to conduct a literature review and use the results of that to understand what would be the best way to meet the requirements that Patria had for the user experience evaluation method. The construction of the method was done by referencing existing methods and discussing the pros and cons of those in Patrias' evaluation context with the leaders of Patria group IT as well as my thesis supervisor. I had also planned to use a focus group for this purpose, but it would have been excessive and was replaced by weekly meetings with Patria group IT leadership. The best way to test the construction turned out to be a weak market validation of the construction using Case Efecte, detailed in section 3.3

Thanks to the literature review and collaboration with stakeholders, I constructed the customizable user experience evaluation method and validated it using Case Efecte. The result of the case study was a user experience score of Efecte, which consisted of eight distinct scores that eight distinct elements of experience of Efecte received from the respondents of a user experience evaluation questionnaire. The questionnaire results were supported by other data and analysed. The analysis revealed, that daily users of Efecte had a statistically significant difference in the way they experienced Efecte, specifically they found it more efficient to use. The reason for this difference could also likely be uncovered with further analysis of the results.

The customizable user experience evaluation method constructed as part of this thesis is the main deliverable of this thesis. The scores it produces when used to evaluate software can be of aid in various projects related to changes in the application landscape at Patria. These scores can be analyzed to uncover the reasons behind positive and

negative elements of the user experience of the evaluated software.

As for how Patria should improve their digital employee experience, the first step would be to utilize the user experience evaluation method constructed as part of this thesis in Patria's various projects related to changing and procuring software. Taking into account the results of the evaluations produced in decision making processes would ensure that in the future the software in use at Patria would have a better user experience on average. This is also how the user experience evaluation method would aid in software consolidation.

A part of the constructive research approach is to show the theoretical connections and the research contribution of the solution concept. I found that the Experiential User Experience Evaluation Method, built based on SUXES and the Experience Pyramid, could be referenced to create a customizable user experience evaluation method that would meet the user experience evaluation needs of a company undergoing shifts in their application landscape.

In conclusion, this thesis and the user experience evaluation constructed as part of it will benefit Patria when they execute the various projects related to changing their application landscape. The benefit will be, that the changes can be done in a way that accounts for the user experience of software Patria retains and procures, which will lead to a improvement in digital employee experience. This thesis also highlights the positive effect of improved digital employee - and user experience on data quality.

5.1 Recommendations

According to research from Gartner, *"The acceleration of digital work has been transformational. To deliver an ambitious digital workplace, application leaders must build an IT-led team seeking partnerships with HR, corporate real estate, business leaders and employees to deliver on the future of work."* They claim, that *"exclusively IT-led digital workplace (DW) programs are too technology-focused and lack alignment to employees' ways of working and the business capabilities they deliver."* [16] Based on this research I decided to build a focus group, as this thesis could be considered a transformational digital workplace program led by IT leaders.

The focus group was gathered from among the employees of Patria's various different group functions. The purpose of the focus group would have been be to evaluate how well the method meets its requirements. It proved unnecessary, with how involved the Patria IT leaders were with the project. The members of the focus group would also have been able to assist in further refining of the requirements as well as identifying limitations. Semi-structured or informal interviews could also have been conducted with select focus group members, should the need have arisen. The focus group could also have been used to gather perspectives from across the company functions make the method more

user-centric. However, I found that I was able to gather these without the aid of the focus group.

I recommend that Patria doesn't waste this pool of people who volunteered to participate in a DEX-improving initiative like this. They could be utilized as a backbone for a special interest group that can shape the direction of DEX improvements at Patria in the future. I recommended creating such a special interest group in the DEX road map I created as part of this thesis in subsection 2.3.4

I also recommend that Patria includes the list of questions I created as part of case Anselmo vs. Patria Group IT Survey, found in subsection 3.2.3 Whether Patria takes into use Anselmo or Patria Group IT Survey, the questions would provide valuable data about user experiences of software. This data could be used as part of user experience evaluation design when using the method constructed as part of this thesis. The tools, or whichever one of them is taken into use, could also be used to distribute the user experience evaluation questionnaire to specific users at Patria. This could be a much more convenient method for questionnaire distribution than Microsoft Forms and email lists.

As it is, the applicability of the construction, the final aspect to examine as part of the constructive method, is for application landscape rationalization. It requires further refinement to be utilized as part of software procurement. In their dissertation, Roto writes that they have found the key to user experience evaluation to be "*to analyse whether the product met the expectations that the user had before starting to use it*." [36] Due to the nature of Patrias needs for the user experience evaluation method, expectations are challenging to take into account. When evaluating software that has been in internal use for a long time, expectations likely will not much differ from the experienced user experience due to how experienced the users will be with the software. In such cases, despite what Roto asserts, expectations will not be of much value for the sake of evaluation.

The Experiential User Experience Evaluation Method and SUXES also both highlight the importance of evaluating the difference between user expectations of the user experience and the actualized user experience they had once they tried the software. [13] [15] Analyzing this difference will be key for conducting user experience evaluations of software Patria will procure. The user experience evaluation method in its' current state is more suitable for use in situations where we have access to a pool of users with experience using the software that is to be evaluated.

In addition to using the gathered expectations for comparisons to actualized user experiences, they can also be used to improve the requirements that Patria sets for the software they wish to procure. My final recommendation is for Patria to iterate on this user experience evaluation method by including the changes required to use this method for software procurement.

REFERENCES

[1] *Patria corporate structure and shareholder voting rights*, 2022. [Online]. Available: https://www.patriagroup.com/governance/corporate-structure-and-shareholder-voting-rights.

[2] *Patria in brief*, 2022. [Online]. Available: https://www.patriagroup.com/about-us/patria-in-brief.

