Investigating Operation Challenges of IoT Systems to Improve the Usability through User Interface Design

Isa Buwalda

A thesis for the master's degree of
Human Computer Interaction

Supervisors:
Thomas Kosch, Utrecht University
Ioanna Lykourentzou, Utrecht University
Tom Gielen, Coffee IT

Faculty of Science

December 27, 2022

Abstract

The Internet of Things (IoT) has been a rapidly expanding network in recent years and widely studied. However, many scholars focus on the technological features of the IoT network. Nevertheless, challenges still arise regarding usability and user experience (UX) of the systems in this network, especially in the consumer domain of IoT. The current study assessed those challenges by conducting a literature review. To confirm these challenges, a prestudy was conducted with ten participants that interacted with an IoT system. Results showed that the most important issues arise due to the lack of standardization amongst IoT systems, interoperability problems, the complexity of the systems, and the lack of feedback in them.

Next, criteria for the design of user interfaces of IoT systems were proposed to overcome the IoT network-related challenges. These criteria were implemented in the IoT system that was used in the prestudy. To test the effect of these criteria, another study was conducted with ten participants that interacted with the system. The System Usability Scale (SUS) showed that the usability of the system had increased significantly. The AttrakDiff results demonstrated a non-significant increase of UX. Furthermore, the qualitative results of the interviews also suggested an increase in UX.

The study thus offers new insights into designing user interfaces of IoT systems that are usable for consumers. Future work could replicate the study with a larger and more diverse sample size to generalize the results to a larger population. Furthermore, it would be interesting to investigate ways to standardize the development of IoT systems amongst companies developing IoT by making the suggested design criteria the norm.

Keywords: IoT; usability; UX; user interface design; design criteria; SUS; AttrakDiff

Contents

List of Abbreviations

EPC	Electronic Product Code
IoT	Internet of Things
QoE	Quality of Experience
QoS	Quality of Service
RFID	Radio Frequency Identification
SUS	System Usability Scale
UX	User Experience
WiFi	Wireless Fidelity
WSN	Wireless Sensor Networks

Chapter 1

Introduction

1.1 Motivation

At the beginning of computing history, computers were very scarce and expensive. Back then, there was a many-to-one relationship between humans and computers, where users were forced to share one computer amongst big groups of people. However, in 1984 the number of people having their own computers surpassed the number of shared computers. The relationship between humans and computers thus quickly evolved into a one-to-many relationship (West, 2011).

In parallel to this development, the internet was a quickly evolving concept. From 1990 to 1995, the number of computers on the internet increased by more than twenty-fold (Glowniak, 1998). Not only computers were connected to the web, but also other devices or "things" got connected to the internet. This gave birth to the Internet of Things (IoT), a network of interconnected intelligent objects (Suresh et al., 2014).

Ever since the birth of this IoT network, the number of devices connected to it has been expanding rapidly. Studies claim that in 2025 there will be 30.9 billion IoT devices[1], creating a complex ecosystem generating huge amounts of data (Fang and Yan, 2020). The growth over time of devices connected to the internet since 2010 can be seen in Figure 1.1. This image also shows the number of non-IoT devices connected to the internet, such as smartphones and laptops. The graph gives a clear overview of how IoT devices surpass the amount of non-IoT devices.

Although IoT has been a widely studied topic, many scholars focus on the technological features of the concept. An example of such a study is by Delgado Rodriguez et al. (2021), the authors developed a smart desk platform that enables the user to interact with multiple IoT devices at once. This *ActPad* uses touch points attached to objects by connectors. The authors claim that this approach of centralizing interaction within the home and controlling IoT devices requires much less additional hardware and software than for example smart home hubs. Other research focuses on improving the IoT architecture. For example, *BlockToIntelligence* is an architecture combining Artificial Intelligence and Blockchain to achieve the goal of scalable and secure IoT (Singh, Rathore, and Park, 2020). However, even though these scholars aim to make IoT systems more usable, they do not take a user-centered perspective.

