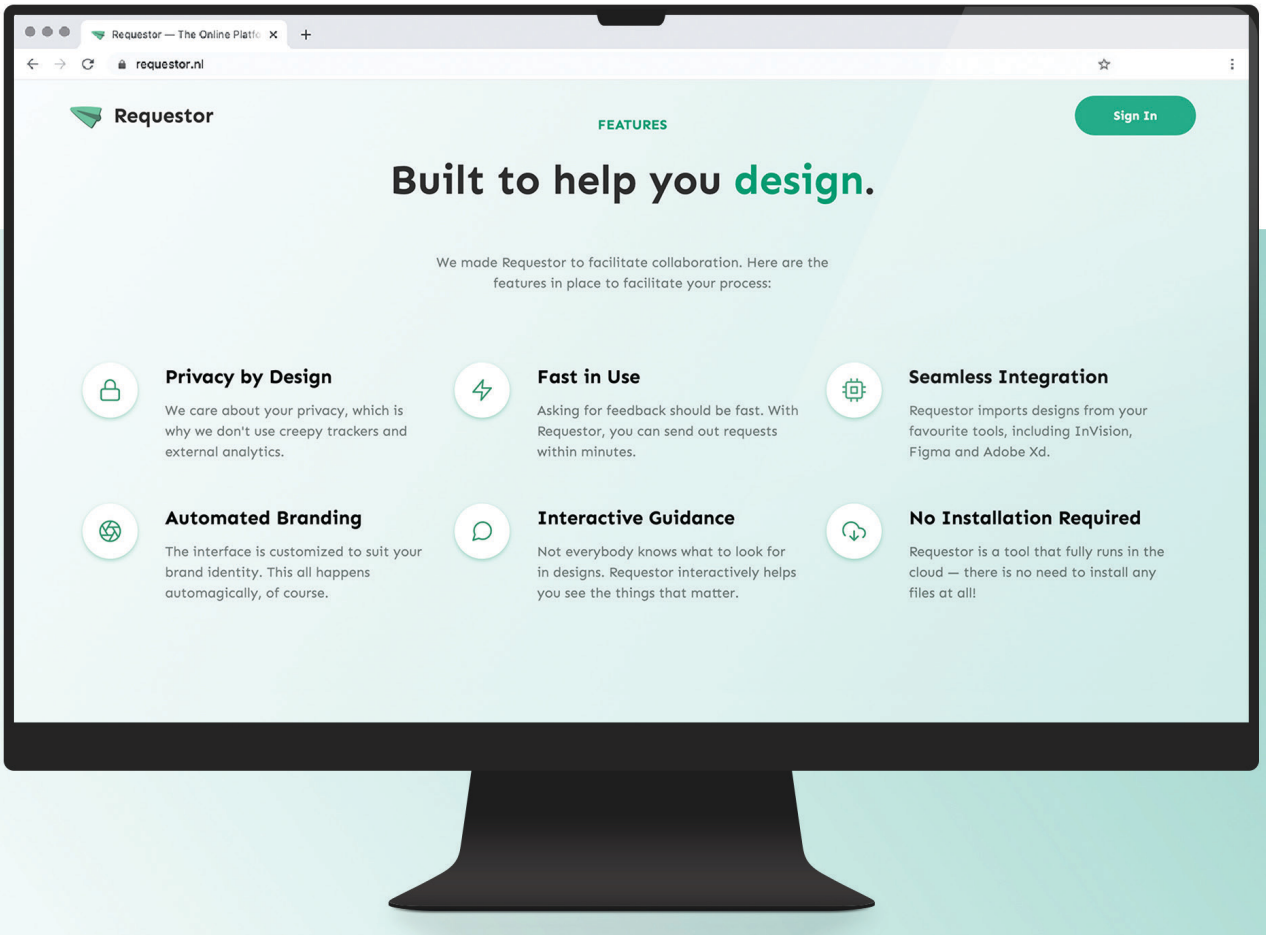
REQUESTOR

'The Online Platform to Share UX Works-in-Progress'



Eindhoven University of Technology

Faculty of Industrial Design

Final Master Project by

Arthur Geel

Supervised by

Dr. Ir. Javed Khan

Published on

June 11, 2020

Word count: 14.226

Executive Summary

This report describes the project *'Requestor — The Online Platform to Share UX Works-in-Progress'*, the Final Master Project by Arthur Geel. This project was conducted as part of the Industrial Design M.Sc. programme at the Eindhoven University of Technology.

The aim of this report is to communicate the process and outcome of this project, giving insight in my overall competence of design.

With this project I seek to explore the dynamic of User Experience (UX) Designers and their enterprise environments, with the aim to contribute a strategy to resolve problems that arise from the limited industry understanding and buy-in on the role of UX Design in enterprise environments.

In this project I seek an answer to the question:

How might we create a better collaborative space for UX Designers and their environment?

This project is framed by a review of literature and trends, explaining the nuances in the problem statement. The design process is described, clarifying the design decisions made when taking an early concept to the final high-quality demonstrator, embodying my vision on addressing the outlined problem.

The later sections of this report describe the set-up, implementation and analysis of a summative investigation of the design concept. In this, research participants operated one of two variants of the design concept to perform a heuristic evaluation of a fictional design case.

Final Master Project by Arthur Geel

Author
Arthur Geel
s149144 — 0907552
a.j.geel@student.tue.nl

First Examiner
Prof. Dr. Panos Markopoulos

Second Examiner
Dr. Ir. Javed Khan

Third Examiner
Dr. Mathias Funk

Disclaimer
All work presented in this document is created by the author unless specified otherwise. Parts of this report have been based on the FMP Proposal written by the author. This original proposal is included in its entirety in the appendices at the end of this report.

Eindhoven University of Technology
Department of Industrial Design
June 11th, 2020

Word count: 14.226

Table of Contents

SECTION ONE

Project Framing

In the first section the context behind the project is described. The section explains the underlying theoretical background, the competency profile and design vision of the author. Furthermore, the design challenge which was formulated in the original project proposal is presented.

1— Project Framing

1.1 Introduction

The Rise of User Experience Design
The increasing importance of computers and the internet in our daily lives has resulted in a discipline that specialises in designing for this context: User Experience (UX) Design. UX Design is characterized by its process which integrates the elements of research, ideation and prototyping to create products, systems or services that elegantly solve problems whilst creating engaging experiences. Similar to other disciplines of design, the UX Design process goes hand in hand with the Design Thinking methodology [5, 25].

An emphasis on UX Design during product development processes is known to increase business success: Returns on Investments are achieved through a reduction in costs for critical areas such as development time [13, 47], maintenance [45] and providing user support [8], while increases are seen in areas such as user adoption, user satisfaction, revenue growth and productivity [44, 45]. However, these benefits have one key prerequisite — those in the enterprise environment need to have sufficient understanding of- and involvement in the design process [31].

Current Challenges in UX
As the awareness on the impact that UX has on business has grown significantly over the past decade, so has the demand for UX Designers [6]. The introduction of a new discipline came with its challenges, as there is ambiguity about what UX Design actually is and the responsibilities practitioners have [31]. For example, while the Design Thinking framework [5] has five main stages *(empathising, defining, ideating, prototyping, testing)*, enterprise environments commonly misconceive designers to merely conceptualize and create the designs (i.e. ideating & prototyping), failing to recognize that this progression is facilitated by other stages of Design Thinking. This misunderstanding results in imbalanced design processes with shortened phases and misaligned priorities, ultimately impacting the UX Designer's ability to perform [10, 44, 45].

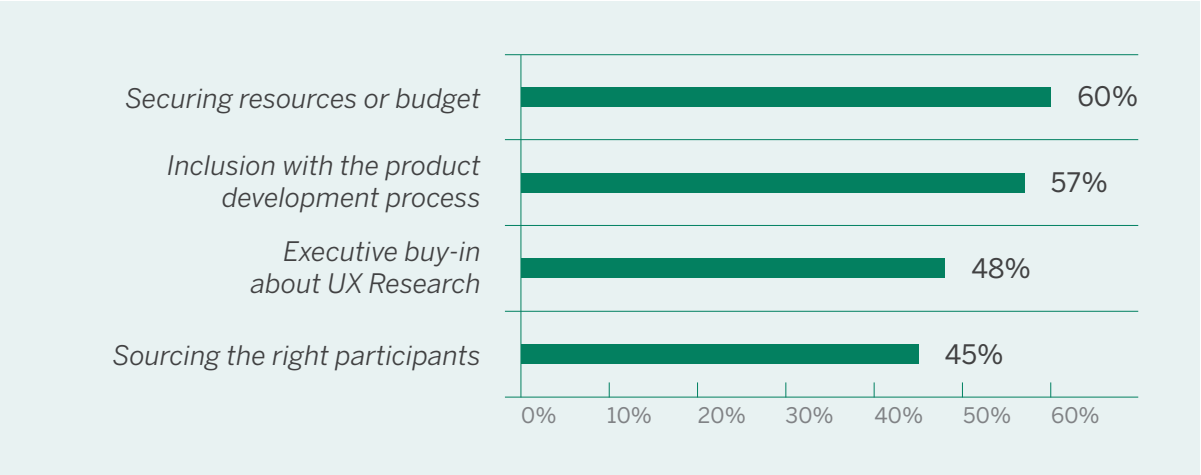


Figure 1: Key challenges faced by UX Researchers in enterprise positions. Adapted from *The State of UX in Enterprise 2018, [52].*

This (lack of) understanding with respect to UX Design and UX Designers is not a recent trend. In fact, it was first recognised more than two decades ago by Jakob Nielsen, who defined eight total stages of *Corporate UX Maturity*. In his work, Nielsen describes a spectrum which ranges from *'hostility towards usability'* to *'user-driven corporations'* [38, 39]. In the same publication, Nielsen offers advice on how to improve Corporate UX Maturity.

More recently, we have seen new academic interest as others build upon the work to create robust models to measure and improve UX Maturity in an effort to help the industry adapt to this new discipline [9, 43, 32].

Nevertheless, recent industry surveys illustrate that the problems outlined by Nielsen in 1996 are still prevalent. A recent survey among over 3,000 designers saw professionals tie their professional top challenges to UX Maturity [54]. Their top challenges included improving UX consistency (59%); testing designs with end-users (53%); securing UX budget or resources (40%) and getting buy-in or understanding from executives (38%). A 2019 edition of an annual UX Designer survey saw similar results [53]: prevailing challenges were including research within the product development process (64%); sourcing the right participants (50%), securing resources or budget (49%) and getting executive buy-in about UX research (49%). Evidently, the integration of the UX Design profession in enterprise environments is a work in progress.

Even though some challenges are becoming less pressing, (the challenge of securing resources or budget decreased from 60% to 49% over the past year [52, 53]) it is evident that the challenges resulting from low UX Maturity will continue to impact the industry for the foreseeable future, especially considering new in-house UX teams are on the rise [32], increasing the amount of environments where UX is newly introduced.

1.2 Context

Identity as a Designer
My name is Arthur Geel — I'm a MSc student Industrial Design at the Eindhoven University of Technology. During my bachelor degree at the ID department I became familiar with a broad range of design applications *(from design research to automotive design)*, but was mostly interested in User Experience Design.

Being a designer in the 21st century means adapting to the needs of a rapidly changing world by having competencies in areas that are not traditionally design-related. My response to this shift is to develop myself as a *Pi-Shaped Designer* [11, 49] . Besides being familiar with a broad range of product design essentials, a Pi-Shaped Designer has two areas of competence in which they excel.

For me, those two areas are *designing for digital user experiences* and *the technological realization of digital products*. I believe that such a varied skillset allows me to better understand contexts by analysing them from both perspectives. Consequently, the two competency areas within the Industrial Design Eindhoven Education Framework I have decided to specialize in are *User & Society* and *Technology & Realization*.

In my process I put an emphasis on creating a thorough understanding of the context and its stakeholders, where I use skills and attitude developed in the User & Society competence to guide me. More specifically, I employ UX methodologies such as design ethnography [26] , contextual inquiries [26] and customer journey mapping [48] to inform my process.

Afterwards, I give shape to the insights by designing digital prototypes in increasing levels of fidelity — from wireframes [2] to complete design system [13] — using Figma [15], my tool of choice for designing digital products and systems. These design artifacts allow me to further explore the design space by conducting in-context evaluations once more tapping into the User & Society competency.

Figure 2: A photograph of me taken in Mostar, Bosnia and Herzegovina, 2019.

Finally, I enjoy breathing life into prototypes by developing them with my skills in the Technology & Realization competency. My preferred tool of choice for this is a front- end development stack (HTML/CSS/JavaScript), using Git for version control. During my Master's degree, Programming has become an integral part of my designer's identity and toolkit.

Throughout my time at the TU/e I was driven to complement my academic knowledge with professional experience. In 2017 I had a six-month internship as a Junior Interaction

Figure 3: Nuntius: an embodiment of my design vision — making the news accessible whilst respecting its end-user.

Designer at creative digital agency Momkai, which I discuss in detail in my portfolio [18] .

Additionally, from September 2018 until January 2020 I have been working part-time as the sole UX Designer at SPIE Nederland. In this position I am responsible for re-designing the usability and user experience of legacy applications.

Experiencing the differences in working as a designer at these companies with very differing attitudes towards design was one of the drivers for this project — I felt that the benefits of UX Design should be available to all, regardless of UX Maturity levels.

The time I spent working as a designer further strengthened my identity as a designer: I want to keep developing myself as a Pi-Shaped UX Designer in my future career.

Vision on Design

Computers and the internet have revolutionalised our way of living. We increasingly perform complex tasks that were previously deemed impossible, including connecting with peers all over the world, automating tasks that are repetitive and sharing vast amounts of data. We have made the world more accessible: computers enable all to contribute to our society, regardless of their location, age, gender or disabilities.

In contrast with physical products, the landscape of digital products is a lot more dynamic. By removing physical materials and real-world shipping from the equation, we see new products make it to the market faster, and see existing products receive new features regularly. Because of this, barriers that prevent us from being productive are rapidly disolving.

However.. that is not always the case. The complexity of the tools we use in everyday life is rapidly expanding, requiring more training and expertise to operate. This clashes with making the world more accessible: not everybody is able to participate anymore as their skillset does not allow them to participate.

Furthermore, we are seeing negative societal trends regarding the digital products we use. Our privacy is being eroded through the increasing intensiveness of web- tracking and transfer of personal data to third parties, and our behaviour is being influenced on a large scale by the party that pays the most.

With my design I try to take a stand against these developments. I am driven to create digital tools and systems that not only make the internet a more accessible place, but also more respecting of the persons operating it.

The Final Master Project
The Final Master Project (FMP) is an educational activity that concludes the Industrial Design MSc programme at the TU/e. In the activity that lasts one full semester the student demonstrates their unique design expertise, vision and identity by working on a topic which fits in with the department's mission on emerging technologies in a societal context.

With this design project I demonstrate the two core competency areas of my personal expertise profile: *User & Society* and *Technology & Realisation*. These competencies have been the guiding factors in the design process of this project — empowering me to transform emerging technologies into experienceable prototypes, placed in a societal context and evaluated to understand the implications of the design.

This project was conducted as part of the Research, Design and Development (RDD) track, which focuses on preparing its students in developing the attitude, skills and knowledge required for a career in (research and-) design departments of commercial organizations.

For this reason, I carried out the project in a way similar to how these departments operate: basing my approach on a hybrid version of the Design Thinking- and Dynamic Systems Development methodologies [5, 42]. Essentially, these allow me to integrate the principles of human-centered design whilst respecting the time-bound nature of work done in commercial environments.

1.3 Project Themes

The foundation of this project is further built up from two themes that substantiate the design decisions made during the process. In this section, a supplementary overview of the context is provided.

UX Maturity Strategies
Where current attempts at addressing the challenges arising from low UX Maturity environments are mostly aimed at providing long-term support for industry executives to diagnose and improve their UX Maturity [9, 32, 43], this project takes a different approach by focusing on helping UX professionals themselves adapt to their challenging environments.

The rationale behind this is that while long-term support for executives may solve the problem once the company has raised their UX Maturity, identical problems will continue to occur in places where UX is newly introduced, which is still on the rise [32]. Environments where UX is newly introduced are are especially prone to erroneous implementations of methodologies as the industry conditions of limited time and resources have a greater say if its benefits are unclear.

Oevad and Larsen propose an interesting strategy that may resolve the challenges: training your software developers in UX to divide workload and foster a stronger UX culture [41]. In order to do so, the authors have created simplified versions of UX methodologies, bundled with documentation materials under the name 'UX Toolbox'.

Oevad and Larsen validated their strategy with a number of case studies. Software developers, previously unfamiliar with UX methodologies, were given access to the UX Toolbox and asked to perform an A/B Usability test. The researchers concluded that their approach is feasible in enabling developers to successfully perform these UX tasks.

In a related study which examined the intricate challenges encountered with the growing importance of UX in software development, Kashfi et al. highlight the issue on power struggles between designers and developers [24]. The interviewees of the study believe that a viable strategy to resolve the power struggles would be to increase the knowledge and awareness about UX in organisations, and vocalize UX responsibilities to other stakeholders involved. A suggestion would be to make the methods more accessible and more well-integrated in the routine of others.

UX Design Evaluation Methodologies

Evaluation is a core component of UX Design: it allows for the assessment of how a person perceives a system before, during or after interacting with it — serving as an important performance indicator in how well a system is designed. However, evaluating user experiences is not simple by any standard: experiences are highly subjective, dependent on context and may change over time [27].

It is no surprise that evaluating UX is an area that receives little priority in most enterprise environments [24]. In fact, UX evaluation is often one of the things that are scrapped when time or budget are limited. However, in order to better understand why this is, we need some more context on evaluation methodologies.

Nielsen Norman Group recognises two types of evaluations: formative and summative [23]. Formative evaluations generally focus on qualitative data: concentrating on learning what parts of a system work well, or do not, and understanding why that is. Formative evaluations play into the iterative aspect of UX Design: the output provides suggestions for improving the system, informing the design process.

Summative evaluations focus on how well the designed solutions perform, and usually are conducted towards the end of a project cycle. They may provide crucial information with regard to business goals: will the system perform well enough, or should it be altered before it can be released?

For both types of evaluation there are number of commonly used methodologies. Perhaps the best known method is the *usability test* [12]. In this method, the designer observes representative end-users while they carry out tasks using a system. Often, the designer asks the end-user to verbalize their thoughts, giving them a better understanding of how the system is perceived and how usable it is.

A/B Tests are a variant of the usability test. With A/B Testing, one out of two slightly different variants of a system are shown to end-users, after which metrics such as click rates or questionnaire responses determine which variant is superior for reaching the goals.

When doing usability tests, a broad range of metrics can be employed which are chosen based on the area the designer is interested in. For example, the *Technology Acceptance Model* [28] provides insights in a person's acceptance of a system based on *Perceived Usefulness* and *Perceived Ease of Use*. The *System Usability Scale* is an industry standard tool for gathering subjective ratings of a system's usability [4]. Hassenzahl's *AttrakDiff* can be used to measure the hedonic and pragmatic qualities a system evokes for end-users. This is just a small selection of the academic methods that have been developed to evaluate interactive systems.

A major drawback of usability tests that involve (representative) end-users is the investment

required to run them — both in terms of financial investment as well as the involvement that is required. The low priority that UX evaluation gets in Low UX Maturity environments starts to make more sense.

Fortunately, there are UX evaluation methods that can be conducted without end-users, such as the expert review [20]. For this type of evaluation, one or more UX experts inspect (parts of a) system in order to identify potential issues in usability. Usually, expert reviews are done based on a list of guidelines, be it specifically compiled for the specific context, or using a generalized list. The most widely used list of guidelines for user interfaces is Jakob Nielsen's Usability Heuristics [36]. This set of ten heuristics are applicable to most projects within web and mobile applications, which is why they receive broad adoption.

1.4 Design Brief

With this project I seek to explore the dynamic between User Experience (UX) Designers and their enterprise environments, with the aim to contribute towards resolving problems that arise from the limited industry understanding and buy-in on the role of UX Design in industry environments. In short, with this project I seek an answer to the following question:

- *How might we create a better collaborative space for UX Designers and their environment?*

This design challenge is guided by the conditions of the modern working environment as outlined in previous sections: limited time, limited resources yet a need for consistent, high-quality UX output.

From these circumstances, a number of requirements for the potential design concept can be derived:

a. *The concept should fit in well with the way UX Designers operate in their respective environments, specifically with regard to limitations in time and resources.*

b. *The concept should contribute to the consistency and quality of the UX Designer's design output.*

c. *The concept should contribute to the understanding of UX Design by stakeholders in the respective environments.*

In the project proposal (Appendix F), the following design concept was presented:

- *A system that allows UX Designers to more easily share their works-in-progress in order to receive formative feedback from colleagues and related stakeholders.*

A schematic overview of the system can be seen in Figure 4. The system has two aims: firstly, it aims to increase the iterative capabilities of the UX Designers by addressing the need for formative design feedback. Secondly, it aims to resolve the tensions existing in low UX Maturity environments through a more inclusive involvement of stakeholders.