[3] D. Wilson and S. Downes, "Innovation insight for the digital employee experience", 2021. [Online]. Available: https://www.gartner.com/en/documents/4008928.

[4] P. Morville, "User experience design - user experience honeycomb", 2004. [Online]. Available: https://semanticstudios.com/user_experience_design/.

[5] T. Cipolla, S. Kumar, and S. Downes, "2023 strategic roadmap for digital workplace infrastructure and it operations", 2023. [Online]. Available: https://www.gartner.com/en/documents/4612099.

[6] E. Kasanen, K. Lukka, and A. Siitonen, *The constructive approach in management accounting research*, 10, 1993, p. 243.

[7] K. Lukka, *Konstruktiivinen tutkimusote*, 2001. url: https://metodix.fi/2014/05/19/lukka-konstruktiivinen-tutkimusote/.

[8] *User experience white paper - bringing clarity to the concept of user experience*, 2011. [Online]. Available: https://experienceresearchsociety.org/wp-content/uploads/2023/01/UX-WhitePaper.pdf.

[9] A. Butler, T. Feaster, D. Miller, and J. Barbone, "Using iterative user experience design to improve electronic clinical outcomes assessment data quality", *Value in Health*, vol. 20, 2017.

[10] M. Davids, U. Chikte, and M. Halperin, "Effect of improving the usability of an e-learning resource: A randomized trial", *Advances in Physiology Education*, vol. 38, 2014.

[11] D. Norman, *The Design of Everyday Things: Revised and Expanded Edition*. Basic Books, 2014, 347 pp.

[12] *Ergonomics of human-system interaction. part 11: Usability: Definitions and concepts (iso 9241-11:2018)*, 2018. [Online]. Available: https://www.iso.org/standard/63500.html.

[13] T. Keskinen, *Evaluating the user experience of interactive systems in challenging circumstances*, 2015.

[14] *Expresso - ux evaluation methods*, 2024. [Online]. Available: https://experienceresearchs
org/ux/evaluation-methods/.

[15] M. Turunen, J. Hakulinen, and A. Melto, “Suxes - user experience evaluation method for spoken and multimodal interaction”, *INTERSPEECH 2009, 10th Annual Conference of the International Speech Communication Association*, 2009.

[16] T. Paulman and D. Wilson, “Build a better digital workplace team to enable a hybrid future of work”, 2021. [Online]. Available: https://www.gartner.com/en/documents/4006586.

[17] J. Wong and M. Cain, “Use gartner’s dex blueprint to mature your digital workplace strategy”, 2023. [Online]. Available: https://www.gartner.com/en/documents/4457099.

[18] D. Wilson and A. Stanish, “Market guide for dex tools”, 2023.

[19] D. Wilson and A. Stanish, “How to successfully deploy a dex tool”, 2023.

[20] J. Brooke, “Sus—a quick and dirty usability scale”, *Usability Evaluation in Industry*, 1996. [Online]. Available: https://www.researchgate.net/publication/228593520_SUS_A_quick_and_dirty_usability_scale.

[21] P. So and B. Gore, *Nasa task load index (tlx)*, 2020. [Online]. Available: https://humansystems.arc.nasa.gov/groups/tlx/.

[22] L. Page, *Beyond the nps: Measuring perceived usability with the sus, nasa-tlx, and the single ease question after tasks and usability tests*, 2018. [Online]. Available: https://www.nngroup.com/articles/measuring-perceived-usability/.

[23] “The pros and cons of net promoter”, *Professional remodeler*, vol. 14, no. 2, p. 18, 2010.

[24] Efecte, “Enterprise service management”, 2024. [Online]. Available: https://www.efecte.com/solutions/enterprise-service-management.

[25] M. Sengpiel and D. Dittberner, “The computer literacy scale (cls) for older adults-development and validation”, 2008.

[26] S. MacKenzie, *Human-Computer Interaction, 2nd Edition*. Morgan Kaufmann, 2024.

[27] “Powerful and consistent analysis of likert-type ratingscales”, *Proceedings of the 28th International Conference on Human Factors in Computing Systems*, 2010.

[28] F. Wilcoxon, “Individual comparisons by ranking methods”, *Biometrics Bulletin*, vol. 1, 1945.

[29] K. Grace-Martin, “Can likert scale data ever be continuous?”, 2008. [Online]. Available: https://www.theanalysisfactor.com/can-likert-scale-data-ever-be-continuous/.

[30] G. Wickström and T. Bendix, “Hawthorne effect” — what did the original hawthorne studies actually show?”, *Scandinavian Journal of Work, Environment & Health*, vol. 26, 2000.

[31] J. Nielsen, *10 usability heuristics for user interface design*, 1994. [Online]. Available: https://www.nngroup.com/articles/ten-usability-heuristics/.

[32] *Requirements of the act on the provision of digital services*, 2023. [Online]. Available: https://www.webaccessibility.fi/requirements-of-the-act-on-the-provision-of-digital-services/.

[33] H. Müller, A. Sedley, and E. Ferrall-Nunge, "Survey research in hci", *Springer Science*, 2009.

[34] B. Caldwell, M. Cooper, L. Reid, and V. G., "Web content accessibility guidelines (wcag) 2.0", 2008. [Online]. Available: https://www.w3.org/TR/WCAG20/.

[35] E. M. Pivetta and D. S. Saito, "Automated accessibility evaluation software for authenticated environments: A heuristic usability evaluation", *Springer International Publishing*, vol. 8516, 2014.

[36] V. Roto, *Web browsing on mobile phones - characteristics of user experience*, 2006, p. 76. [Online]. Available: http://lib.tkk.fi/Diss/2006/isbn9512284707/isbn9512284707.pdf.