However, there are still challenges arising when taking a user-centered perspective regarding IoT, despite technological innovations. One such challenge is that interaction with the different IoT devices is handled separately. This can create cognitive overload for the users as the number of devices and their functional complexity is still increasing (Nazari Shirehjini and Semsar, 2017). Although *ActPad* attempts

[1] https://iot-analytics.com/state-of-the-iot-2020-12-billion-iot-connections-surpassing-non-iot-for-the-first-time/

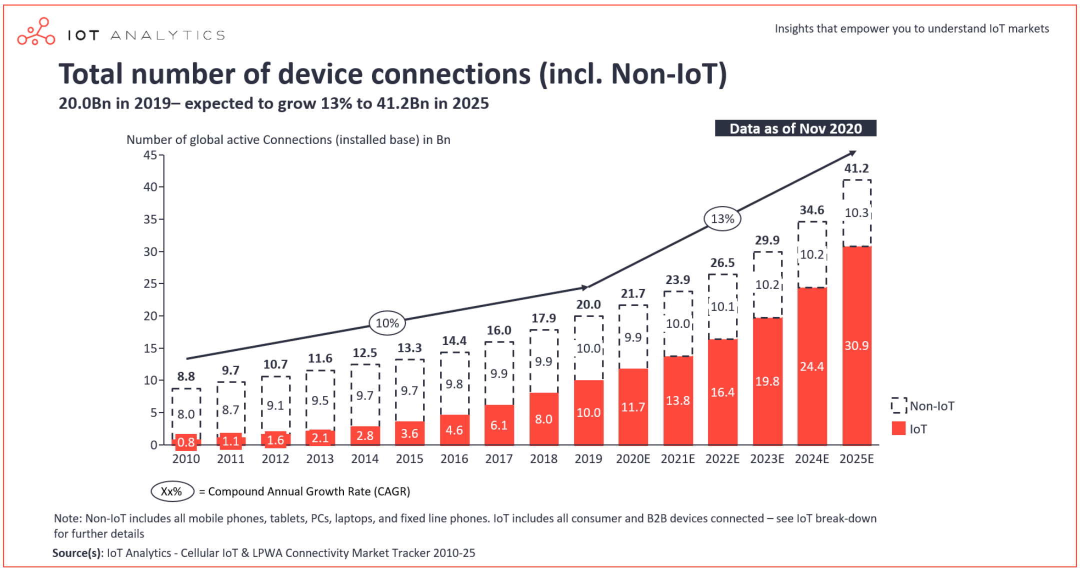


FIGURE 1.1: Number of devices connected to the internet from 2010 to 2025[1].

to handle this phenomenon, the user experience (UX) of this platform is not tested in the study. Another important challenge relates to the interoperability of IoT systems. The ecosystem of IoT devices is still growing, and often the devices have to interact and communicate with each other. Bergman et al. (2018) state that issues arise in the interaction and communication amongst devices because there is no standardization amongst IoT developing companies. These issues result in a decreasing UX.

The current study aims to close the gap found in literature by taking a user-centered perspective. The scope of this research will be narrowed towards consumer IoT as research by Bergman et al. (2018) suggests that this is the most interesting application domain of IoT to look into. **In this study, it will thus be investigated what challenges users currently face when operating consumer IoT devices. Consequently, design criteria will be developed and evaluated to support the users in operating IoT devices.**

1.2 Research Problem & Questions

There are no usability guidelines for evaluating and designing IoT devices due to the absence of user-centered research in IoT. This might cause issues with UX. Since UX is a significant factor in the overall quality of a device (Shin, 2017), it will be studied in current research. First, it is essential to identify what challenges arise when taking a user-centered approach. This gives rise to the first research question of this thesis.

RQ1: What are current UX challenges when interacting with IoT systems?

Here, a focus will be on the usability of consumer IoT systems. Usability is a subdomain of UX, and increasing usability thus increases UX. The challenges will be identified through a systematic literature survey. This survey will provide an initial overview of UX studies within the IoT domain. The knowledge will be used to derive a procedure for a prestudy. Participants will be interviewed after using an IoT device to identify UX challenges concretely. The showcase device to be used will be an IoT lighting system, namely the in-lite system[2] with the accompanying app developed by Coffee IT[3]. This app is used to control the lighting system. It is predicted that the results of this study can be generalized to other consumer IoT systems, as the in-lite system is an IoT system currently sold on the consumer market, and its operation is similar to other consumer IoT systems. Namely, with a graphical user interface on a separate device, such as a phone or a tablet. The system can be automated or remotely monitored and operated.

In the next part of the study, design criteria will be developed to overcome the challenges established based on the prestudy and the interviews. This leads to the second research question.

RQ2: How can graphical user interfaces of IoT systems be designed to overcome UX challenges?

The design criteria developed will be applied to the lighting control user interface of the aforementioned application. In the next step, a user study will be conducted to assess whether the derived design requirements improve the operation of the IoT system.

1.3 Outline

In the following chapter, literature related to the research topic is discussed. First, the concepts of IoT, usability, and UX are defined. Next, research related to the usability and UX of IoT is set out. Lastly, the challenges found in literature regarding user experience are listed.