Additionally, a wireframe representation of the design concept was presented in the project proposal (Figure 5). In this, we can see the two components of the system. The system contains a way to share the experience of the design work-in-progress (on the left), and a way to guide non-experts in understanding- and evaluating the design decisions taken in the work-in-progress

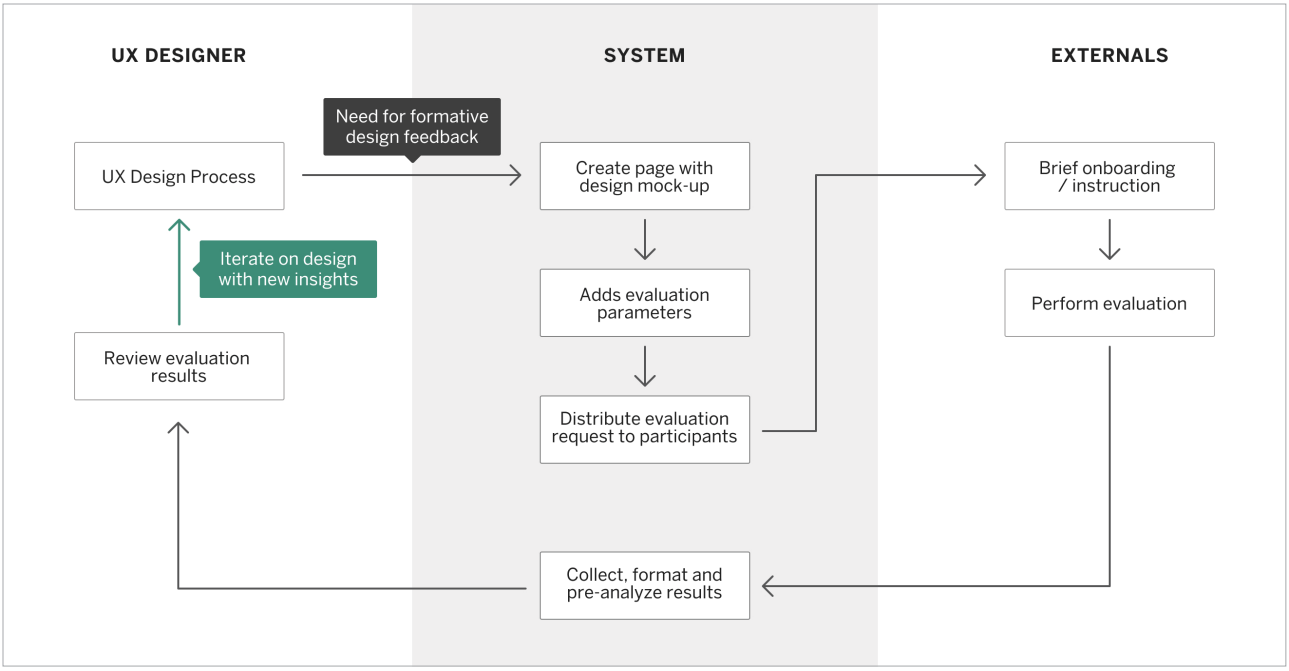


Figure 4: Schematic overview of the proposed design concept: a system that allows UX Designers to more easily iterate on their work by more inclusively engaging stakeholders in the process.

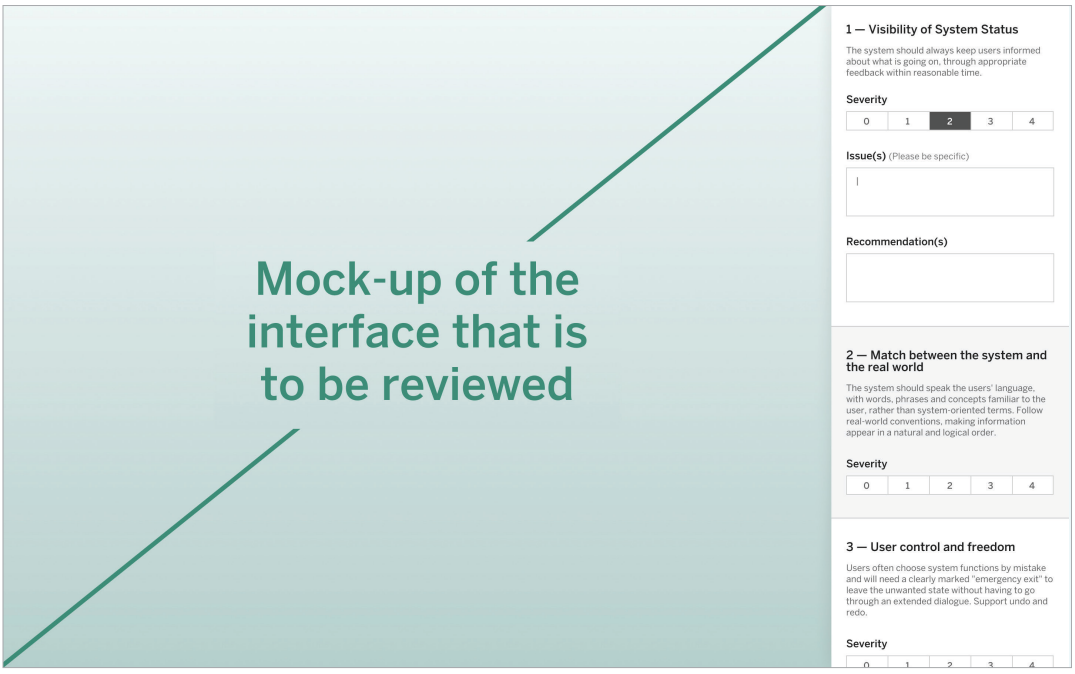


Figure 5: Wireframe representation of the proposed design concept, showcasing the prototype viewer (left) and the evaluation component (right).

SECTION TWO

Design Process

In the second section the process from design brief to a high fidelity demonstrator is discussed in detail — describing the potential solutions and discussing the rationale behind the subsequent design decisions.

2— Design Process

2.1 Continuing the Design Brief

The project proposal ended with the proposition of a design concept that allows UX Designers to more easily share their works-in-progress with non-expert colleagues and stakeholders. In the design brief, the design concept was presented in a typical UX deliverable format: a *wireframe*. Wireframing is a methodology used early in design processes where a system's basic structure is compiled graphically, while the refined visual design is disregarded for the time being.

The great benefit to using wireframes in early phases is the fact that it allows adaptations to be made more easily, since the resolution of the work is in lower fidelity. Wireframes are a great vehicle to communicate design concepts visually to stakeholders before a large investment of time is made. The logical next step in this design process was to gather feedback on the early concept by demonstrating it to others, and discussing it.

Early Concept Review

The deliverables were presented to a number of (former) master students in the faculty who had affinity with UX design, ranging from *one*- to *four* years of experience. The sessions had an unstructured format: the emphasis was on seeing the overall response to the concept. The students were given the freedom to interact with the prototype using the mouse.

A common question that was asked was *"Does Figma not already support the functionality of sharing works with others?"*. Figma is a collaborative design tool which supports the design of user interfaces that can be made interactive.

Figma does indeed already support sharing works-in-progress with others. Recently, Figma has made a toolkit available that allows third parties to create their own *'plug-ins'*, which add custom functionalities to the tool. However, virtually all of these third parties focus on facilitating user testing and are modeled around the business of offering user test participants. The focus for this project is

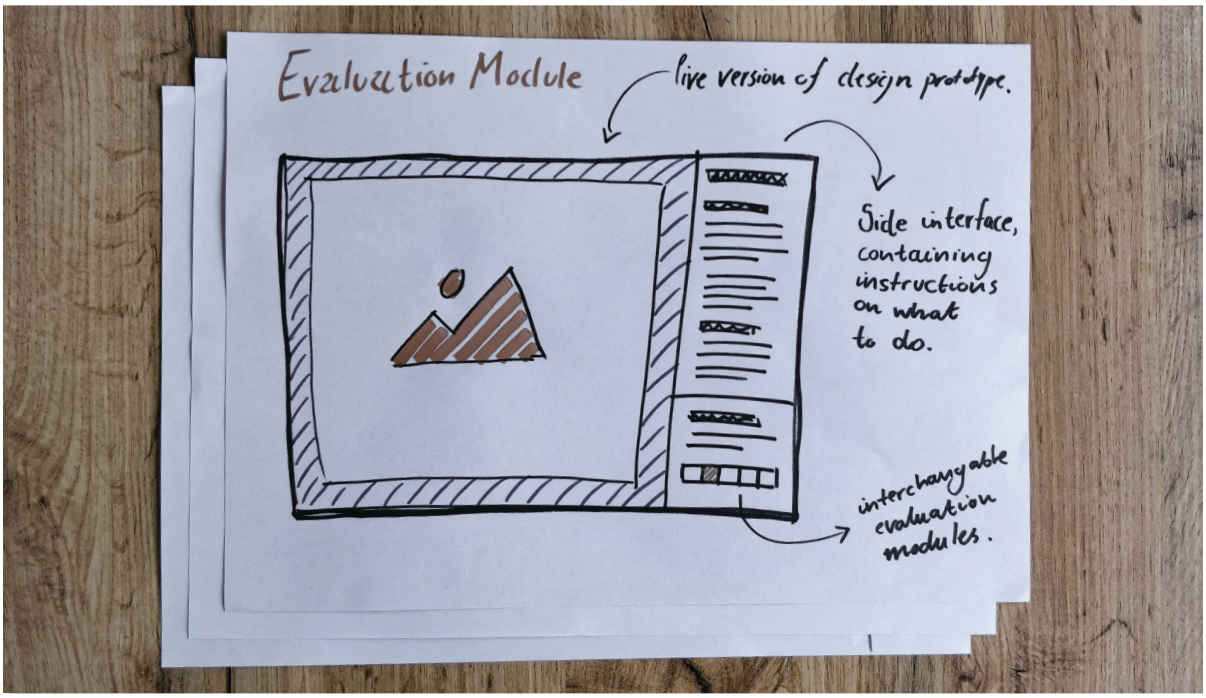


Figure 6: Early wireframe sketches, exploring proportions and layout of components.

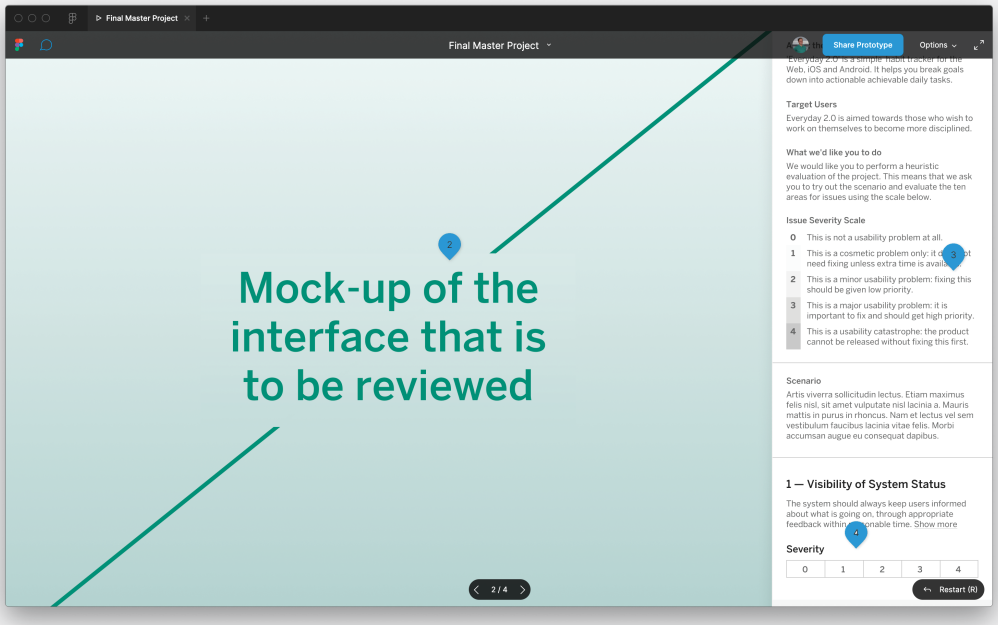


Figure 7: Using Figma's 'Share Prototype' feature to demonstrate the concept in a early reviews.

to help designers *connect with non-designers* in order to facilitate collaboration. At the time of the review, no such tool was available. Once this was clarified, the sentiments were positive:

> *"I definitely see the value of that. I often experience resistance from others when discussing important decisions [...] as they don't seem to grasp that design is more than what the end product looks like".*

Open discussions resulted in criteria that the students felt would 'make or break' the design concept.

> *"It's important for the concept to have a seamless integration with Figma: I don't want to spend a lot of time setting things up."*

Furthermore, the topic of *motivation* had a prominent place:

> *"Why would colleagues want to help me in evaluating the designs? If there are no willing users or if the quality of their feedback is poor, the service does not offer a lot of value."*

A suggestion offered to make the process of providing feedback more inviting was looking into the *Self-Determination Theory* (SDT) [17]. Using SDT it is possible to review components of an individual's needs (i.e. *autonomy*, *competence* and *relatedness*) in order to increase the willingness of the user base to help with such design evaluations.

Furthermore, tailoring the contents of the evaluation to the context and personality of the end-users was suggested. For example, there are opportunities to vary the tone of voice used when addressing them, or even by adapting the contents of the evaluation: this could gradually increase in complexity if the system learns that the evaluator has shown they can handle the complex matter.

Re-designing the Onboarding Experience

First impressions are crucial in UX Design — they have a significant subconscious effect on how credible, relevant and even useful a product is perceived [14]. For that reason the very first thing end-users of the tool was given an overhaul (Figure 8). In the original design, all information is displayed in a modal message which has two buttons — not very engaging.

The re-design sees the introduction of a number of elements that are based on the comments of the early review. Firstly, the message is transported through the metaphor of a postcard with an included letter. Most of its details are hidden, creating a compelling invitation to learn more.

Additionally, the notion of *relatedness* from SDT is used as a design material. Far more prominent in hierarchy than before, the first name of the person requesting the feedback is displayed (Figure 8). The rationale behind this is that in realistic contexts UX Designers have built up some degree of relatedness with their environment which may help motivate others to repay the favor to them.

Re-designing the Workflow

Revisiting the comment on seamlessly integrating the tool with Figma, the design decision was made to develop part of the tool inside Figma. Recently, Figma opened up their Application Programming Interface (API), which means that third parties can develop custom tools that may extend the functionality of Figma to their liking.

This provided unique opportunities for this project: rather than making a user operate multiple interfaces to complete the task of

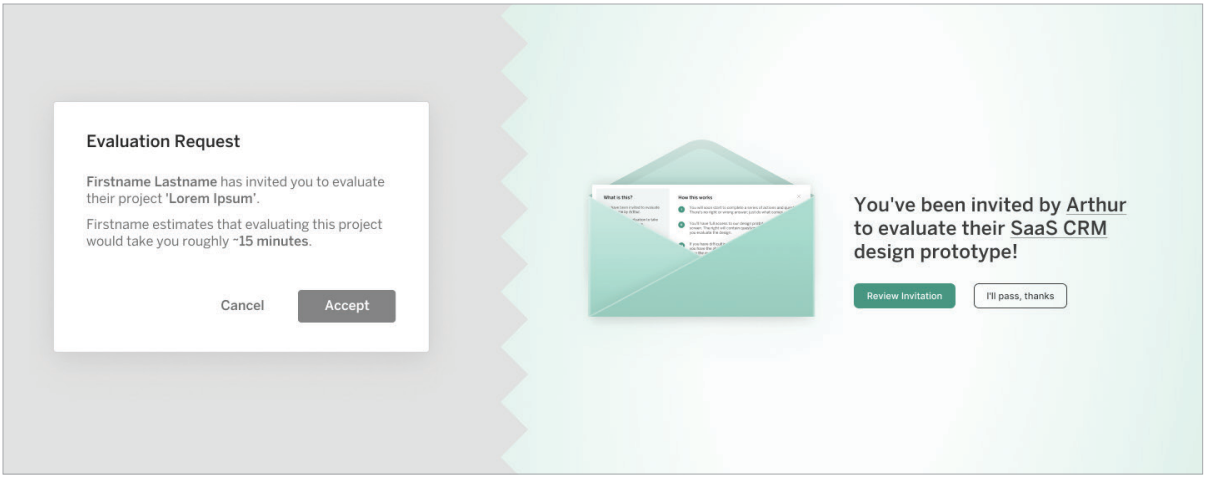


Figure 8: Before (left) and after (right) views of the evaluation request received when works-in-progress gets shared through the Requestor system.

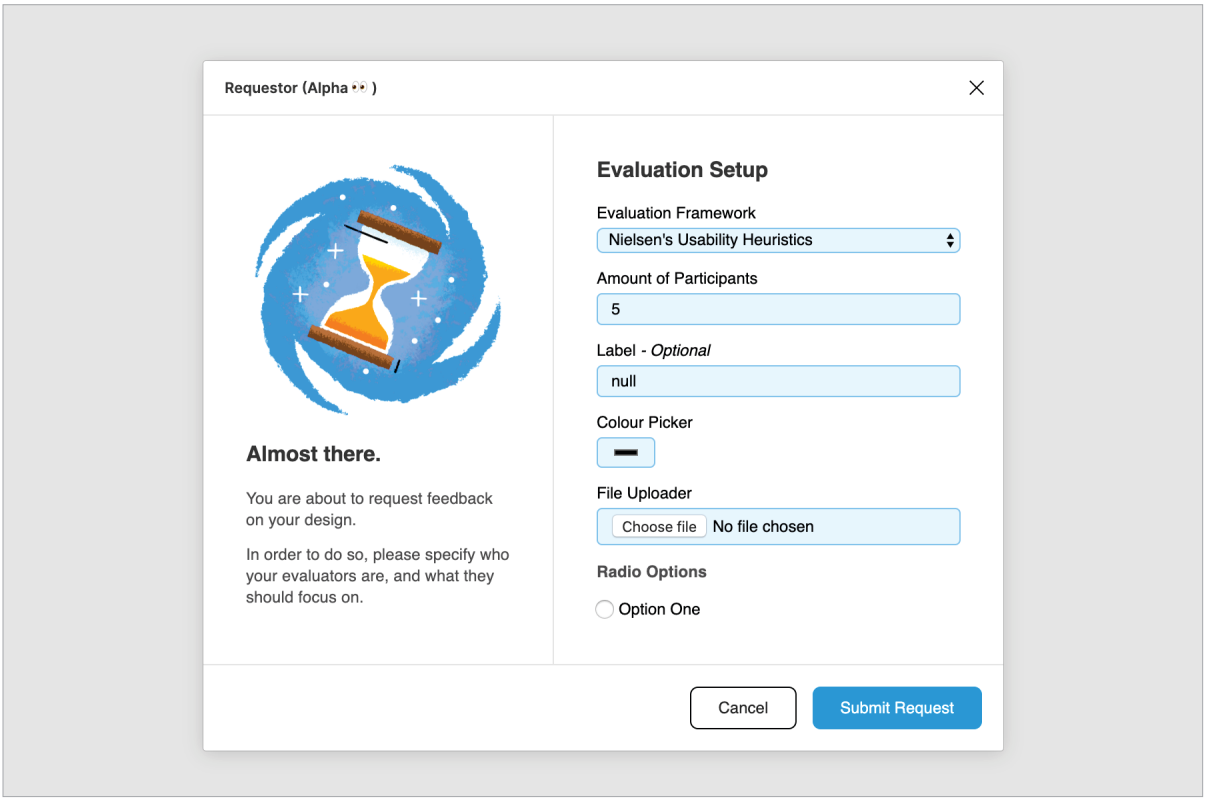


Figure 9: Early version of the Requestor Figma plugin. This plugin can be opened within the Figma application, creating a truly seamless experience for the UX Designers in sharing their work.

sharing their works, it could be localized within Figma, presumptively increasing the efficiency of the workflow. For that reason, I started developing a custom plugin in Figma.

Figma's API [16] is based on REST architecture, and handles requests made through HTTP endpoints. Furthermore, plugins made in Figma are written in TypeScript (a stricter variant of JavaScript), and have a graphical user interface that is created with HTML.

I created an early version of the Figma Plugin (Figure 9 on the previous page). This custom plugin can be opened within the Figma application within two clicks, offering a fast workflow. Additionally, the interface made use of 'smart defaults': reducing the cognitive load required to operate the interface. The plugin was developed to match the visual identity of Figma (typography, buttons) to not clash with the user experience resulting from the use of the 'native' Figma application.

At this point, the UX Designer could use the interface to manage the set up of their evaluation: allowing them to choose a framework which would guide the users in performing the evaluation. For the early prototypes, Nielsen's Usability Heuristics were used as the methodology included in the evaluation.