The conducted prestudy is described in Chapter 3. The method and the results of this study are set out. These results give quantitative and qualitative insight into the UX challenges users face when interacting with the showcase IoT system.

Chapter 4 discusses the design criteria for IoT derived from the results of the prestudy. This chapter also presents how these design criteria were implemented in the showcase IoT system.

In Chapter 5, the main study is discussed, which is conducted to evaluate the derived design criteria. The method of this study is described in this chapter, and the results are set out.

The whole study is discussed in Chapter 6. This chapter reviews the results of the prestudy and the main study. This chapter also gives insight into the limitations of the study and gives suggestions for future work.

Lastly, the thesis is concluded in Chapter 7.

[2] https://in-lite.com/nl/
[3] https://coffeeit.nl/

Chapter 2

Related Work

Previous work investigated the functionality of IoT systems. At the same time, past research looked into how the usability of IoT devices can be improved. This section elaborates on both aspects.

2.1 Internet of things (IoT)

The term 'Internet of Things' was first introduced by Kevin Ashton in 1999; he used it to refer to a system in which objects in the physical world could be connected to the internet by sensors (Rose, Eldridge, and Chapin, 2015). However, the idea of connecting physical objects to the internet was not new. In 1990, John Romkey created a toaster that could be controlled over the internet. This toaster is seen as one of the first IoT devices presented in public. One of the first commercial IoT devices was a smart refrigerator presented by LG in 2000. This refrigerator would be able to determine whether the items in it were fresh or not (Suresh et al., 2014).

In recent years, the term IoT has been widely used for discrepant devices and technologies, inducing confusion about what the term actually comprises. Present-day, there is no universally accepted specific definition of the term IoT. This is mainly because the idea of IoT has developed over time and will probably do so in the upcoming years due to the constant development of innovative technologies. These technologies include cloud computing, big data, and social networking (Atzori, Iera, and Morabito, 2017). One definition by Suresh et al. (2014) that seems to clearly cover the meaning of IoT today is:

> *An open and comprehensive network of intelligent objects that have the capacity to auto-organize, share information, data, and resources, reacting and acting in face of situations and changes in the environment.*

Thus, the IoT network not only allows computer devices to be connected but also other physical objects. This is done by equipping these objects with sensors and actuators. This way, the objects can be connected to and can communicate with the IoT network (Madakam et al., 2015). Consequently, IoT can be found almost everywhere around us, for example, in healthcare, education, homes, cars, and industrial factories.

IoT can thus be applied in various domains. Xenofontos et al. (2021) distinguish between three major application sectors, which are the following:

1. **Consumer IoT**: this domain targets the end-user. It comprises personal devices, such as wearables and smartwatches, and internet-connected home devices and appliances that can collect data and can be remotely monitored and controlled, like smart lamps and thermostats.

2. **Commercial IoT**: this refers to IoT applied in enterprises and bigger infrastructures. While minimizing operational cost and service latency, they are used to automate, coordinate and respond to their environment.

3. **Industrial IoT**: IoT in this domain is created to enable real-time exchange of industrial system information, provide better situational awareness, improve the control over system processes and increase productivity and efficiency.

The shared and common goal of all IoT applications is to provide smart services to increase the quality of human life. Technologies are created and used to monitor, manage and automate human activities to satisfy this goal (Asghari, Rahmani, and Javadi, 2019). One place where IoT clearly eases human life is in houses, these types of IoT are part of the consumer IoT systems. Smart home technologies have been quickly developing over the past years. In a smart home, management functions are integrated that enable the user to control their house optimally. IoT allows objects to be connected to a bigger network of objects. The data gathered through the sensors on these objects is provided to the network and can be used to manage and automate human activities (Choi et al., 2021). Examples of such human activities are managing the temperature in the house or turning the lights on and off.

It is important to note that IoT is closely related to ubiquitous computing. This term refers to the presence of technology everywhere around us but concealed to the background, which enables tasks to be completed with little or no interaction with the user. The term goes hand-in-hand with IoT because, generally, this technology is connected to the internet. The interaction between the user and these systems is often limited to the system's set-up or the device's management and communication of the recorded data (Resnick, 2013).

2.2 IoT Architecture

Because of the quick development of IoT and its broad range of services and devices, there is no uniform architecture for IoT. In literature, different architectures or models are adopted (Madakam et al., 2015). Farooq et al. (2015) claim IoT is generally divided into six layers, which are shown in Figure 2.1. Each layer will be described below.

The Coding Layer

In this first layer, the to-be-connected objects are assigned a unique ID. This allows for differentiating the objects.