Further Validating the Problem Statement

Finally, perhaps the largest insight gained during the early review was that like in any design project, the concept I proposed is based on a number of assumptions and unknowns. For that reason, I decided to take a step back from designing, and assess the validity of the originally posed problem statement. I created a questionnaire on *design maturity* (Figure 10, Appendix G), with the intention to better understand its impact on designers working

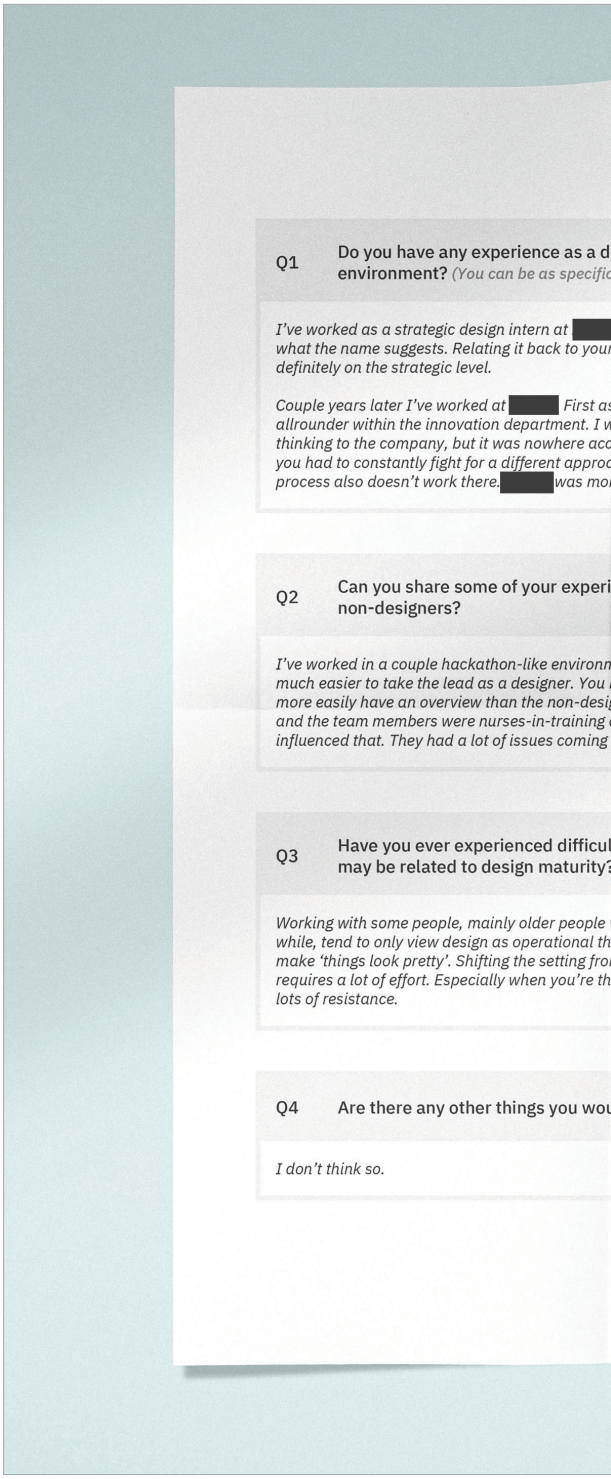
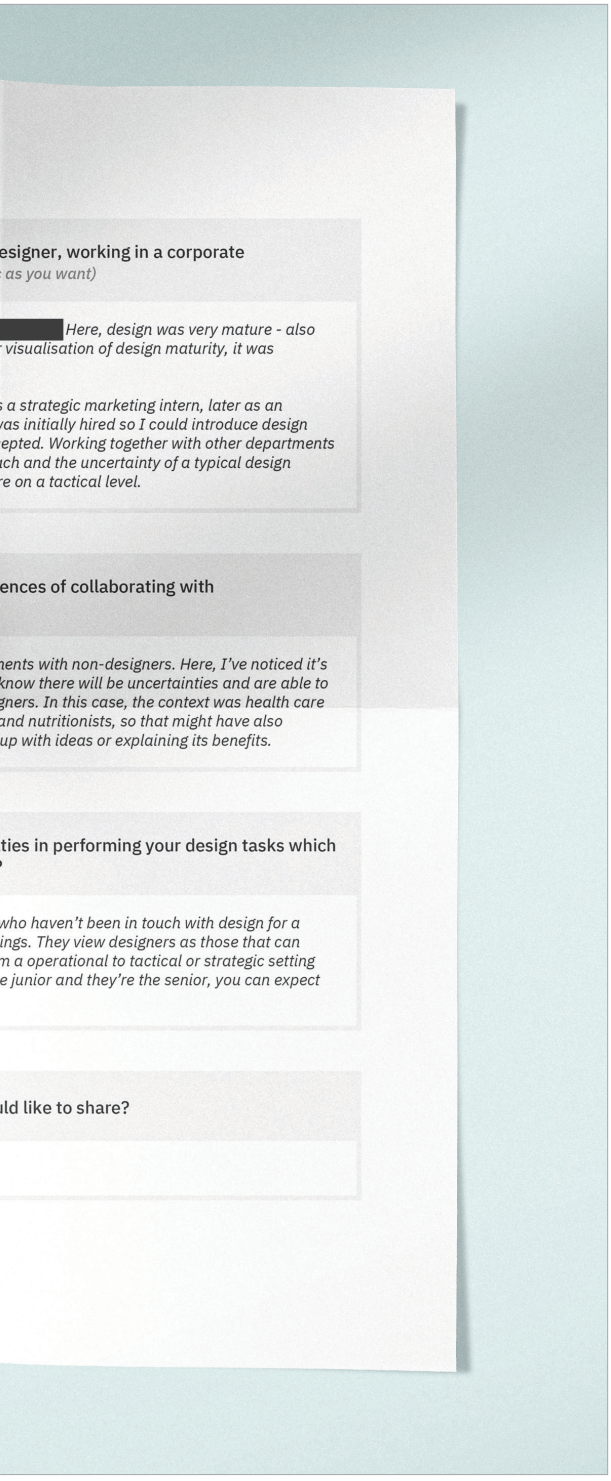


Figure 10: One of the respondents' answers to the
the described companies have been edited out to



e questionnaire on design maturity. The names of
respect their privacy.

in corporate environments. The questionnaire was distributed to (former) design students of the faculty who have experience working in corporate environments.

The results coincide with the sentiments gathered from the work presented in Section One: designers reported that not all stakeholders had an appropriate view of the characteristics and responsibilities seen in design.

> *"... some people [...] tend to only view design as operational things. They view designers as those that can make 'things look pretty'."*

This is in line with the findings from earlier presented work: designer's responsibilities are trivialized, and making strategic decisions using arguments from the design angle is met with resistance.

This resistance most commonly occurs in environments where design is newly introduced. These insights further justify the design direction: the best way to develop empathy towards others is by making their presence more apparent.

2.2 First Fully Functional Prototype

Increasing Visual Fidelity

Having explored the design space through an early review of the concept and problem statement, I resumed the design- and development towards a functional proof-of-concept.

Within UX Design, the visual fidelity of designed products are gradually built up as the design space is explored. For this project, that meant that sketches of wireframes were digitized, and those digital wireframes were refined to meet the visual design principles of

scale, *visual hierarchy*, *balance* and *contrast*. To reduce cognitive load, icons were prepended to headers on the page to provide visual structure and easier *scannability* (Figure 11).

The visual identity of the platform should lend itself to that of the content that was to be evaluated. Not only would this give the embedded works-in-progress a higher importance, it would also extend the brand experience of the work to create a familiar user experience. For this reason, a dynamic colour scheme was introduced, derived from the design work that was embedded in the platform (Figure 12).

The technology to realise such a feature is publicly available, as the JavaScript library vibrant.js extracts prominent colours from images by quantizing RGB histograms [59]. However, when dynamically using colours, it is important to keep the contrast in mind, mainly for web accessibility. According to the Web Content Accessibility Guidelines (WCAG), large text should have a contrast of at least 3:1, while normal text should have a contrast of at least 4.5:1 [57]. Dynamically generated colours should be checked for contrast.

Finally, this iteration saw *microinteractions* being included. Microinteractions are small moments of feedback on user actions that disclose the status of a system, communicate additional information or add to the brand experience. In this specific case I saw the opportunity to include microinteractions to reduce cognitive load. Rather than having to remember all specific error severities, the system now displays them on-demand (Figure 13).

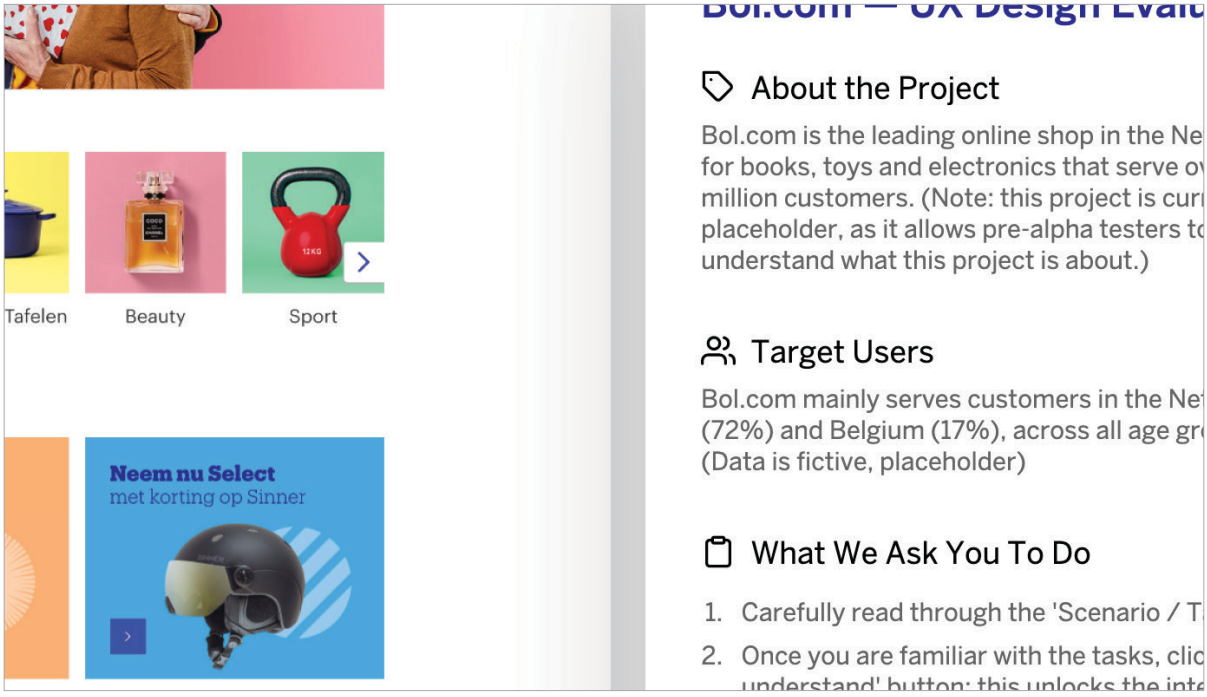


Figure 11: Iconography being used to increase the scannability of content, and to improve the visual aesthetic by sticking to a coherent design language used throughout all branding.

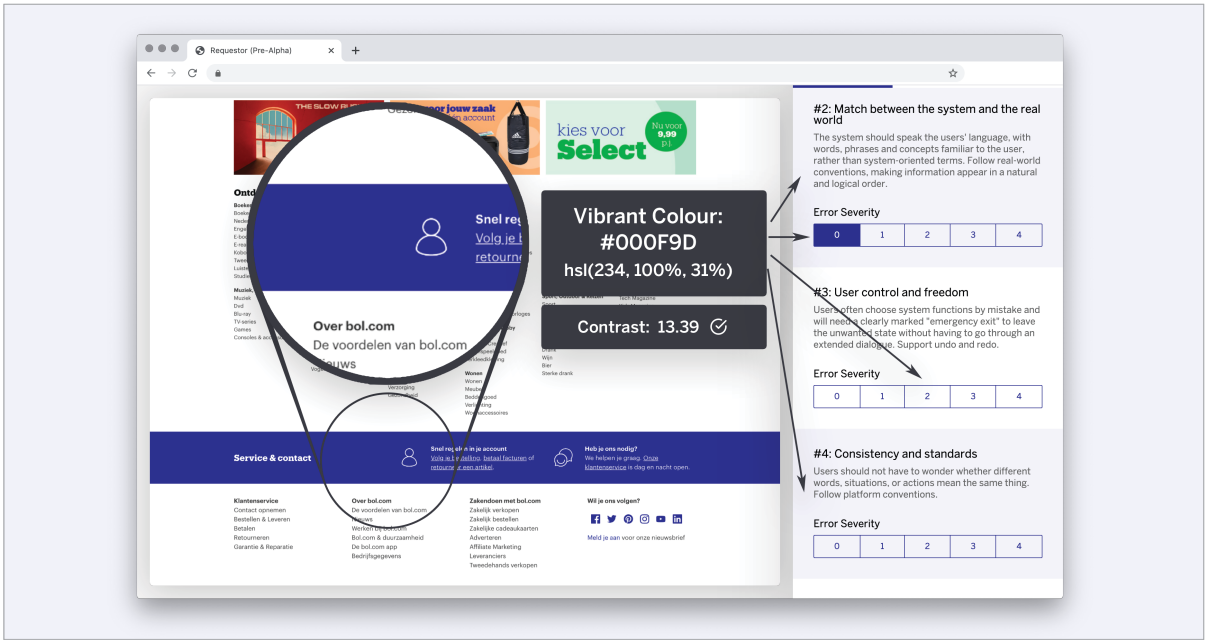


Figure 12: Proof-of-concept of automatic detection of vibrant colours.

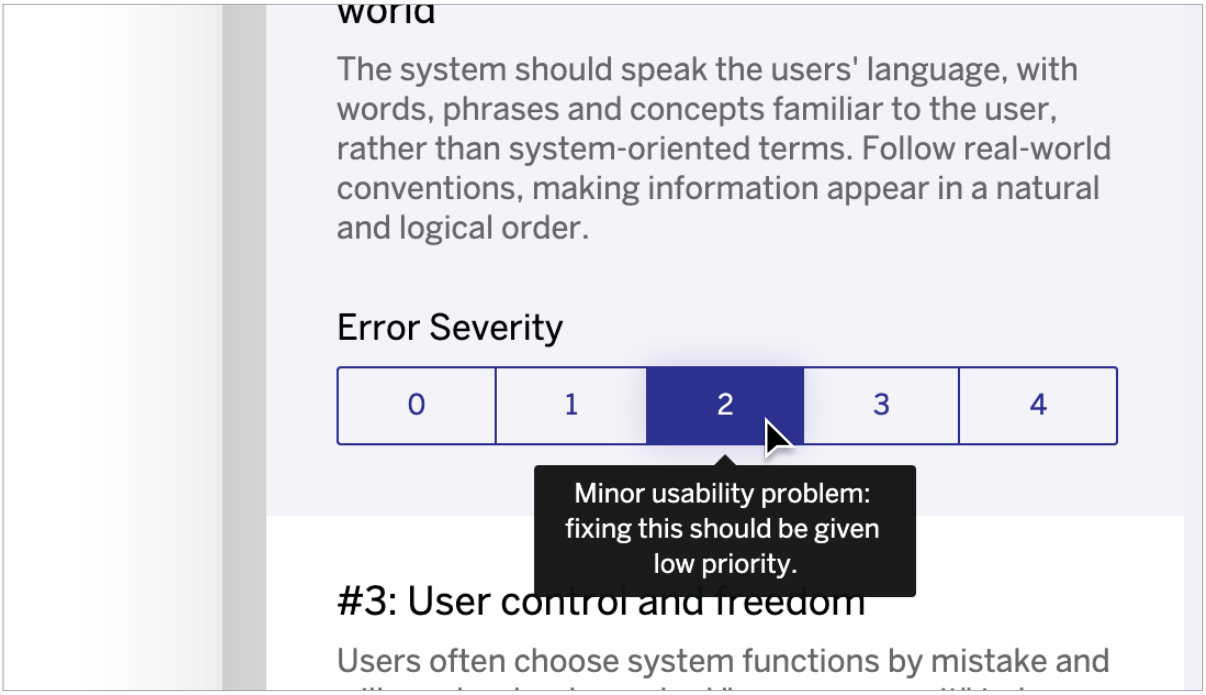


Figure 13: An example of a microinteraction. Upon hovering over the error severity rating, the user is presented with the context-specific information required to make a good judgement.

Technical Realization
The platform is built as a web application using HTML, CSS, JavaScript on the front-end. Web applications can be accessed directly through an internet browser. Considering the design context (limited resources available, differences in skill-sets and knowledge) it made the most sense to make the experience of using the platform as frictionless as possible. No installation is required to access the platform, saving precious time for all involved. This also meant that later in the project, sharing it with testers would be a lot easier.

The back-end of this platform was realized using a LAMP stack: Linux as operating system, Apache as web server, MySQL as the database and PHP as the programming language which is connected to the front-end (Figure 14).

User Flow Overview
Graphical user interfaces cannot be communicated well over text. Instead, to provide an impression of the first fully functional prototype, a user flow is shown in Figure 15.

Through these user flows, the process of evaluating a design work-in-progress is shown. The work-in-progress is imported from Figma, and has full interactive capabilities — giving the evaluators the feeling that the project is in a far stage of development.

It is important to note that the screens of the interactive prototype depicted in the user flow are made of scraped screenshots of a Dutch e-commerce website, which I do not claim authorship of. Rather, the content was used to give the testers a relatable case to evaluate.

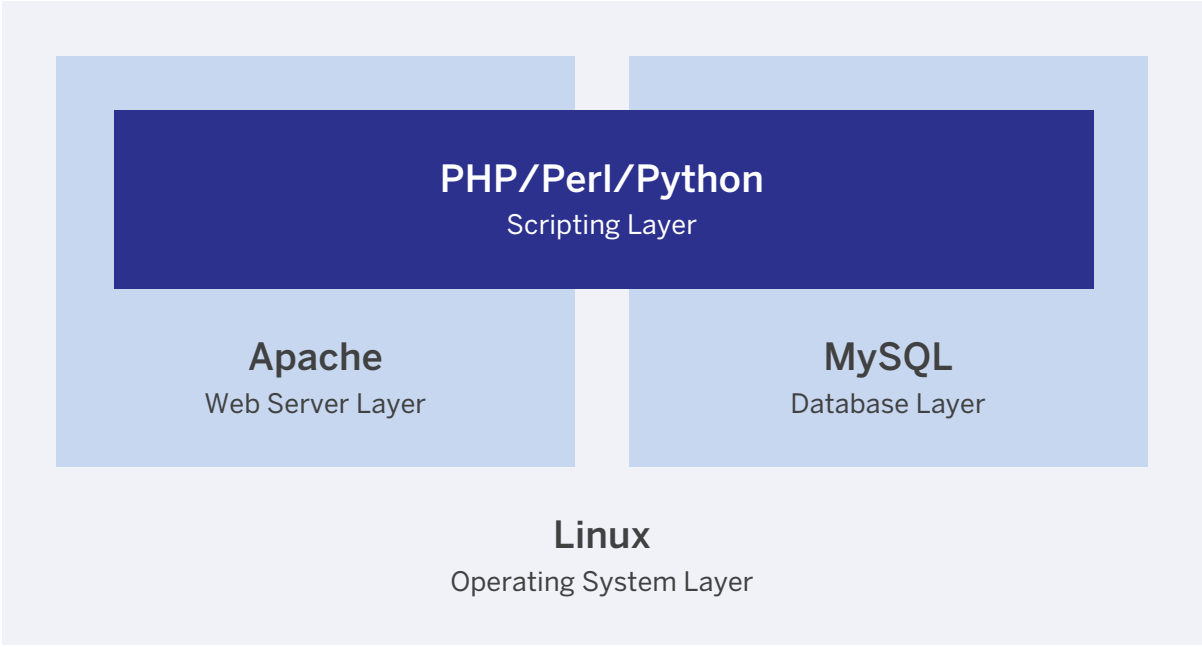
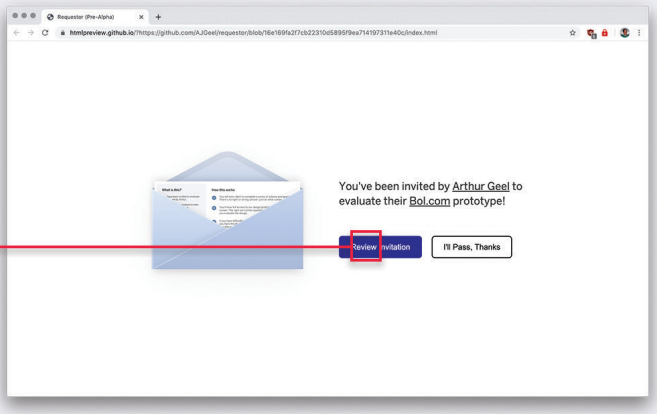
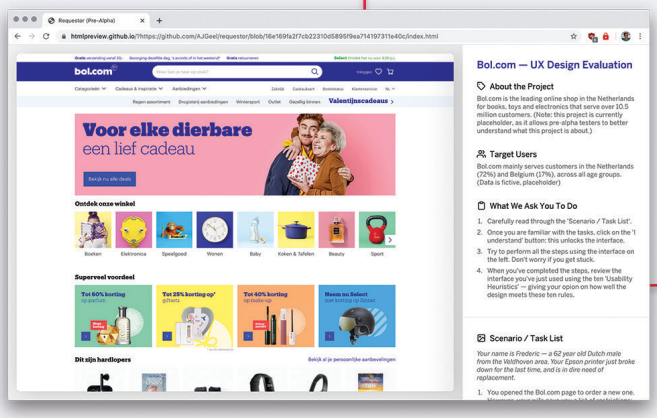


Figure 14: Visual schematic of the back-end technology stack used in this project.

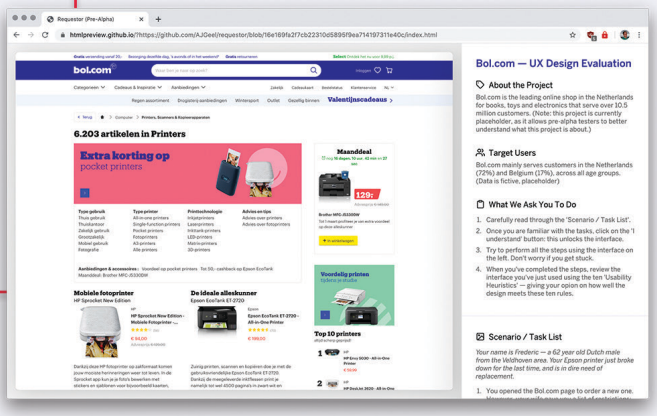
1. The user enters the website, and is greeted with an invitation that explains the context.



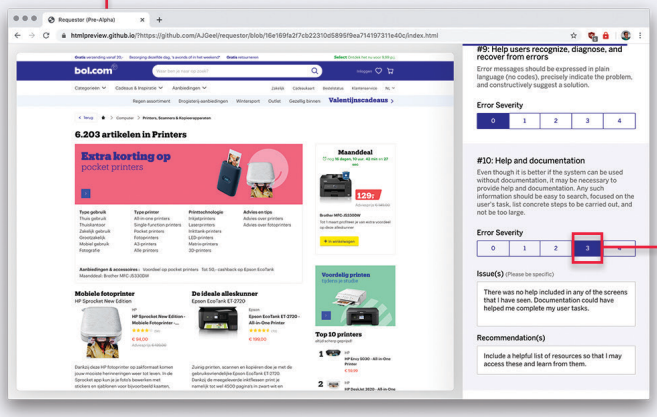
2. The user accepts the invitation, and is shown an interactive prototype alongside instructions for the evaluation.



3. The user performs a specific scenario using the interactive prototype on the left.



4. After completing the scenario, the user is able answers questions on the usability of the design.



5. The user has completed the full evaluation and ends the interaction by submitting it.

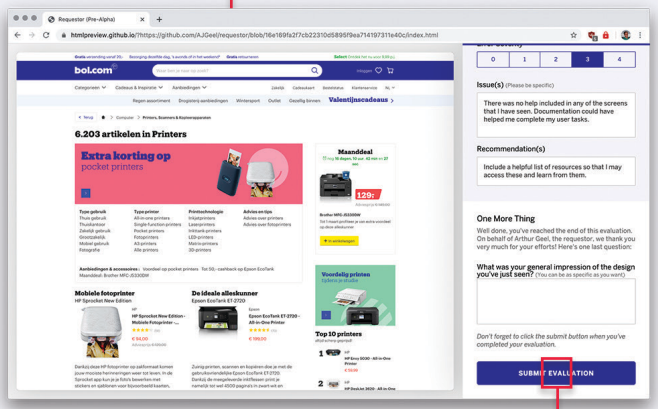


Figure 15: User flow of the first fully functional prototype.

2.3 Diverging Once More

Usability Testing
The first fully functional prototype presented in the previous pages was evaluated in a small-scale usability test. The goals for this usability test were to assure the technical implementation worked soundly, to diagnose potential problems in the system, as well as discover opportunities for improvement.

The rationale for only conducting a small-scale usability test is that experts agree that the best results come from testing no more than five users [37]. Rather than conducting very thorough usability tests in the active development of a project, experts advocate saving time and resources by running as many small tests as possible.

For this usability test four end-users were recruited. Because at this point in time I had not completed the Ethical Review Form yet while the design process required intermediate evaluation of the concept before further development, the decision was made to once more recruit master students from the Industrial Design faculty to minimize the potential unethical impact of this test.

Participants of this usability test were asked to use the user interface with the goal of completing a heuristic evaluation of the embedded e-commerce website. In total, the tests lasted no more than fifteen minutes each.

Overview of the Results
All four participants were successful in conducting the heuristic evaluation. However, that was to be expected as the user group had above average design expertise. The focus for this usability test was mainly to validate that everything worked as it should.

A comment was made on the copywriting used for Nielsen's Usability Heuristics:

> *"Take a look at the descriptive pieces of text below the questions as it's quite ambiguous and a-personal how it comes across."*

The textual content provided to aid the evaluators in performing the heuristic evaluation was directly cited from the original, academic work to retain its authenticity. However, adjusting the copy to make it more accessible may be sensible for this context.

Furthermore, a number of technical issues were uncovered. While the application broadly functioned in the way it was programmed to (*i.e. serving- and saving the appropriate content to- and from the database)*, a number of minor bugs resulted in a worse user experience.

For example, an issue was found which related to overflow on the horizontal axis, caused by erroneous CSS. Additionally, some of the internal processes of the application caused an unexpected CPU usage.

One participant remarked feeling overwhelmed at the start of the interaction. This comment was understandable: the interface did indeed display a lot of content at once when the invitation was opened, and the embedded prototype loading from Figma's servers caused some flickering on the screen.

A comment was made on the applicability of the tool: currently only available for prototypes that were designed in Figma. The participant felt left out, as they exclusively used Adobe Xd as a tool to create design prototypes, which currently was not supported.

Finally, a participant remarked the fact that they found the interface *'somewhat bland'* —

the interactions were straightforward, but felt very formal and goal-oriented. In conclusion, the usability test resulted in constructive new insights that should be reflected upon.

Trend Evaluation

Prompted by the last comment, I started critically looking at the interaction style: perhaps the interface which was designed to be efficient in use, was not the optimal solution for the design challenge.

To refresh my perspective, I hosted a small co-creation session with two fellow master students in which we collected examples of web interactions which we felt had unique qualities.

I analysed the collected cases for themes, which resulted in a number of opportunities to improve the interaction of this project's concept (a more detailed elaboration can be found in Appendix H).

By far the most eye catching trend of the highlighted cases was the strong presence of *brand identities*. The combination of visual design, tone-of-voice and motion design created a enjoyable user experience that made visiting the websites interesting.

Furthermore, the highlighted cases permitted themselves flexibility in their interaction styles. Rather than displaying all content textually, the cases had developed interesting ways to progressively disclose the content, by emphasising a conversational narrative.

The evaluation of case studies resulted in a number of guidelines to be used for improving the design concept.

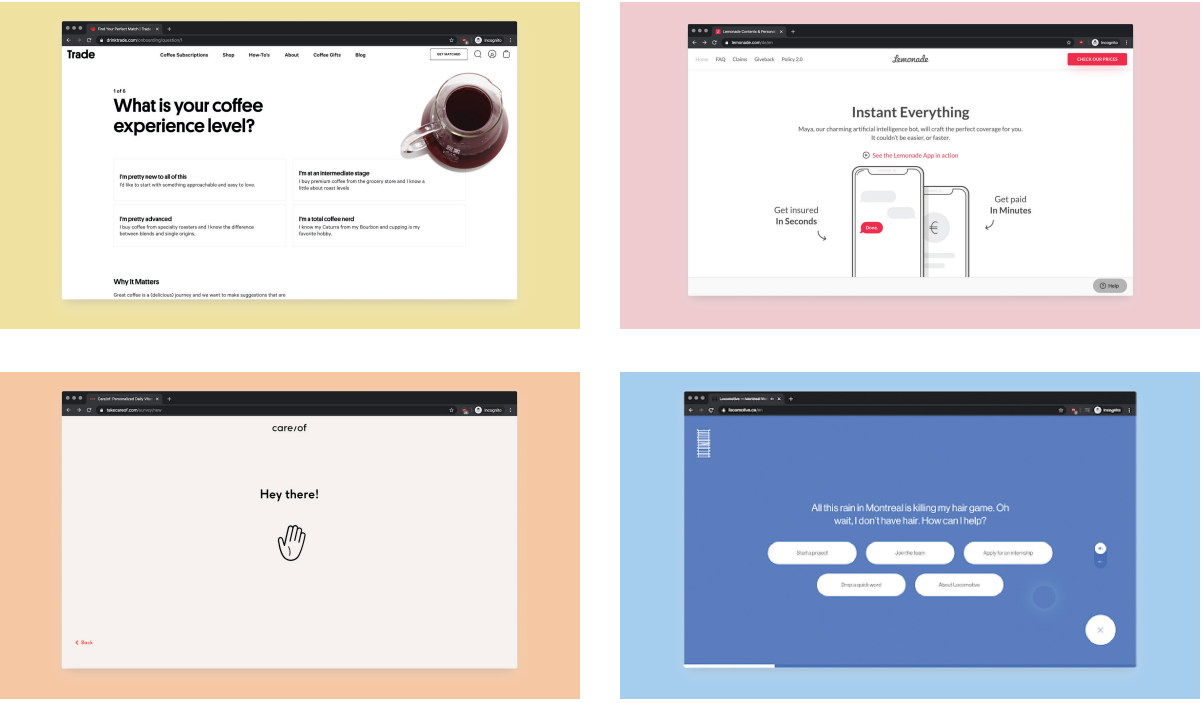


Figure 16: A number of engaging web interfaces resulting from a co-creation session. Source, in clockwise order from top left: Trade [51], Lemonade [29], Locomotive [30] and Care/Of [7].

2.4 Converging Again

Creating a Brand

In order to create a more compelling narrative in support of the design concept, a brand identity was defined. The platform was named *'Requestor'*, referring to the system taking care of requesting feedback from stakeholders.

In order to move towards meeting the design requirements, the platform was positioned to contain the intersection of values of *'reliable'*, *'friendly'* and *'easy-going'*. The colour scheme, typography and tone-of-voice used throughout the concept's assets were designed with these values in mind (Figure 17).

For the type face, the 'Sen' font was used (Figure 18), giving typographic elements a more rounded, friendly feel. The 'Sen' font is licensed under the Open Font License, making it suitable for commercial purposes. Its inclusion in the Google Fonts API ensures that it can reliably be served in digital channels.

Finally, a logo mark was created to suit the brand identity (Figure 19). A paper plane was introduced as a metaphor for the act of playfully repurposing materials found in a traditionally serious space: the office environment. With this, I tried to position the Requestor system as a like-minded 'friendly' and 'fun' activity in an otherwise serious environment.

The paper plane logo further introduced opportunities for expanding the brand design, as it is very suited for motion design. For example, the paper plane can be 'thrown' as a means of sending out the requests, also adding a feeling of delight in this interaction.

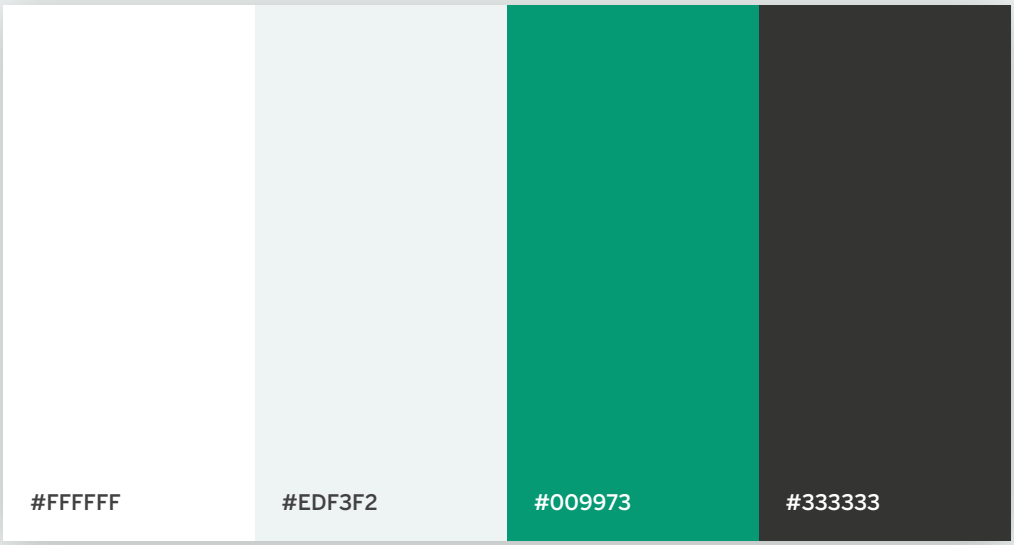


Figure 17: An overview of colours used in the Requestor brand.

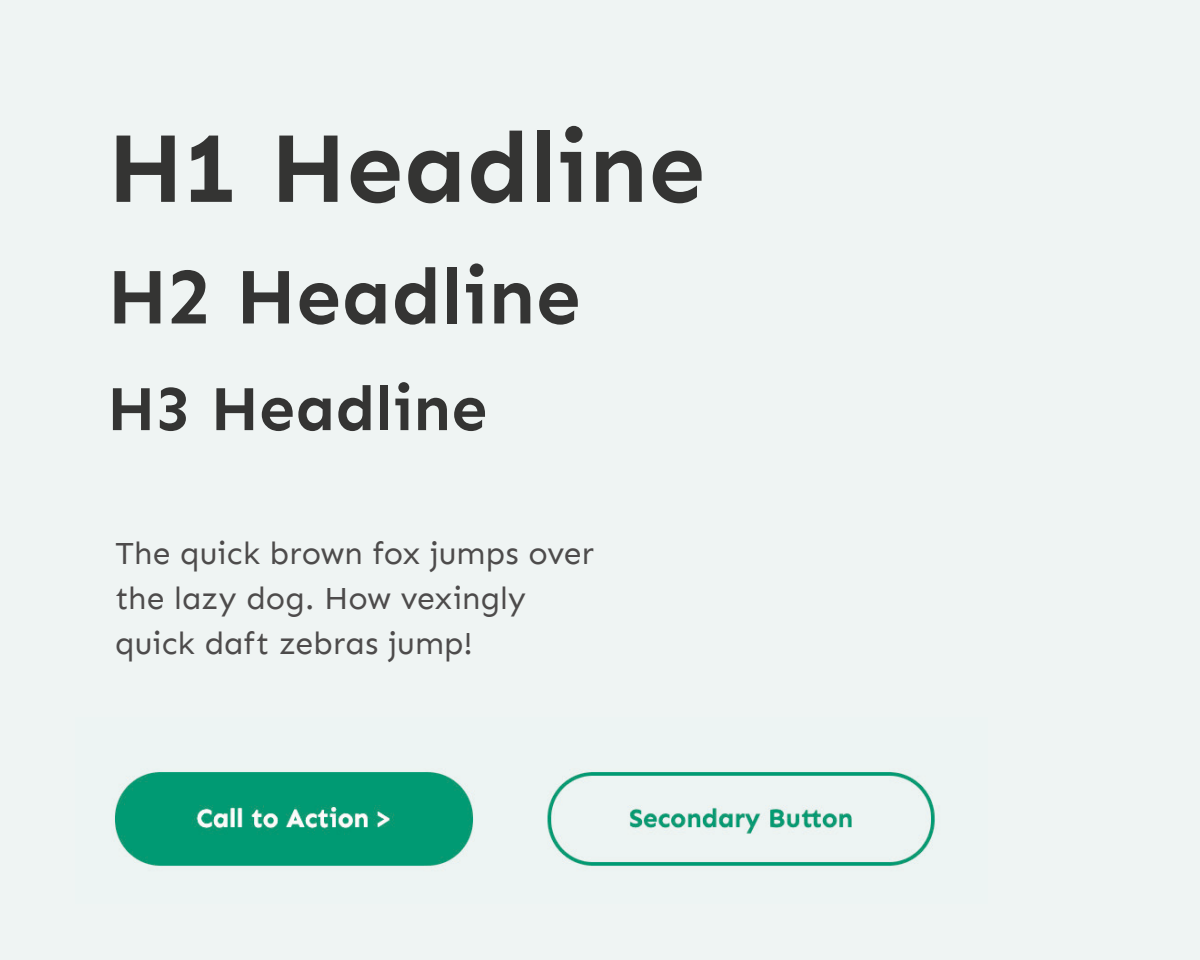


Figure 18: An overview of typographic elements used in the Requestor brand.

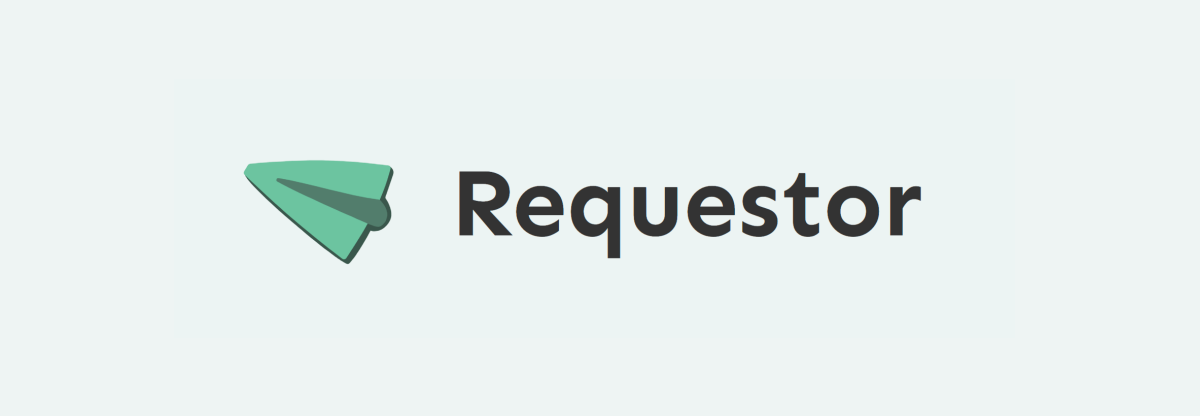


Figure 19: The paper plane logo mark for the Requestor brand.

Expanding the Utility

As previously mentioned, the platform is built as a web application and can be accessed in any browser using the internet. However, importing design projects for use in evaluations happened in Figma, using a custom built plugin. This supplied an iFrame on the page seen by evaluators with the contents of the protoype, which are interactive.

The way this platform is built allows for more flexibility beyond supporting designs built in Figma. The HTML5 iFrame technology used can support a variety of content which even includes other web pages. As such, it would be a shame if only the Figma prototyping tool was supported, since UX Designers also use different tools and the environment is bound to change in the future.

For this reason, this iteration saw the way the platform handles the content input change. To understand the technical possibilities, firstly we need to review the landscape. Design prototyping tools increasingly are cloud-based, which makes sharing works with others easier.

Recently, a lot of prototyping tools have been releasing functionality upgrades that allow interactive prototypes to be shared through an URL in browsers. Commonly, this functionality also includes the option to *embed* this prototype in online documents and web pages.

This development is huge for the Requestor platform: it means that integrating prototyping tools other than Figma is technically feasible. Therefore, the interaction pattern of formulating evaluation requests from the Figma plugin was scrapped, and replaced by a dashboard system which accepts URLs (Figure 20).

This change in inner workings greatly expands the utility and thus marketability of the Requestor platform. Beyond prototypes created in Figma, the platform now supports widely used tools such as *Adobe Xd*, *Invision*, *UXPin* and *Axure*. Finally, any website can also be embedded.

From an interaction perspective, this approach makes a lot more sense. Rather than introducing uncertainty as to where existing Requestor projects can be viewed and managed, they all reside in a centralized place. This also made the process of creating new evaluation requests easier and more intuitive. Taking into account the insights gathered from the trend review in the previous pages, the process was re-designed to better include the design pattern of progressive disclosure. The new interaction uses a sequence of full-screen modals (Figure 21), reducing the cognitive load experienced.

Exploring Interaction Modalities

The final activity conducted in the final iteration was reviewing the way non-experts experienced the process of conducting the design evaluation. In earlier concept prototypes, a typical interaction was designed which guided the users through the process of evaluating by using statically placed text and input fields.

However, the results from the usability test conducted in the previous iteration suggest that there is room for improvement in this regard. Firstly, an incremental adjustment was made by addressing the issue related to feeling overwhelmed by the excess influx of information at the start of the interaction.

A modal overlay was introduced (Figure 22 on the next page) which restored visual hierarchy in the interface. The modal contained a to-the-point instruction, guiding the interaction.

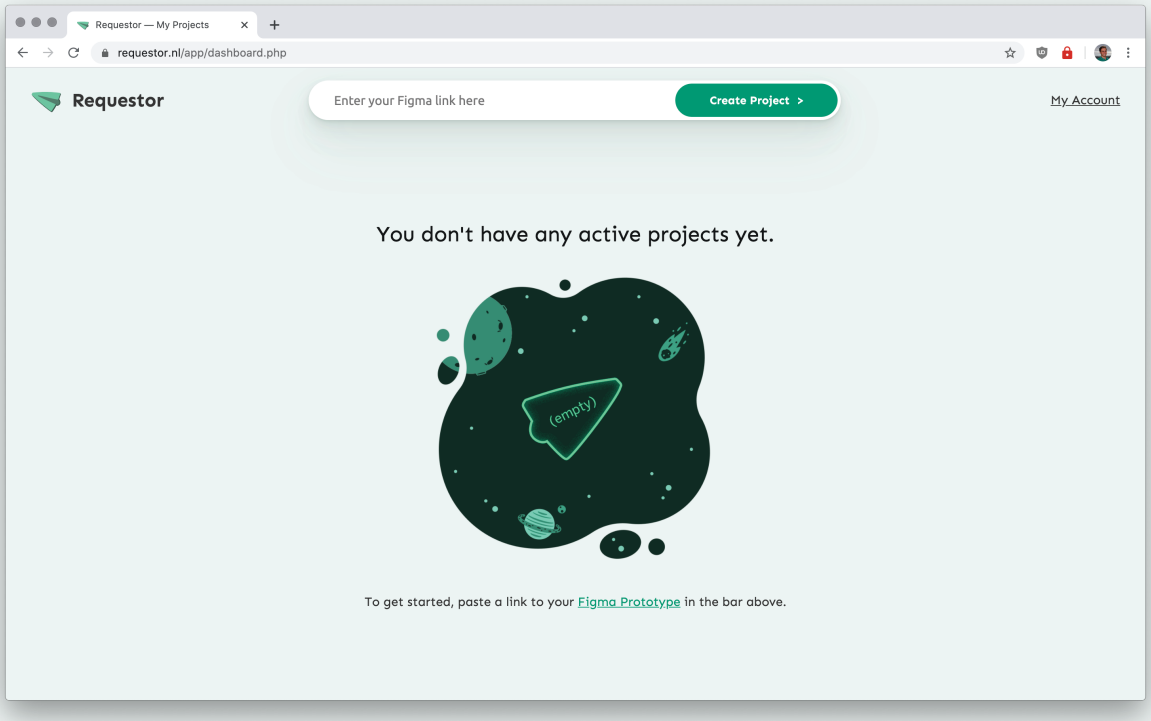


Figure 20: The Requestor Dashboard: a central place where projects can be created, shared and managed. The input bar at the top is used to import projects from external tools.

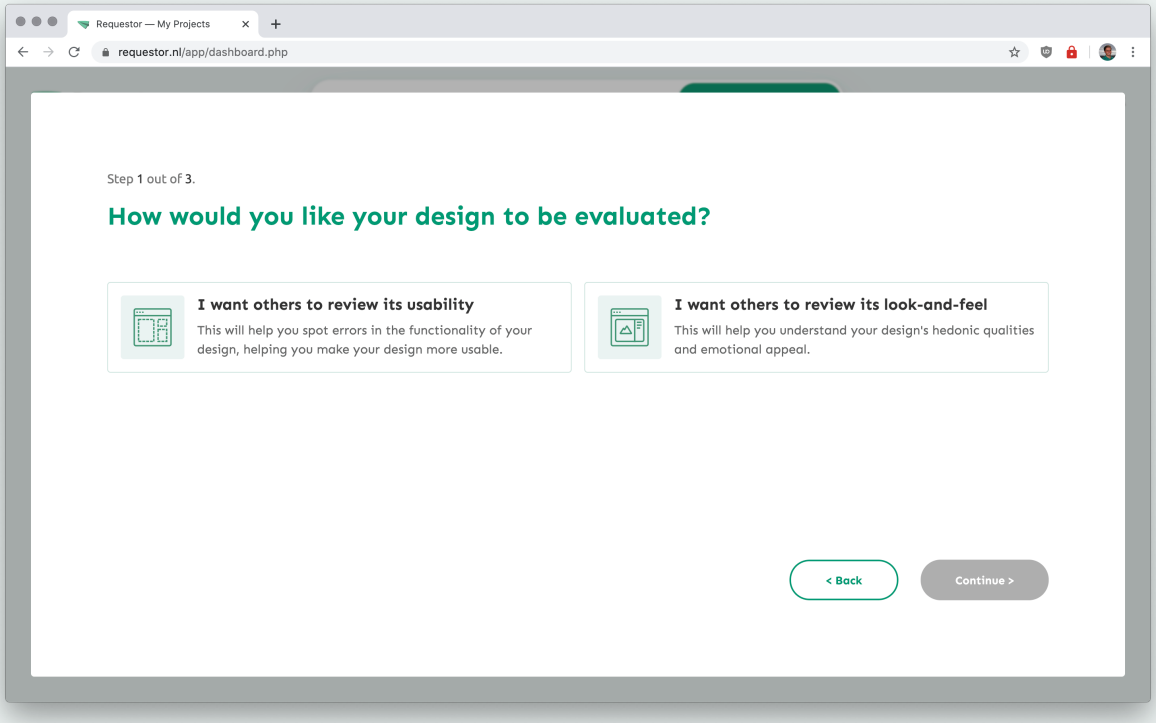


Figure 21: Stage one of the evaluation request creator, specifying the scope of the evaluation.

A subtle animation further helped in directing the attention towards the most important information. Finally, the modal overlay slightly blurred the embedded prototype in the background, while also reducing its opacity. These implementations also helped reduce the flickering nature of content loading in the iFrame.

The trend evaluation conducted earlier resulted in a number of ideas on how the interaction might be improved in order to better fulfill the design requirements described in Section One. Two options stood out: *Natural Language Forms* and *Conversational User Interfaces*.

Natural Language Forms

Natural Language Forms (NLF) are web forms that deviate from the standard label-input field convention of laying out forms. Instead, the form is inserted in a sentence, often written from the reader's perspective (Figure 23). The reader is subtly invited to provide their input.

NLFs have not received mainstream attention within UX Design, although a handful of studies show promising results. In [58], Wroblewski performed an A/B test on the conversion rates of a contact form (Figure 23). The results favour the NLF, increasing conversion rates by 25 to 40%. An experiment conducted with a pet insurance website [32] (n=9523) saw slightly better conversion rates by formulating their call-to-action button using a NLF (42.3%) when compared to a traditional variant (41.8%). However, a study by McKenzie [33] saw different results. In an A/B test (n=736), the conversion rate of the NLF (21.7%) was lower than the traditional variant (27.6%).

Furthermore, the key drawback of a NLF is complexity. Simple forms translate well into a NLF, yet forms with more than 10 items quickly start feeling strenuous. Therefore, a NLF would not be a good fit for this project.

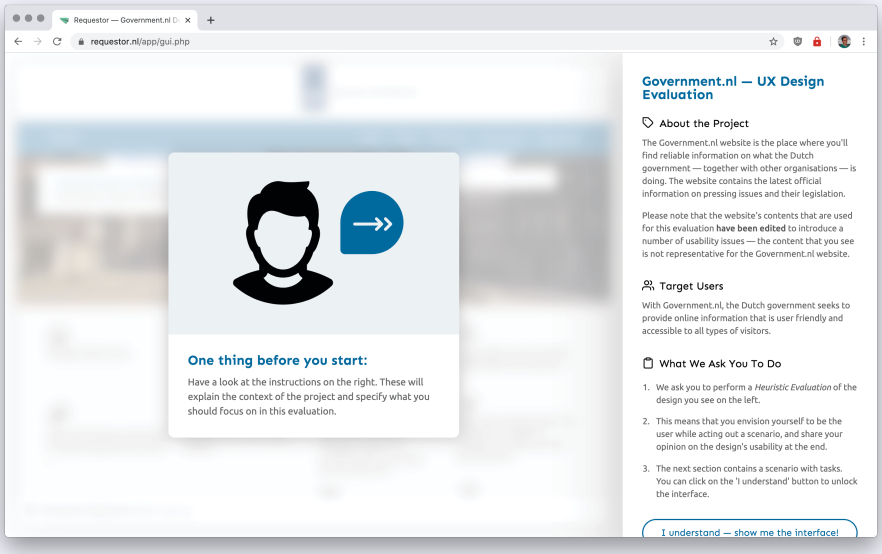


Figure 22: The modal overlay (left), directing the user's attention to the task at hand.

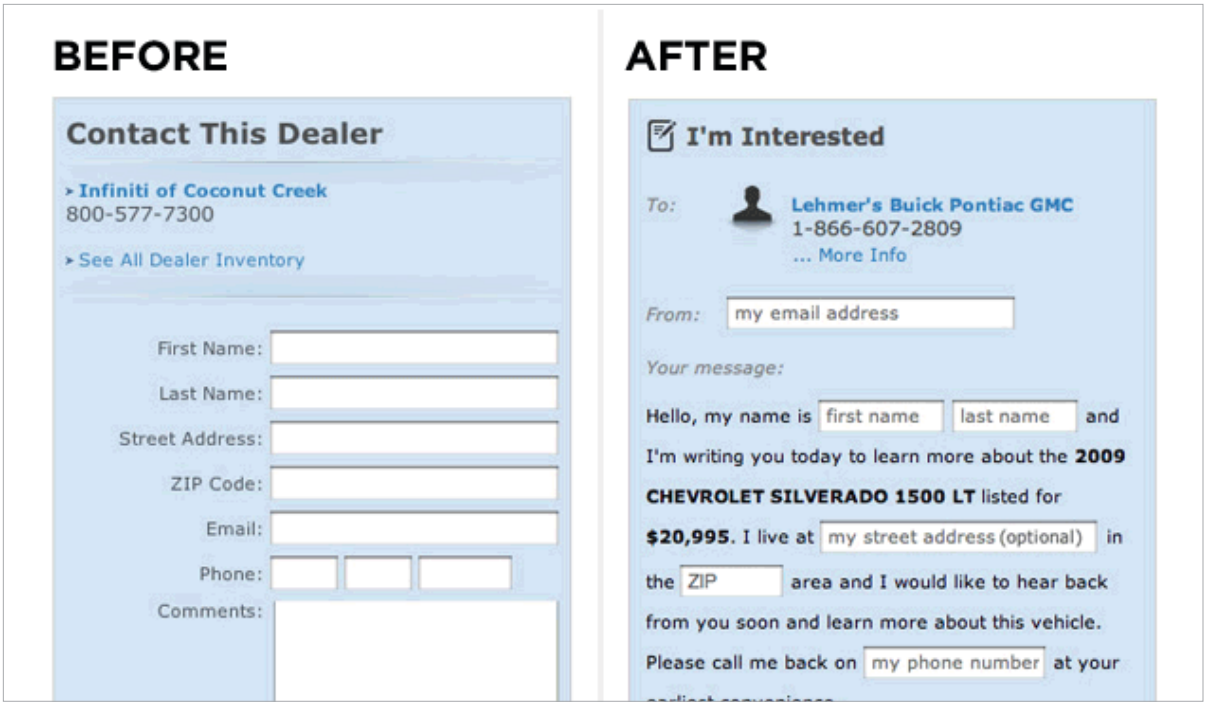


Figure 23: Comparing a traditional form to a NLF. Source: Luke Wrobolewski [58].

Figure 24: An example of a Natural Language Form. Source: Oscar Insurance [40].

Conversational User Interfaces
Conversational User Interfaces (CUIs) are unlike regular Graphical User Interfaces (GUIs) as they allow the users to interact by using principles of real-life communication. Instead of requiring the user to examine graphical elements to understand the system status, a CUI allows the use of speech or writing, as a conversation is held [34].

Often, *Natural Language Processing* (NLP) is used to allow the system to interpret the human's input, although simpler CUIs may not need NLP to function.

CUIs can have many different implementations. Voice Assistants [35] are user interfaces that enable users to perform tasks in smart systems by speaking a command, including instructing speakers to play music, querying specific questions and more.

Chatbots are another common implementation of a CUI [1]. A chatbot can be defined as a web-based interface that enables users to collect- and store information. Chatbots are often integrated in back-end systems of existing services, allowing them to access and update content in these. Chatbots are commonly featured in customer service,
removing the need for human operators.

When applied in the context of the Requestor platform, a CUI could have numerous benefits over the previous, 'traditional' user interface discussed in earlier sections. For one, interacting with a CUI generally is more engaging, as there is a more consistent exchange in communication between the system and the user. Gradually disclosing the information required to perform the evaluation using the system can reduce the experienced cognitive load.

Furthermore, allowing the evaluators to interact with the interface using a language that is more in tune with the way they would naturally communicate may increase their motivation to participate in activities like these.

Out of the two proposed CUIs the Chatbot implementation was chosen because of the better fit in this project. The modular approach to developing the platform made it feasible to replace existing components (Figure 25) with a similar component.

In this new variant of the Requestor platform (Figure 26), the graphical elements are replaced by a conversation with an agent that is based on the chat-bubble.js library [50]. The agent follows a predefined script which has conditional branches based on the user's input.

In similar fashion to the 'traditional' user interface, this CUI guides the evaluators through the process of conducting a design evaluation, in this case using Nielsen's Heuristics.

2.5 Final Design — Requestor

Concept
Requestor is an online platform that was designed to help UX Designers collaborate with stakeholders in their environment. Interactive design prototypes created in common UX software can easily be imported.

Projects created in Requestor can easily be shared with others to ask them for feedback. However, providing meaningful feedback is challenging without any design expertise. For that reason, Requestor offers guidance that enables anyone to give meaningful feedback by introducing commonly used methods.

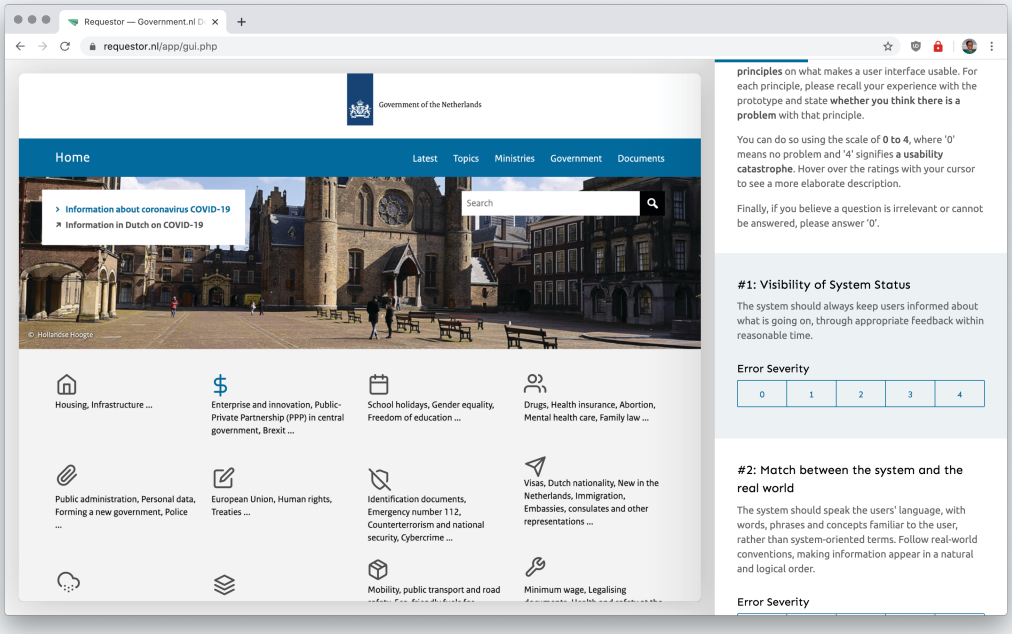


Figure 25: The 'traditional' variant of the Requestor user interface. This variant of the demonstrator prototype can be accessed via https://requestor.nl/app/gui.php

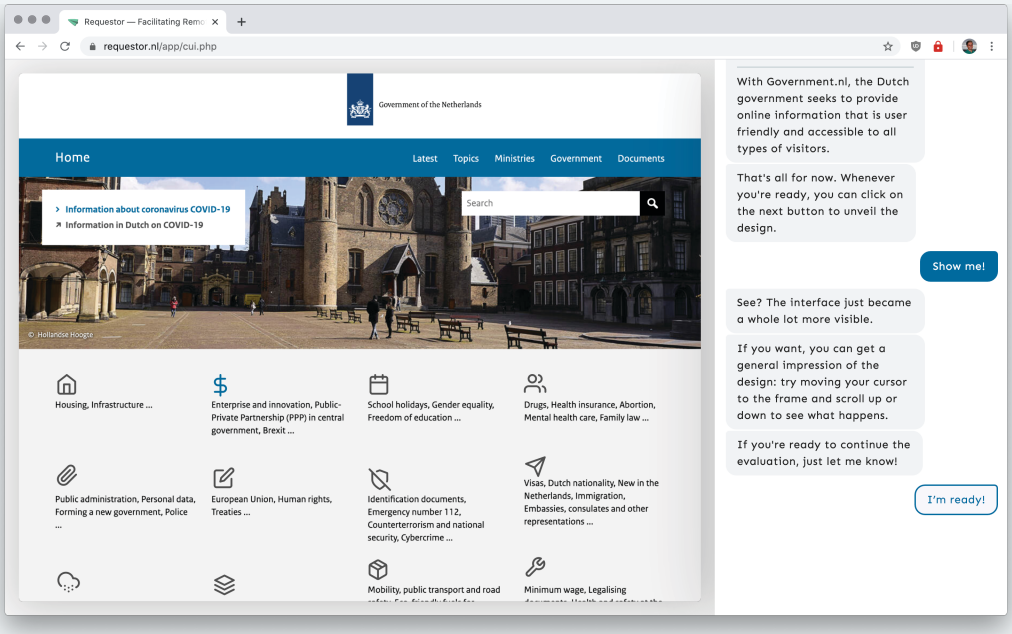


Figure 26: The 'conversational' variant of the Requestor user interface. This variant of the demonstrator prototype can be accessed via https://requestor.nl/app/cui.php

Design and Interaction
The user interface for conducting design evaluations in Requestor is available in two variants: as a *'traditional'* user interface and as a *'conversational'* user interface. These interfaces can be accessed in any modern web browser, and provide interactive guidance in performing evaluations of design projects.

On the left side of the interface the design project that is evaluated can be seen. This component is highly interactive: allowing the users to click, scroll and use other gestures to navigate the design, providing them with a realistic experience of how the eventual project will be.

The right side of the interface contains actionable instructions towards conducting the design evaluation, and informs the users about the project's context and focal group of users. The instructions also contain a use-case scenario that can be carried out using the interactive design on the left. Finally, the instructions contain interactive fields that can be filled in by the user, guiding them through the evaluation This module contains two variants: the *'traditional'* and the *'conversational'* user interface.

Technology
Requestor is a web application that uses a front-end of HTML, CSS and JS that handles user interactions. The application was developed with *responsive web design* in mind, allowing users to resize their browser without degrading the user experience due to layout issues. In order to assure that the embedded prototype displays properly, a notice is displayed whenever the window is less than 840 pixels wide, asking them to increase the size.

The back-end of the system is built using a LAMP stack (Linux, Apache, PHP and MySQL). This part of the system takes care of serving the correct content to its visitors. Additionally, the back-end securely handles the retrieval and modification of user data. In its design, measures were taken to guarantee security, such as the implementation of SSL encryption, and the prevention of SQL Injection.

Final Demo Day
The results of this Final Master Project were presented during the Final Demo Day, a biannual event where students of the Industrial Design faculty present their semester's work.

Due to the circumstances related to the novel Coronavirus pandemic of 2020, the Final Demo Day was held online, on a custom built platform: https://demoday.id.tue.nl/.

In order to present this project in satisfactory standard, I created a number of deliverables specifically for this event. I invested a substantial amount of time in creating a short promotional video as I felt this would be the most compelling way to tell the story in this context The promotional video (Figure 27) explains the design challenge and presents the Requestor platform.

Additionally, I designed and developed a *landing page website*, aimed at explaining and promoting the Requestor platform (Figure 28).

An overview of all Final Demo Day deliverables can be found in Appendix E.

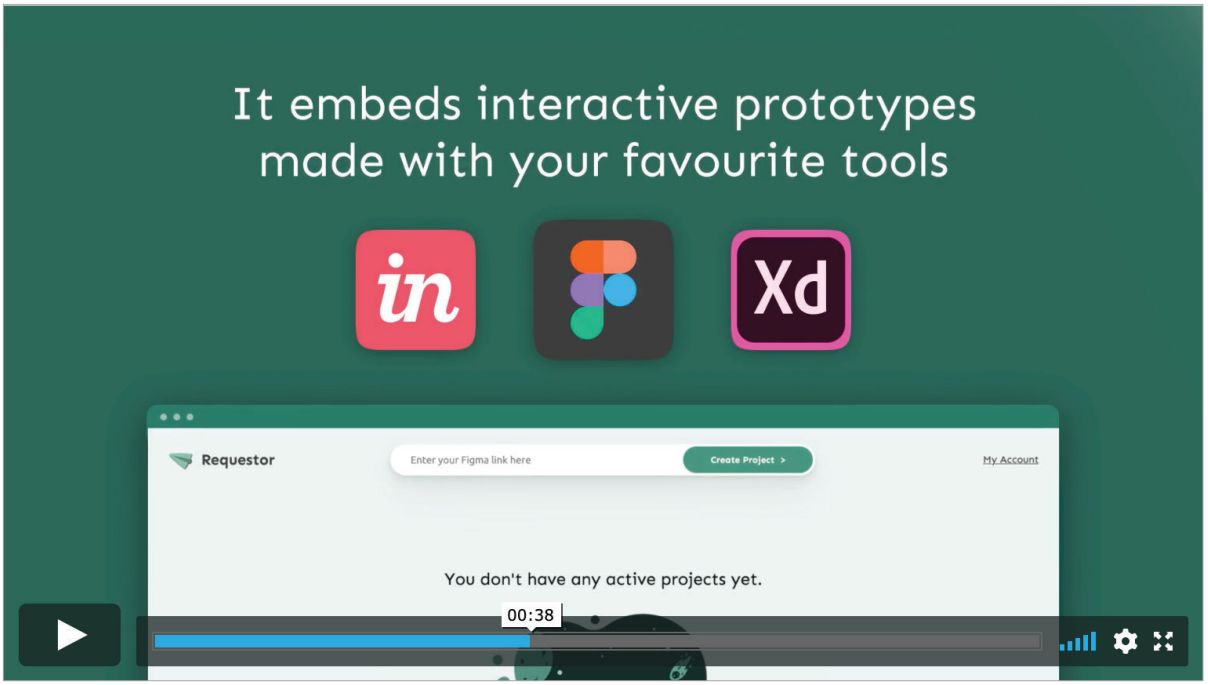


Figure 27: A promotional video created to explain the design challenge and resulting concept: Requestor. The video (01:30) can be accessed via https://vimeo.com/425870546

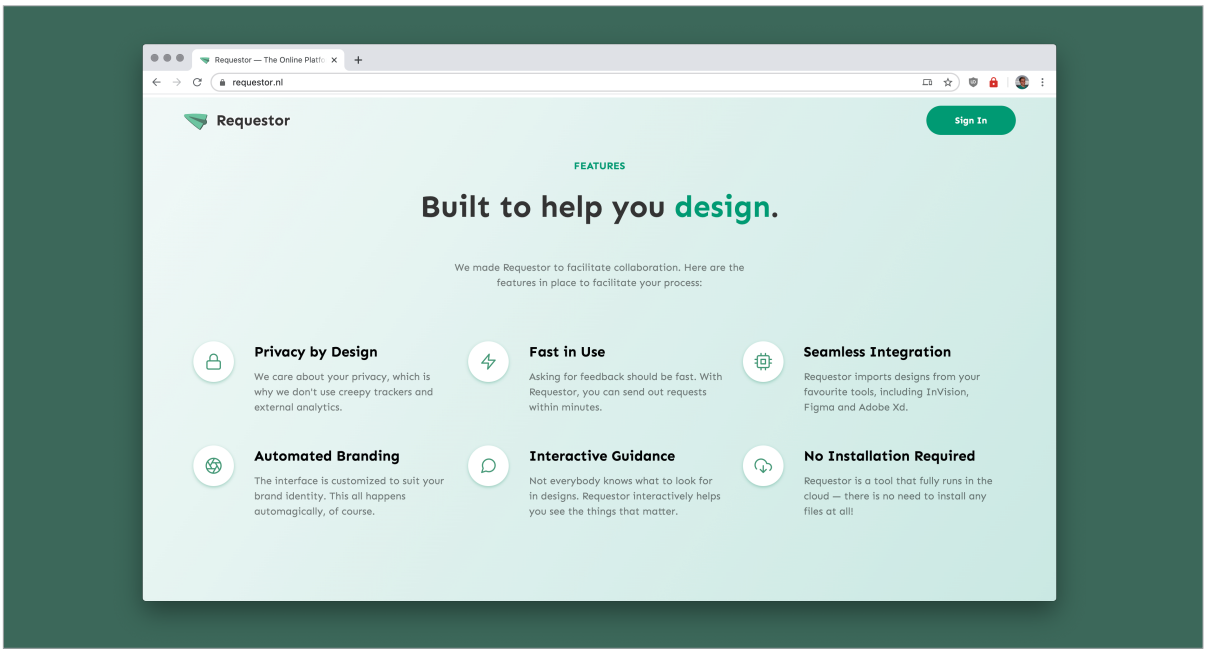


Figure 28: The Requestor landing page, accessible via https://requestor.nl/

SECTION THREE

Design Evaluation

In the third section the aims for- and process of setting up a summative design evaluation of the Requestor concept is described.

3— Design Evaluation

3.1 Study Setup

The aims for this study are bifold: firstly, I seek to evaluate the viability of the Requestor platform as a collaborative tool for design. Secondly, I seek to explore the effects that User Interfaces (UIs) with different styles of interactions have on a person's ability to perform an evaluation of a design prototype.

In order to progress towards these aims, a study was designed which features a realistic design case that is to be evaluated by participants recruited for this study. The participants are supported in evaluating the design case with the interactive guidance of the *Requestor* tool. Each participant was randomly assigned to either the 'Traditional' variant or the 'Conversational' variant of the tool.

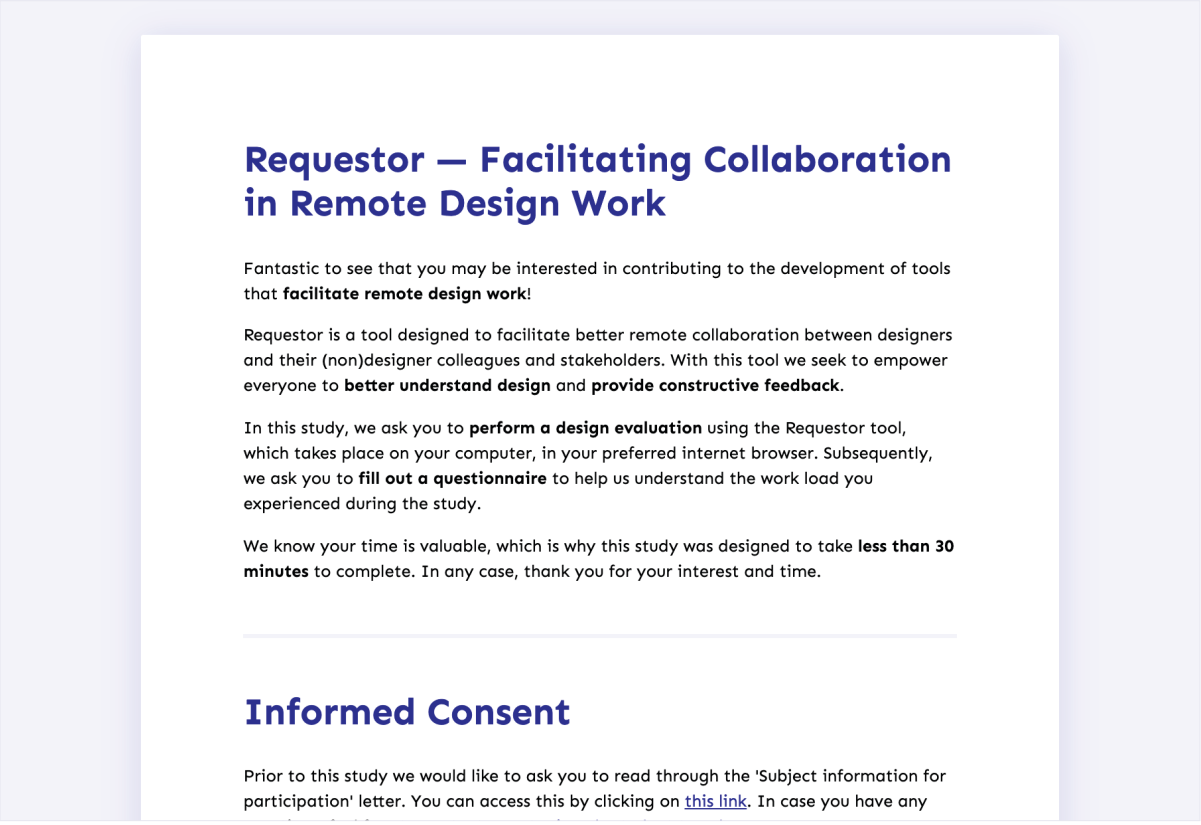
The inclusion criteria for this participating in this study were tuned to fit in with the project framing — focussing mainly on those with little to some expertise in UX Design.

In this experiment, the functionalities of the Requestor platform as decribed in Section Two were extended to keep track of the performance of participants such as the *completion of tasks* and the *time spent* doing so.

After the participants had completed evaluating the design case using the Requestor interface, they were asked to complete a questionnaire on their perceived task load. On average, participants spent 25 to 30 minutes performing the tasks in this study.

To summarise, this study consisted participants taking part in an unmoderated session that consisted of two activities:

1. *Conducting heuristic evaluation of a fictional design case, using one out of two variants of the Requestor prototype.*
2. *Completing a post-task questionnaire to evaluate the work load (NASA-TLX) they perceived for the prototype variant used.*



Requestor — Facilitating Collaboration in Remote Design Work

Fantastic to see that you may be interested in contributing to the development of tools that **facilitate remote design work**!

Requestor is a tool designed to facilitate better remote collaboration between designers and their (non)designer colleagues and stakeholders. With this tool we seek to empower everyone to **better understand design** and **provide constructive feedback**.

In this study, we ask you to **perform a design evaluation** using the Requestor tool, which takes place on your computer, in your preferred internet browser. Subsequently, we ask you to **fill out a questionnaire** to help us understand the work load you experienced during the study.

We know your time is valuable, which is why this study was designed to take **less than 30 minutes** to complete. In any case, thank you for your interest and time.

Informed Consent

Prior to this study we would like to ask you to read through the 'Subject information for participation' letter. You can access this by clicking on this link. In case you have any

Figure 29: An overview of the informational page of the study. This document and the related 'Subject information for participation' letter may be found in Appendix A.

3.2 Fictional Design Case

The design case that was selected for this study was a fictional case involving the Dutch Government and their website https://government.nl. In this fictional case, the participants were asked to perform a *Heuristic Evaluation* of the website's usability.

In these tumultuous times of a pandemic, this website is important to find reliable official information on pressing issues and their legislation, which made it an interesting case for this project. Additionally, the case used throughout development of the platform (pages 23-26) was based on a Dutch e-commerce company, which might have been problematic due to reasons related to copyright.

As part of the design case, the participants were asked to empathise with a fictional end-user, and carry out a number of tasks from their perspective. The scenario was:

> *Your name is Frederic: a 44 year old expat living in the Eindhoven area. As the owner of a business without employees (ZZP), the uncertain times caused by the Corona Crisis have put a strain on your financial situation.*

Moreover, the following was communicated:

1. *You recall that the prime minister discussed options for financial support in a press conference, some time ago. However, you forgot what it was called, and whether it is applicable for you.*

2. *Your task is to **use the government.nl website** to find reliable information on the financial support: you want to know the measure's name and the amount of financial support you can get.*

For this scenario, screenshots of the original Government.nl website were scraped and imported into the Figma prototyping tool. An expert review of the website found that there were little to no usability issues in this website. This is no surprise, seeing as the Government's goal for the website is to provide online information that is user-friendly and accessible to all types of visitors.

Injected Usability Issues
For that reason, the screenshots were modified to introduce a number of usability issues, examples can be seen in Figures 30 and 31.

The injected issues were problematic in six of Nielsen's heuristics:

1. *Visibility of system status;*
2. *Match between the system and the real world;*
3. *User control and freedom;*
4. *Consistency and standards;*
5. *Error prevention;*
9. *Help user recognise, diagnose and recover from errors.*

3.3 Prototype Variants

The participants were asked to use the custom built Requestor tool to help them perform this evaluation. Essentially:

- *It displays an interactive demonstrator of the design project. Users of the tool can use their mouse to scroll and click to browse the project as if it were a fully operational website.*
- *It contains an interactive module that is used to provide instructions for the evaluators. This interactive module guides the evaluators through the process of conducting a design evaluation, and contains questions the evaluators can answer to share their input.*

Participants were programmatcially assigned to a variant of the Requestor tool upon visiting the page. Roughly half of the participants were assigned the *'traditional'* user interface and half of the participants were assigned the *'conversational'* user interface.

3.4 Post-Test Questionnaire

After participants completed the design evaluation by submitting their results, they were provided feedback on their performance through an intermediary screen (Figure 32 on the next page). This was done to provide some mental rest in between the parts of the study, as it could be experienced as demanding.

In this screen, the participants received a message of appreciation for their efforts so far, and given some insights as to how they were doing (time spent, amount of usability issues found). Furthermore, The screen contained an optional

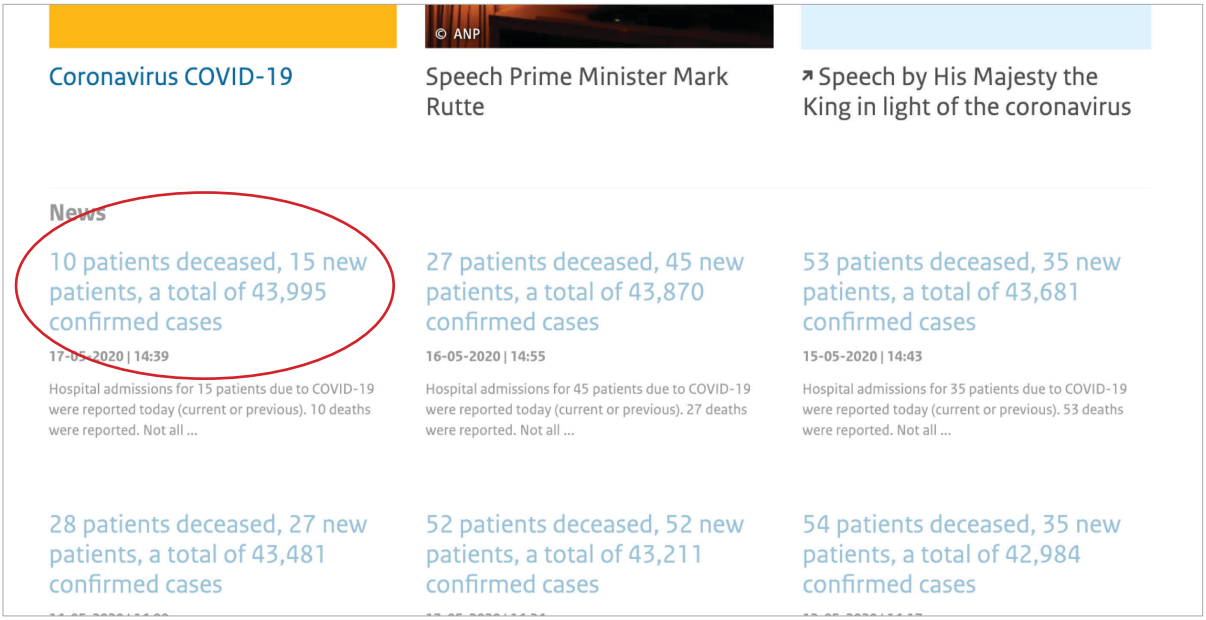


Figure 30: Header and paragraph text for the 'News' items were adjusted to decrease the contrast. The resulting contrast is only 1.8:1, which violates the Nielsen's first heuristic on visibility of system status. Currently, the system status regarding the 'News' items is poorly communicated: *have they been disabled, were they loaded improperly or is there something else going on?*

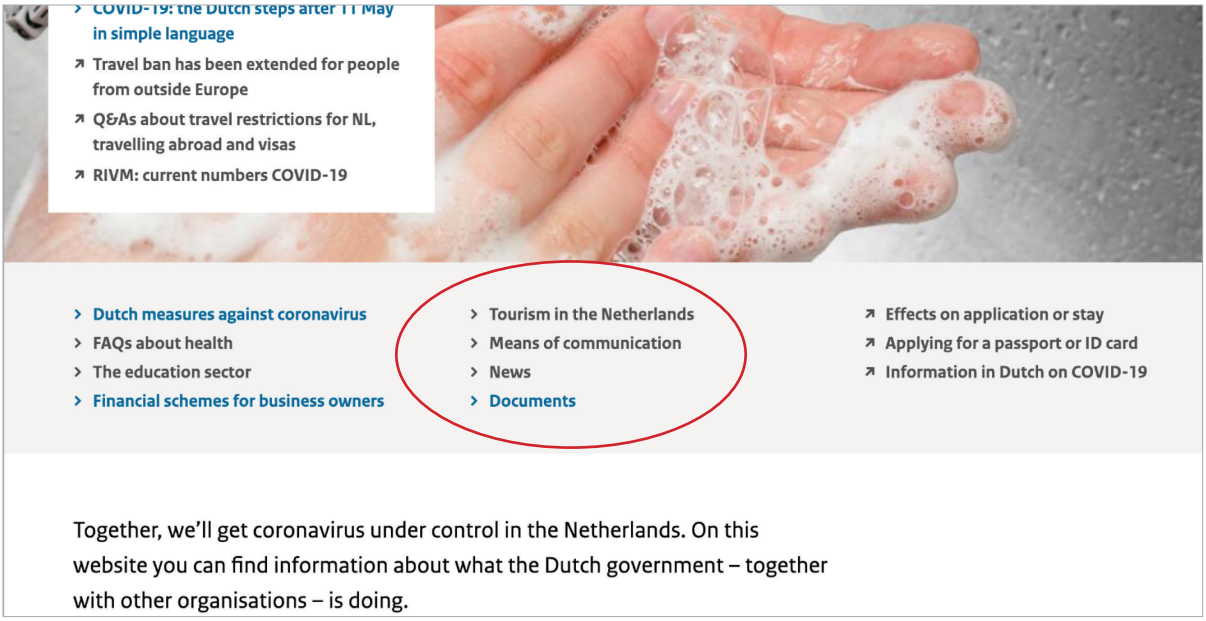


Figure 31: Violation of heuristic 4 on consistency and standards. Links to articles and pages are displayed in different colours, which makes it hard to guess what exactly the difference is.

question which the participants could answer:

> *Optional: Would you be interested in helping us with evaluations in the future?*

This question was intended as an additional measure to determine how the participants had experienced each variant of the prototype. Once the participants felt ready, they could use the button at the bottom of the page to continue to the final questionnaire.

The final questionnaire consisted of a digital implementation of the *NASA Task Load Index* (NASA-TLX) questionnaire [21]. NASA-TLX is a subjective assessment tool which is widely used to assess tasks or systems in terms of perceived workload. The perceived workload is composed of six sub-scales: *mental demand, physical demand, temporal demand, performance, effort* and *frustration*.

In the questionnaire, participants were given descriptions for each of these subscales before these are rated. Ratings are given on a scale that ranges from 0 to 100 points in intervals of 5 points.

After participants had rated the six dimensions of the task, a pair-wise comparison is made between combinations of all subscales, where participants pick the subscale that they perceived as more important for the workload. This results in a weighted score that can range from 0 to 100 points which is known as the overall task load index.

Due to the remote nature of this study, the NASA-TLX questionnaire was not administered as a *pen-and-paper-version* nor through the official Apple iOS App. Instead, a digital implementation in HTML/CSS/JS was used [46]. The implementation was published under the MIT license. I made adjustments to the styling to make the test fit in better with the aesthetic of the other parts of the evaluation, and developed the back-end which allowed the results to be saved to the database.

Once the participants had completed the questionnaire, their participation was concluded. The participants were informed of the end of the study, and were thanked participating. The results of the NASA-TLX questionnaire were displayed graphically, which they could review if they were interested in doing so.

3.5 Summary

Independent variable:
The independent variable is the element that was varied between two sample groups. This was the *user interface type, which had* two categories: *Traditional* and *Conversational*.

Dependent variables:
The dependent variables are the factors that are being studied in this experiment. A total of four dependent variables were recognised: the perceived task load (NASA-TLX), participant task performance (time, success rate), participant responses (amount, details) and the intent to participate in future evaluations.

Controlled variables:
The controlled variables are the things that are kept exactly the same between both sample groups. This experiment has two controlled variables: *evaluation content* (Government.nl case) and *evaluation methodology* (Nielsen's Usability Heuristics).

Extraneous variable:
Finally, extraneous variables are things that may affect the results. For this, we recognized the participants' experience in UX Design.

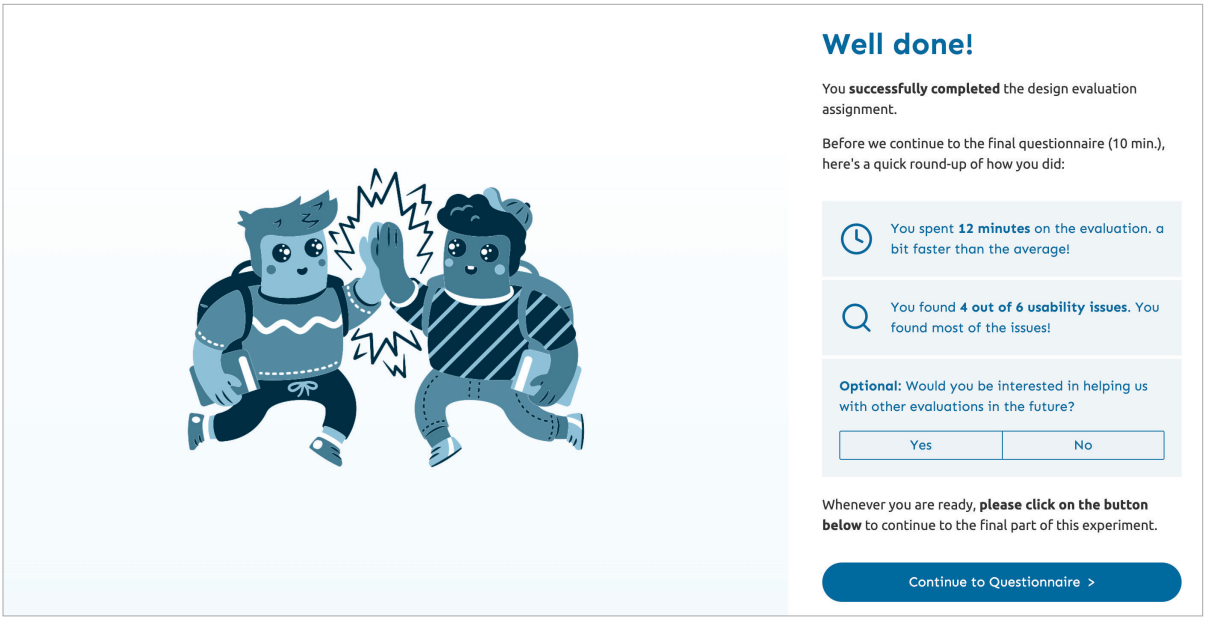


Figure 32: Intermediary screen that was shown to the participants in between the two parts of the evaluation in order to keep them motivated throughout the experiment. *(Illustration shown on the left is used under the Icons8 Universal Media License, and is part of Icons8's 'Flame Illustrations')*

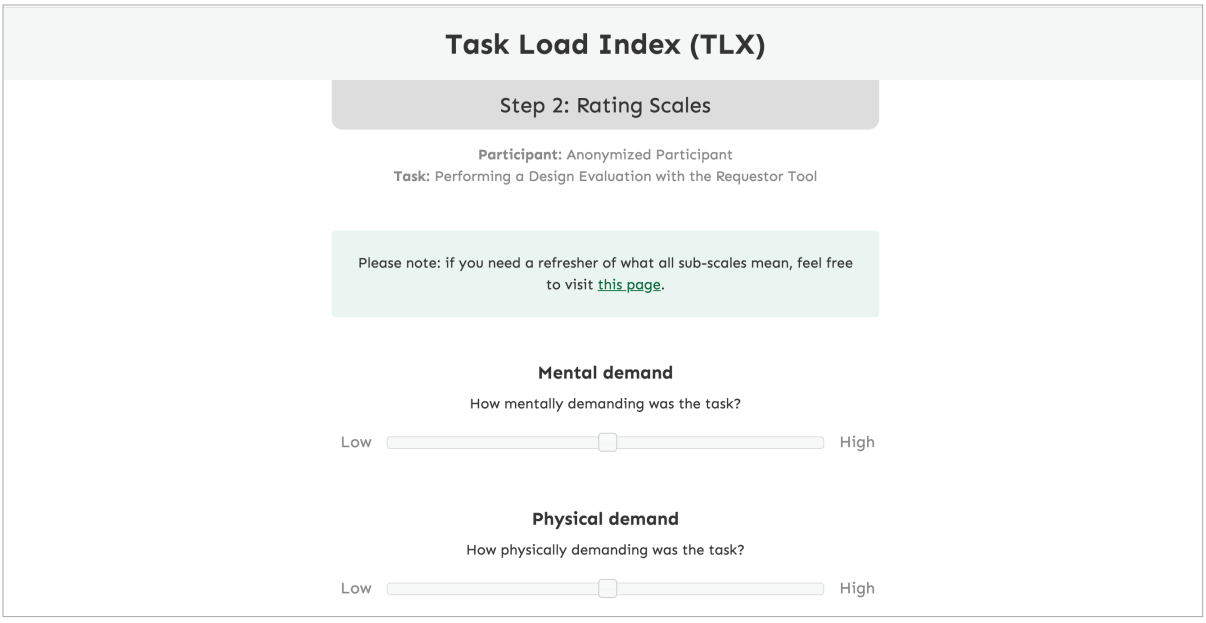


Figure 33: Digital implementation of the NASA Task Load Index questionnaire. The questionnaire can be accessed via https://requestor.nl/nasa-tlx/gui.html

SECTION FOUR

Evaluation Results

In the fourth section the results of the evaluation study of the Requestor concept are presented and analysed. Additionally, the implications of this work are discussed.

4— Evaluation Results

4.1 Overview of Results

After the study's proposal had been reviewed and approved by the local Ethical Review Board from the Industrial Design department (Appendix X), the research study was advertised on social media channels (i.e. LinkedIn, Facebook Groups and Whatsapp) in order to recruit participants that met the inclusion criteria.

A total of 33 individuals consented to participating in this research by submitting the Informed Consent Form. They shared their perceived knowledge of User Experience Design by answering the following question:

> *Please select the option that best indicates how you perceive your knowledge of User Experience (UX) Design below.*

Out of the **33** total participants, **24** completed the design evaluation using the Requestor tool, of which **22** completed the NASA-TLX questionnaire on perceived workload.

An overview of composition of the 22 participants that completed all parts of this study can be seen in Figure X. Of these 22 participants, **12** used the 'traditional' user interface *(from here on referred to as 'GUI')* while **10** used the 'conversational' user interface *(from here on referred to as 'CUI')*.

Furthermore, a quick review of the input and answers provided by the participants revealed that there were no untrustworthy entries present in the dataset, removing the need to clean the dataset. The datasets generated in this study can be found in Appendix D. Finally, it is important to note that this evaluation is presented as a collection of emergent trends in quantitative data, further substantiated by qualitative data.

4.2 Time-Related Performance

Both Requestor variants were equipped with the same functionality that allows the time

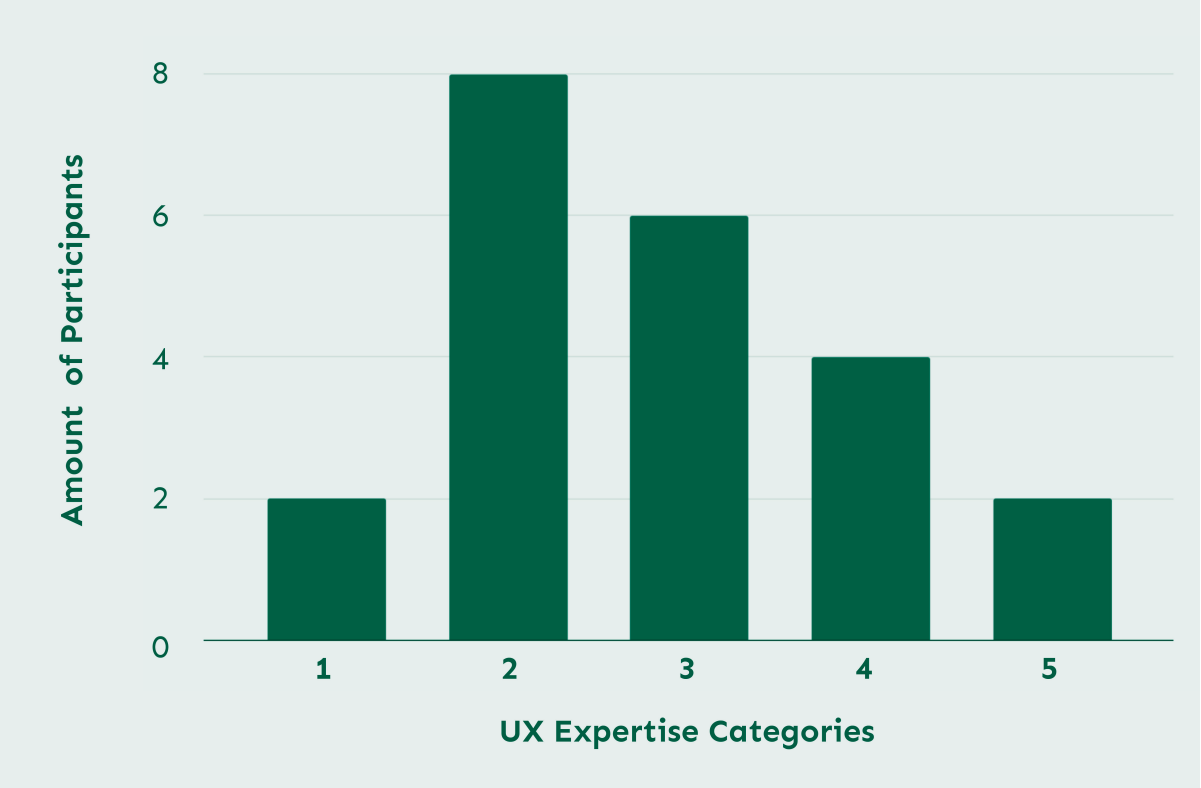


Figure 34: Distribution of the study participants among self-identified expertise categories (n=22).The categories had the following descriptions:

1. *I am unaware of what UX involves and the added value it may bring.*
2. *I have some understanding of what UX is, but have no experience in practical application.*
3. *I know what UX is and have limited experience in practical application.*
4. *I know what UX is and have considerable experience in practical application.*
5. *I am known as an expert and have professional experience in practical application.*

taken to complete tasks to be measured. In this experiment, four intervals were of interest: *Total*, *Onboarding*, *Scenario* and *Evaluation*.

- ***Total:*** *The total amount of time that passed from the moment the page was opened until the evaluation is submitted.*
- ***Onboarding:*** *The amount of time spent by the participant reading 'onboarding materials' such as the project's context, its users and the list of tasks.*
- ***Scenario:*** *The amount of time spent by the participant performing the scenario in the embedded design prototype.*
- ***Evaluation:*** *The amount of time spent providing answers to the questions on the heuristic evaluation.*

The time intervals were recorded in seconds. Descriptive statistics of these four intervals, split by interface variant, are shown in Tables 1 and 2 on the next page. For improved

legibility, the values in seconds have been converted to the format of hours, minutes and seconds (hh:mm:ss). Additionally, boxplot visualisations of all four time intervals for both prototype variants were made in R Studio, are shown in Figures 35 through 38.

Firstly, the results show a great variation in the amount of time that was spent on tasks by different participants. For instance, the median value for the total time spent on completing all tasks using the GUI was 18 minutes and 48 seconds, whereas the maximum recorded value was over 1.5 hours (Figure 35). This suggests that this metric may not reliable enough to draw trustworthy conclusions.

In semi-structured interviews held with participants after they had completed the tasks, an explanation for this range in time was given: one participant stated that they were interrupted by other activities while they were performing the evaluation. As a result, they left the interface open, and resumed the task at a later time. In the way that the time tracking functionality was realized, there was no way to detect such events, rendering the time-related stats unreliable to draw strong conclusions.

Still, the results suggest a number of trends. The median values for time spent performing tasks with the GUI generally are higher than those of the CUI. For example, the median value for completing the design evaluation with the GUI was 15 minutes, whereas the same task took participants operating the CUI 8 minutes and 37 seconds.

While these stats are not fully reliable, one participant shared why they believed the CUI

Table 1: *Descriptive statistics for time-related stats (hh:mm:ss) of GUI users (n=12).*

	Min.	Q1	Median	Mean	Q3	Max.
total	0:07:29	0:10:21	0:18:48	0:32:18	0:51:24	1:32:43
onboarding	0:00:05	0:01:09	0:01:44	0:03:30	0:02:36	0:20:40
scenario	0:01:13	0:01:58	0:02:21	0:04:28	0:03:27	0:25:35
evaluation	0:04:33	0:07:34	0:15:00	0:24:20	0:46:28	0:59:40

Table 2: *Descriptive statistics for time-related stats (hh:mm:ss) of CUI users (n=10).*

	Min.	Q1	Median	Mean	Q3	Max.
total	0:04:32	0:10:45	0:14:45	0:16:14	0:19:21	0:34:22
onboarding	0:00:26	0:00:35	0:00:43	0:02:19	0:01:14	0:17:10
scenario	0:00:32	0:01:56	0:02:39	0:03:16	0:03:54	0:08:28
evaluation	0:01:47	0:04:39	0:08:37	0:10:40	0:11:59	0:27:04

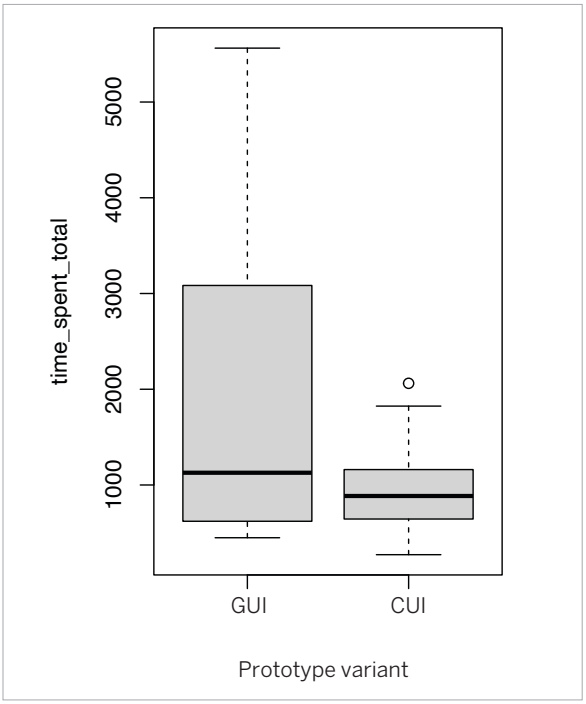


Figure 35: Boxplots for total time spent.

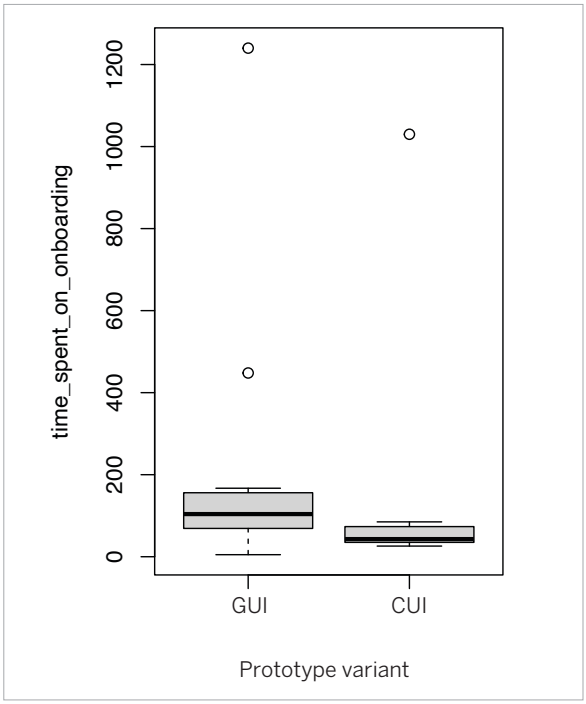


Figure 36: Boxplots for onboarding times.

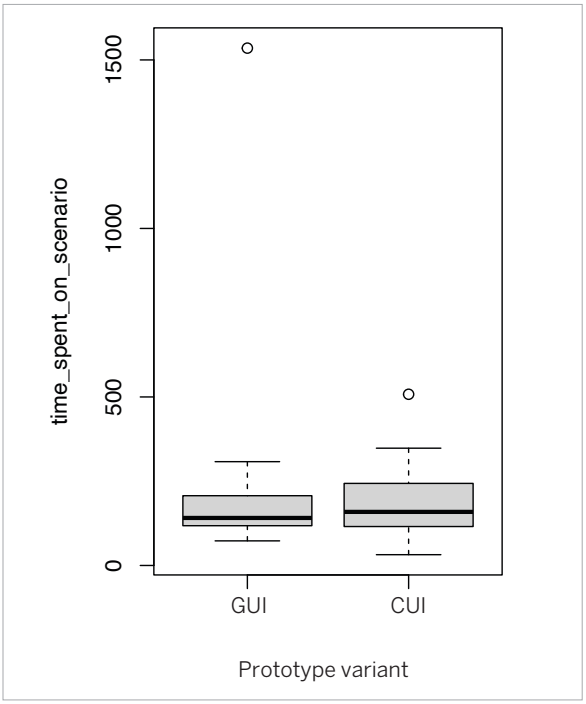


Figure 37: Boxplots for scenario times.

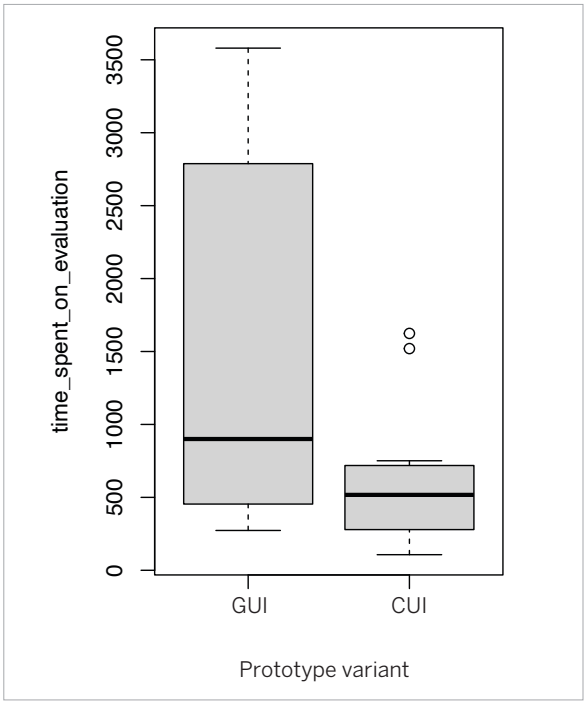


Figure 38: Boxplots for evaluation times.

made them perform the design evaluation faster. The participant mentioned that because the CUI was represented by an agent which used human-like communication in helping them conduct the evaluation, they perceived the agent as 'observing them'. The participant attributed human emotions to the agent, stating:

> *"The interface made me think I was talking to someone. I didn't want to keep them waiting."*

Another potential explanation is the fact that graphical user interfaces can, but do not force a linear interaction path from start to end. With a GUI, it is easier to re-visit information, or to update input that was given earlier, as this interaction style is better suited for that.

Conditional paths in CUIs can also be programmed to allow the conversation to be re-directed to segments that occurred earlier, yet technically implementing such a feature is resource-intensive, which is why this functionality was not included in the research prototype. A such implementation would contribute to a more reliable metric on time-related stats.

4.3 Perceived Work Load

Participants of this study completed a questionnaire on the workload that they perceived while operating one of Requestor's variants. This questionnaire was an implementation of NASA-TLX, which yields scores ranging from 0 to 100, where lower scores indicate a lower perceived work load.

The implementation of NASA-TLX generates a weighted average of the perceived workload *(TLX Score or TLX)*, and also provides indices of the workload experienced in the six sub-scales out of which it is built up. These sub-scales are *Mental Demand, Physical Demand, Temporal Demand, Performance, Effort* and *Frustration*. The mean task load indices for all aforementioned, separated by prototype variant, can be found in Table 3.

Overview

The first thing to note from is the consistently low scores in physical demand for the GUI (M=9.2, SD=8.48) and the CUI (M=9.5, SD=7.6). This makes perfect sense: both prototype variants were used on computers, requiring little physical effort. Still, seeing such uniform ratings of a standardised measure implies that all participants were truthful in answering the questionnaire (Figure 39).

The mean value for the Task Load Index for the GUI (M=46.8, SD=27.9) is higher than that of the CUI (M=38.0, SD=14.2). However, this result is not statistically significant (p=.3516), as determined by a two-sample T-test (Table 4 on the next page). Therefore, the conclusion that the CUI is perceived to have a different perceived work load than the GUI may not be drawn.

Similarly, there is no statistical evidence that the constructs of physical demand (p=.9237), temporal demand (p=.8618) or frustration (p=.6348) vary among the prototype variants (Table 4 on the next page). Consequently, the results do not provide conclusive evidence that the prototype variants have different perceived work loads for these sub-scales.

Perceived Effort

However, there is a statistically significant (p=.048) difference between the perceived effort between the two variants of the prototype, as shown by a two-sample T-test (Table 4 on the next page).

Table 3: *Mean task load indices for the weighted TLX score and TLX sub-scales for the cumulative sample and both individual prototype variants.*

	TLX	Mental	Physical	Temporal	Perform.	Effort	Frustr.
Cumulative (N = 22)	42.7	44.1	9.3	34.1	40.2	37.5	42.0
GUI Var. (N = 12)	46.8	52.9	9.2	28.3	41.7	47.5	45.4
CUI Var. (N = 10)	38.0	33.5	9.5	41.0	38.5	25.5	38.5

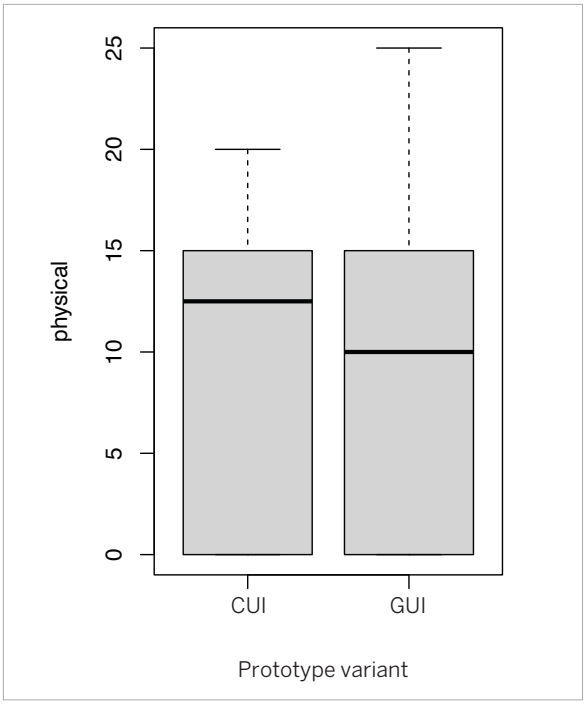


Figure 39: Boxplot for perceived physical demand for both prototype variants.

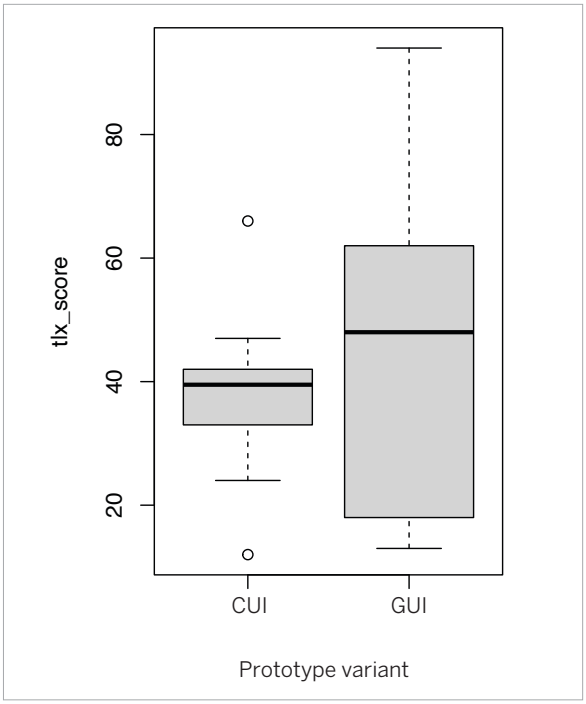


Figure 40: Boxplot for overall Task Load Indices for both prototype variants.

Table 4: *Two-sample t-test results for the NASA-TLX Score and sub-scales.*

	T-value	Degrees of freedom	P-value
TLX_Score	-0.9489	16,883	.3561
Mental	-18.251	18.885	.08384
Physical	0.097025	19.859	.9237
Temporal	10.229	17.836	.32
Performance	-0.23473	19.818	.8168
Effort*	*-21.136*	*18.909*	*.04807*
Frustration	-0.48228	19.988	.6348

When comparing the perceived effort for the both variants, visualised in a boxplot in Figure 41, we see that the CUI (M=25.5, SD=18.9) scores *significantly lower* ($p<.05$) than the GUI (M=47.5, SD=29.5).

In post-study interviews, some comments made by participants provided some insight as to why that was. Participants that had used the GUI variant commonly remarked on the amount of reading required:

> *"There was a lot of text in the website. It felt hard to get through."*

> *"When I opened the page, it felt like there was too much to read. In the end it was okay, but my first impression was feeling overwhelmed."*

Meanwhile, participants that had used the CUI did not share those sentiments:

> *"The conversation felt playful, like a game. It indeed was a lot to process, but not too much."*

> *"I mostly felt in control of the conversation: if I was ready for the next question, I could click on the button. However, there was one instance where I think my screen was too small, because all of the new messages did not fit on my screen."*

The way in which the CUI presents the content to the evaluators may be the reason why the perceived effort is lower. The GUI displays most of the information and steps that the participant was required to carry out at the same time, with some interactive element to it. The interaction design pattern of progressive disclosure is applied to more gradually display content, such as the follow-up questions required when a user signifies a usability problem.

However, the interaction in the CUI is a lot more dynamic. The amount of information conveyed per message is shorter, and new messages are posted in the chat window through an animation, adding some delight to the interaction.

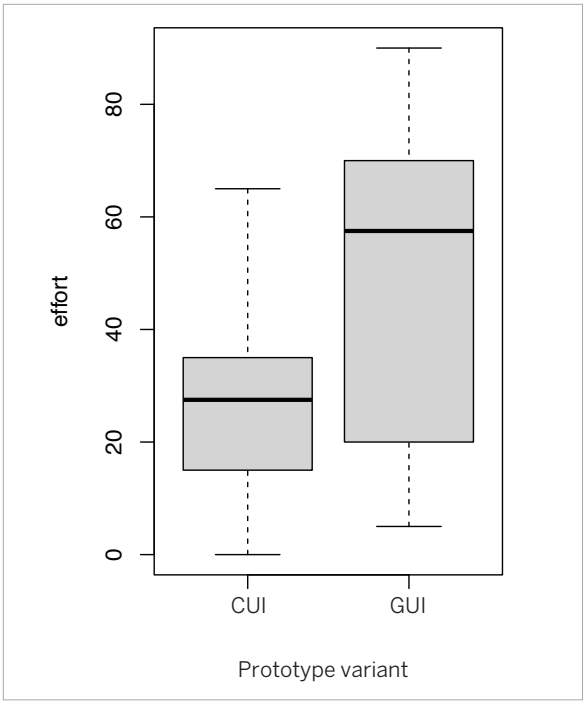


Figure 41: Boxplot for perceived effort for both prototype variants

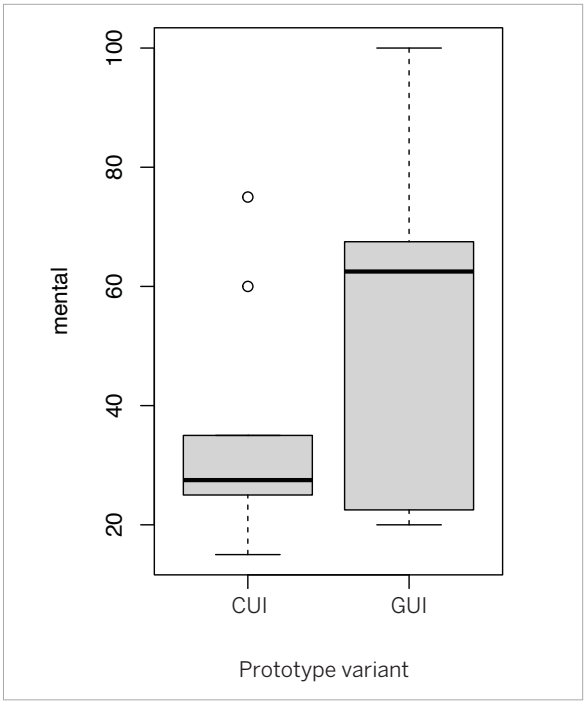


Figure 42: Boxplot for perceived mental demand for both prototype variants.

Perceived Mental Demand

At a first glance, the mean values for perceived mental demand of both prototype variants (Figure 42) also seem to be significantly different. The scores for mental demand for the GUI (M=52.9, SD=30.2.) and the CUI (M=33.5, SD=19.3) contribute to this belief.

A two-sample t-test of these variables deemed this hypothesis statistically insignificant (p=.084). However, if we look at the boxplot in Figure 42, we see two outliers, values that are without 1.5 times the Inter-Quartile Range of the third quartile. The absolute values of these outliers were scores of 60 and 75, respectively.

Another t-test was performed after these outliers had been removed to see if it impacted the outcome. This test resulted in values of t=-3.08, df=12.75 and p=.0090.

With this, we have statistically significant evidence to conclude that the perceived mental demand of the CUI (M=25, SD=7.07) is *lower* than that of the GUI (M=52.9, SD=30.2).

Once more, this may be explained by the difference in nature of the two interaction styles. The GUI is more static whereas the CUI is more dynamic which levels out the perceived mental demand over a broader sequence of interactions.

Furthermore, the two sub-scales that were found significantly different among the prototype variants have some overlap, as 'effort' is described as *'How hard did you have to work (mentally and physically) to accomplish your level of performance?',* which also includes some of the characteristics present in mental demand.

Table 5: *Participant answers to the question on whether they would be interested in helping out with future evaluations.*

	GUI Variant (N=12)	CUI Variant (N=10)	Cumulative (N=22)
Not Interested	6 (50%)	1 (10%)	7 (32%)
Interested	6 (50%)	9 (90%)	15 (78%)
Total	12 (100%)	10 (100%)	22 (100%)

4.4 Attitude Towards Participating

After the participants had completed the design evaluation using the Requestor tool, they were shown a screen where they were provided feedback on their performance, and thanked for their efforts so far. The screen also included a question which the participants could answer:

> *Optional: Would you be interested in helping us with evaluations in the future?*

Participants could choose to answer 'yes', 'no' or choose to leave the field open. Open answers were registered as 'null' in the database, but were converted to 'no' for this analysis. The results of this question are displayed in Table 5.

From this, we see some interesting results. Participants using the GUI variant (N=12) are divided, as half are interested while the other half are not interested. Participants using the GUI (N=10) however are more unanimous: 9 out of 10 participants stated that they would be interested in doing further evaluations.

In order to evaluate the significance of this observation, a Chi-Squared test was performed. There was no significant relationship found between the prototype variants and participant interest, however $X^2 (2, N=22) = 1.89, p=.17$. F

The sample size of this experiment is not yet large enough for a such test, and one of the fundamental assumptions for performing a Chi-Square test are violated, as more than 20% of the cells have expected counts of <5.

Regardless, the participant responses suggest a certain trend which may be met with a larger sample size.

4.5 Discussion

The aims for this study were bifold: firstly, I sought to evaluate the viability of the Requestor platform as a collaborative tool for design. Secondly, I sought to explore the effects that User Interfaces (UIs) with different styles of interactions have on a person's ability to perform an evaluation of a design prototype.

In this study, a total of 24 participants completed a design evaluation of a fictional design case using the Requestor platform, of whom

22 completed the subsequent questionnaire on the workload they had perceived. Participants were randomly assigned either variant of the Requestor platform, with the options being a 'Traditional' Graphical User Interface (GUI) and a 'Conversational' User Interface (CUI).

In this study, the methodology and contents used throughout the participants' design evaluation were kept identical, where the interaction in place to guide them through the evaluation was varied.

A first note on this experiment is that the final sample size (n=22) can not be considered enough to produce conclusive evidence of the trends being present in the data resulting from this experiment. Furthermore, the strategy employed to recruit participants for this study may make it difficult to extrapolate the findings presented to populations other than the sample of this experiment.

Nevertheless, this experiment resulted in a number of valuable results. Firstly, the Requestor platform enabled all participants of this study to complete a heuristic evaluation of the fictional design case, despite the variety in their self-identified expertise of UX Design.

Even though it is difficult to attribute absolute value to a subjective scale such as NASA-TLX, the scores resulting from the perceived workload for both variants received satisfactory scores in perceived workload. This becomes more apparent, especially considering the fact that participants carried out a task usually performed by a design expert, which they did using an interface they had never used before.

Furthermore, by investigating the data resulting from the specified independent variables, a number of trends and themes surfaced. Firstly, the statistical analysis of sub-constructs of the NASA-TLX questionnaire provided statistically significant evidence that the Conversational User Interface was perceived to result in a lower work load than the 'Traditional' User Interface, both in *Perceived Effort* as well as in *Perceived Mental Demand*.

These results may be taken as design guidelines for the design of user interfaces that support non-experts in performing tasks that normally are performed by experts. I would especially like to highlight the inherent qualities that Conversational User Interfaces have which may make the process of understanding- and learning these tasks less steep.

Furthermore, while this study did not produce statistically significant evidence, its results suggest that the CUI is the preferred style of interaction among participants of this study for the task at hand.

However, further work is needed to provide a more comprehensive overview of the matter. A significant limitation of this study was its sample size, as between-subjects comparisons generally require 86 participants or more to ensure statistical significance. A larger sample size would have resulted in more substantial conclusions to be drawn, especially since it would have been very likely that it would have provided a conclusive answer to Section 4.4.

Furthermore, it would be worthwile to more thoroughly examine the contents of the participants' answers given in both variants of the prototype, as well as evaluate their performance based on the amount of correctly identified usability issues. Initially, the intention had been to do exactly that. However, this was given a lower priority due to the qualitative nature of these data, making it challenging to present a conclusive and dependable judgement on the matter. Nevertheless, this could provide interesting insights that may be used for the development of similar systems.

4.6 Conclusion

In this report the design process of 'Requestor', the online platform to share UX works-in-progress was presented. First, the design decisions taken in the iterative realization of the concept were discussed, seeking to find an answer to the question:

> *How might we create a better collaborative space for UX Designers and their environment?*

Afterwards, the aims of an evaluation of the design concept are explained, and its process is described. In a statistical analysis of resulting quantitative data, the results of the evaluation are presented, providing insights that are backed by qualitative data from interviews.

In the next section I take a reflective look at the Final Master Project, and conclude this report by sharing my personal growth.

SECTION FIVE

Reflection

In the fifth and final section I look back and reflect on this Final Master Project.

5— *Reflection*

5.1 On the Past

Throughout my Bachelor's degree I learnt to make physical and digital prototypes, acquired some experience in programming and started to develop the ability to direct my own design processes. It was not until my B3.1 internship where I also found a design space which fits me well: UX Design.

In my Master's degree I had one overarching goal: to continue developing myself as a UX Designer. However, as time passed I started to realise that by itself it may not be enough. I wanted to develop a more unique competency profile by specialising in digital technologies, allowing me to better collaborate with others in my preferred career of working in digital, high-tech environments.

When I was reflecting on my experiences throughout both studies I found that one thing that always bothered me was the fact that all I seemed to be doing was creating prototypes, never truly creating products that I felt were finished. Often this would be because of restrictions in time, resulting in questionable design decisions. Additionally, I felt that I never had all the skills required to complete the full process by myself, or learn enough during the project to do so.

5.2 On the Process

The Final Master Project provided me with the opportunity to integrate all of the skills, attitude and knowledge developed at the faculty over the past years. In this project, I wanted to do everything right.

I went through a highly iterative design process in which I sought to get ideas out of my head and into the hands of others as soon as I could. Above all, I prioritised ending this project with a fully functioning system.

However, before I was there, there were a number of hurdles to be taken. Prior to this

project, I had never created any system that was able to reliably retrieve- and save persistent data.

One of the biggest moments of growth for me during this project was bridging the gap to back-end development. Bolstered in confidence by the electives in Computer Science I took in my M21 option, I took it upon myself to start learning PHP as a server-side language and MySQL to access databases.

I integrated these new skills in the first 'fully functional prototype', which felt like magic — I got the feeling that I was finally moving beyond the territory of making prototypes and into the field of making actual products.

This project however was also characterised by the circumstances of the world during a pandemic, which complicated things. Firstly, there was the issue of staying motivated, and keeping inspiration levels high. Secondly, I was unable to complete the project in the way that I had initially envisioned, which was by doing an in-depth, qualitative evaluation of the concept in a 'real-life' environment, preferably within a company in Eindhoven that shared the problem of low UX Maturity.

Fortunately, I had a great group of fellow master students with whom I had daily 'stand-up meetings', discussing our projects. I learnt that as a designer it is impossible to do things all alone, even if I would like it to be that way. One very valuable skill which I have developed during my Master's degree is the ability to adapt, especially by allowing perspective from the outside.

I adapted my expectations on the evaluation of this project, as the in-context evaluation as initially formulated no longer was feasible. Instead, I saw an opportunity to take advantage of the fact that the design process had resulted in multiple fully functioning prototypes.

For me, the research study conducted to conclude this project was the most thorough and well executed out of any I had done so far. The prototypes presented in the study worked well, and the experiment was set up to generate objective insights.

5.3 On the Future

Firstly, I would like to share my ambition in further working on this project, specifically on the evaluation. My biggest point of criticism on the study was the fact that the number of recruited participants was not enough to draw a reliable conclusion for all components of the experiment. Nevertheless, the trends suggested in the results are definitely interesting, which was affirmed by the comments made by my graduation mentor Dr. Khan. For that reason, I am interested in extending the experiment to accept a greater number of participants in order to produce well-grounded insights.

Furthermore, I feel that I am ready to make the change to start working on long-term projects in a professional career as a UX Designer. The Final Master Project has given me confidence in my ability to independently direct a design process from an early brief to a strong realization of a final concept. The world will remain in a strange place due to the restrictions we impose upon ourselves to contain the virus, yet I feel well-equipped to start contributing — be it through remote means or in physical presence.

Acknowledgements

I would like to express my sincerest gratitude to everybody that contributed to this project. In particular, I would like to thank:

Dr. Javed Khan, for mentoring me over the past 1.5 years. I could not have wished for a better mentor, guiding me towards my vision of becoming a Pi-Shaped UX Designer.

Marco and *Melvin* for the hours spent adapting to the new 1.5-meter-society, where we discussed design in video conference-calls and collectively grew an interest in baking bread.

All friends, family and others that took time out of their days to hear me talk about this project, or took part in the various user tests throughout the process.

All members and staff of the ID Faculty who have helped me in my pursuit of better understanding design by providing courses, holding discussions and giving feedback on my work.

Kees, *Frederika*, *Dieuwertje* and *Johan* for their unconditional support and their recurring interest in this project, even though the exact point of what I was doing may have eluded some of you at points in time.

Nienke, for being there for me when I most needed it. This project was a big challenge, and I could not have done it without you.

References

1. Abdul-Kader, S. A., & Woods, J. C. (2015). Survey on chatbot design techniques in speech conversation systems. International Journal of Advanced Computer Science and Applications, 6(7).

2. Allen, J. J., & Chudley, J. J. (2012). Smashing UX design: Foundations for designing online user experiences (Vol. 34). John Wiley & Sons.

3. Andrews, A. (June, 2017). Scrum Of One: How to Bring Scrum into your One-Person Operation. Retrieved December 1, 2019, from https://www.raywenderlich.com/585-scrum-of-one-how-to-bring-scrum-into-your-one-person-operation

4. Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An empirical evaluation of the system usability scale. Intl. Journal of Human–Computer Interaction, 24(6), 574-594.

5. Brown, T., & Wyatt, J. (2010). Design thinking for social innovation. Development Outreach, 12(1), 29-43.

6. Brumberger, E., & Lauer, C. (2015). The evolution of technical communication: An analysis of industry job postings. Technical Communication, 62(4), 224-243.

7. Care/of. (n.d.). Care/of: Personalized Daily Vitamin Packs. Retrieved from https://takecareof.com/

8.Chai One (2014, December). Improve ROI with User Experience Research — The Secret to Enterprise Software Success. Retrieved from http://chaione.com/wp-content/uploads/2015/10/Improve-ROI-with-User-Experience-Research.pdf

9. Chapman, L., & Plewes, S. (2014, June). A UX maturity model: Effective introduction of UX into organizations. In International Conference of Design, User Experience, and Usability (pp. 12-22). Springer, Cham.

10. Charlette, R. (2005, September). Why Software Fails — We waste billions of dollars each year on entirely preventable mistakes. Retrieved from https://spectrum.ieee.org/computing/software/why-software-fails

11. Demirkan, H., & Spohrer, J. C. (2018). Commentary—cultivating T-shaped professionals in the era of digital transformation. Service Science, 10(1), 98-109.

12. Dumas, J. S., Dumas, J. S., & Redish, J. (1999). A practical guide to usability testing. Intellect books.

13. Fanguy, W. (June 24, 2019). Design Systems — A comprehensive guide to design systems. Retrieved from https://www.invisionapp.com/inside-design/guide-to-design-systems/

14. Fesseden, T. (October, 2017). First Impressions Matter: How Designers Can Support Humans' Automatic Cognitive Processing. Retrieved from https://www.nngroup.com/articles/first-impressions-human-automaticity/

15. Figma. (2019). Figma | The all-in-one design tool built for collaboration. Retrieved from https://www.figma.com/collaboration/

15. Figma. (2019). Figma Developer Platform — Build for your team to create your perfect workflow. Or extend what's possible in Figma for millions of users. Retrieved from https://www.figma.com/developers

16. Gagné, M., & Deci, E. L. (2005). Self-determination theory and work motivation. Journal of Organizational behavior, 26(4), 331-362.

17. Geel, A.J. (July, 2018). Momkai — Half-year Internship. Retrieved from https://arthur-geel.com/momkai.html

18. GoodUI. (n.d.). Exploring: Natural Language Forms. Retrieved from https://goodui.org/blog/exploring-natural-language-forms/

19. Harley, A. (February, 2018). UX Expert Reviews. Retrieved from https://www.nngroup.com/articles/ux-expert-reviews/

20.Hart, S. G. (2006, October). NASA-task load index (NASA-TLX); 20 years later. In Proceedings of the human factors and ergonomics society annual meeting (Vol. 50, No. 9, pp. 904-908). Sage CA: Los Angeles, CA: Sage Publications.

21. Hassenzahl, M., Burmester, M., & Koller, F. (2003). AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In Mensch & computer 2003 (pp. 187-196). Vieweg+ Teubner Verlag.

22. Joyce, A. (July, 2019). Formative vs. Summative Evaluations. Retrieved from https://www.nngroup.com/articles/formative-vs-summative-evaluations/

23. Kashfi, P., Nilsson, A., & Feldt, R. (2017). Integrating User eXperience practices into software development processes: implications of the UX characteristics. PeerJ Computer Science, 3, e130.

24. Knemeyer, D. (2015). Design thinking and UX: two sides of the same coin. Interactions, 22(5), 66-68.

25. van Kollenburg, J., & Bogers, S. J. A. (2019). Data-enabled design : a situated design approach that uses data as creative material when designing for intelligent ecosystems. Eindhoven: Technische Universiteit Eindhoven.

26. Koskinen, I., Zimmerman, J., Binder, T., Redstrom, J., & Wensveen, S. (2011). Design research through practice: From the lab, field, and showroom. Elsevier.

27. Law, E. L. C., Roto, V., Hassenzahl, M., Vermeeren, A. P., & Kort, J. (2009, April). Understanding, scoping and defining user experience: a survey approach. In Proceedings of the SIGCHI conference on human factors in computing systems (pp. 719-728).

28. Lee, Y., Kozar, K. A., & Larsen, K. R. (2003). The technology acceptance model: Past, present, and future. Communications of the Association for information systems, 12(1), 50.

29. Lemonade. (n.d.). Lemonade Contents Insurance & Liability Insurance. Retrieved from https://www.lemonade.com/nl/en

30. Locomotive. (n.d.). Locomotive — Montreal Web Agency. Retrieved from https://locomotive.ca/en

31. Lull, D. (2017). UX and Society. In Discussions in User Experience (pp. 5-5). Apress, Berkeley, CA.

32. MacDonald, C. M. (2019, June). User Experience (UX) Capacity-Building: A Conceptual Model and Research Agenda. In Proceedings of the 2019 on Designing Interactive Systems Conference (pp. 187-200). ACM.

33. McKenzie, P. (2010, February 27). Lesson from Madlibs Signup Fad: Do Your Own Tests. Retrieved from https://www.kalzumeus.com/2010/02/27/lesson-from-madlibs-signup-fad-do-your-own-tests/

34. McTear, M. F. (2002). Spoken dialogue technology: enabling the conversational user interface. ACM Computing Surveys (CSUR), 34(1), 90-169.

35. Nasirian, F., Ahmadian, M., & Lee, O. K. D. (2017). AI-based voice assistant systems: Evaluating from the interaction and trust perspectives.

36. Nielsen, J. (1992, June). Finding usability problems through heuristic evaluation. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 373-380). ACM.

37. Nielsen, J. (2000, March). Why You Only Need to Test with 5 Users. Retrieved from https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/

38. Nielsen, J. (2006, April 23). Corporate UX Maturity: Stages 1-4. Retrieved from https://www.nngroup.com/articles/ux-maturity-stages-1-4/

39. Nielsen, J. (2006, April 30). Corporate UX Maturity: Stages 5-8. Retrieved from https://www.nngroup.com/articles/ux-maturity-stages-5-8/

40. Oscar Insurance. (n.d.). Enroll with Oscar. Retrieved rom https://www.hioscar.com/individuals

41. Øvad, T., & Larsen, L. B. (2016). How to reduce the UX bottleneck–train your software developers. Behaviour & Information Technology, 35(12), 1080-1090.

42. Plonka, L., Sharp, H., Gregory, P., & Taylor, K. (2014, May). UX design in agile: a DSDM case study. In International Conference on Agile Software Development (pp. 1-15). Springer, Cham.

43. Sauro, J., Johnson, K., & Meenan, C. (2017, May). From Snake-Oil to Science: Measuring UX Maturity. In Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (pp. 1084-1091). ACM.

44. Sand Hill Group & Neochange (2008). Achieving enterprise software "success": A study of buyer and seller perspectives on the drivers of enterprise software success.

45. Sand Hill Group, Neochange, & TSIA (2009). Achieving enterprise software success: Insight from software buyers and sellers on the drivers, expectations and responsibilities for success.

46. Schwarz, F. (2012). Nasa-tlx: an implementation of the NASA Task Load Index (NASA-TLX). (Github). Retrieved from https://github.com/isellsoap/nasa-tlx

47. Spillers, F. (2014, July). Making a Strong Business case for the ROI of UX [INFOGRAPHIC]. Retrieved from https://www.experiencedynamics.com/blog/2014/07/making-strong-business-case-roi-ux-infographic

48. Stickdorn, M., Hormess, M. E., Lawrence, A., & Schneider, J. (2018). This is service design doing: Applying service design thinking in the real world. " O'Reilly Media, Inc.".

49. Tervo, V. (January 8, 2015). From T to Pi: design skill expectations in change. Retrieved from https://www.futurice.com/blog/from-t-to-pi-design-skill-expectations-in-change/

50. Therbadji, D. (2020). Chat-bubble.js (Github repository). Retrieved from https://github.com/dmitrizzle/chat-bubble

51. Trade Coffee. (n.d.). Trade Coffee | The Best Coffee Subscription in the Nation. Retrieved from https://www.drinktrade.com/

52. Userzoom. (2018). The State of UX In Enterprise 2018. Retrieved from https://info.userzoom.com/the-state-of-ux-in-the-enterprise-report.html

53. Userzoom. (2019). The State of UX In Enterprise 2019. Retrieved from http://info.userzoom.com/rs/293-RDJ-600/images/The-State-of-UX-in-the-Enterprise-2019_UserZoom.pdf

54. UXPin. (2018). Enterprise UX Industry Report 2017-2018 — the most in-depth look at designing B2B products today. Retrieved from https://www.uxpin.com/enterprise-ux-design-2017-2018-industry-report

55. UXTools. (2019). Design Tools Survey 2019. Retrieved from https://uxtools.co/survey-2019/

56. UXTools. (2019). 2019 Design Tools Survey Responses [Data file]. Retrieved from https://docs.google.com/spreadsheets/d/1bP2v_K5loavHRR8PdrabgZFU-ReXyMvJr1qsVGHtTfsA

57. World Wide Web Consortium. (June 5, 2018). Web Content Accessibility Guidelines (WCAG) 2.1 — W3C Recommendation 05 June 2018. Retrieved from https://www.w3.org/TR/WCAG21/

58. Wrobolewski, L. (February 25, 2010). "Mad Libs" Style Form Increased Conversion by 25-40%. Retrieved from https://www.lukew.com/ff/entry.asp?1007

59. Zwarts, J. (2017). Vibrant.js (Github repository). Retrieved from https://github.com/jariz/vibrant.js/

Appendices

Please note: due to the quantity of appendices and the format of some (i.e. Prototypes & Code, Dataset), they are not included in this report. Rather, they can be accessed with the following link:

https://requestor.nl/documents/appendices

To protect the privacy of the participants, all data has been anonymised. Furthermore, it can only be unarchived by those that have the password, which is: 'DER200' .