The Perception Layer

This layer measures the useful information of the objects, which can be done with data sensors of different forms. Lee and Lee (2015) describe two of those data sensors. The first is radio-frequency identification (RFID). This system uses radio waves, a tag, and a reader for automatic identification of an object and for data storage. The tag carries data in the form of an Electronic Product Code (EPC), which allows the identification of an object. This EPC is thus assigned in the coding layer. There are three types of RFID tags, namely passive, semi-passive and active. They are distinguished by their battery source. The second technology is wireless sensor networks (WSN). These networks consist of devices equipped with sensors and can monitor physical or environmental conditions. They are spatially distributed, and

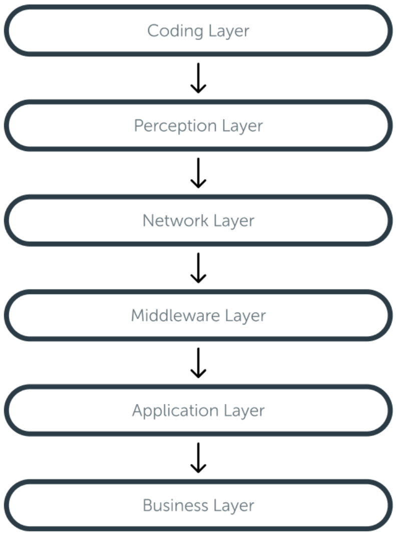


FIGURE 2.1: Proposed six-layer IoT architecture by Farooq et al. (2015).

in cooperation with the RFID systems, they can measure features like temperature and movement at different locations. Furthermore, the devices in these networks communicate and share their measured data.

The Network Layer

This layer receives the data from the previous layer in the form of digital signals. Then, it is transferred to the next layer via networks like WiFi, Bluetooth, and 3G.

The Middleware Layer

Next, there is an essential software layer called middleware. This layer is often described as the glue between the IoT device or object and its application. This layer allows for communication and interoperability between the devices connected to the internet (Ngu et al., 2016). It includes technologies like cloud computing. IoT has led to a huge amount of data from all the connected devices that has to be stored and processed somewhere. Cloud computing acts as a solution to this problem. Data from all different resources can be stored and processed with this model (Lee and Lee, 2015).

The Application Layer

One requirement for IoT is IoT applications. These applications allow for device-to-device and human-to-device interaction. Device-to-device applications are there to receive data and act upon that data. These applications are used for devices and processes that do not need human intervention. Human-centered IoT applications are necessary for devices that users interact directly with, like the devices in smart homes. These human-to-device applications should provide data visualization to present information to the user. All the IoT applications need to be built so the IoT devices can monitor the environment, identify problems, communicate with each other (and if necessary with the user), and potentially solve problems, with or without human intervention (Lee and Lee, 2015).

The Business Layer

This layer manages the applications and services of IoT and is responsible for all research related to IoT.

2.3 Usability and UX

2.3.1 Definitions

Usability is an important characteristic of computing systems because it is one of the main factors consumers consider when purchasing systems. According to Shackel (2009) usability should be defined as:

> *The capability in human functional terms to be used easily and effectively by the specified range of users, given specified training and user support, to fulfill the specified range of tasks within the specified range of environmental scenarios.*

Usable systems are also often referred to as user-friendly. Usability is seen as one of the elements contributing to UX. This is more broad and is defined by ISO FDIS 9241-210[1] as:

> *A person's perceptions and responses that result from the use and/or anticipated use of a product, system, or service* (Bevan, 2009).

UX takes a more holistic view and aims for the balance between task-oriented aspects of systems (like usability and utility) and non-task-oriented or hedonic aspects. These hedonic aspects are related to subjective reactions and perceptions. It can thus be said that increased usability leads to increased UX; however, the inverse is not necessarily true (Petrie and Bevan, 2009).

The usability of systems and their user interfaces is often evaluated through the ten heuristics proposed by Nielsen (2005). These heuristics can be seen as rules of thumb when designing usable systems. The first heuristic is the **visibility of system status**, which states that the system should give timely feedback to the user about what is going on in the system. The second heuristic is the **match between system and real world**, which specifies the system has to use real-world conventions. It should provide information in a natural and logical order in a way that the user understands it. The third heuristic is concerned with **user control and freedom**. This heuristic simply states that the system should support undo and redo. The user should be able to easily recover from a mistake. The fourth heuristic is about **consistency and standards** and states that the words, actions, and situations should be consistent about the platform. The user should not have to worry the same action has different outcomes in different conditions. The fifth heuristic is **error prevention**. This guideline states that systems should prevent problems from occurring as much as possible. Error-prone conditions should be deleted, and if that is not possible, confirmation should be given by the user to make sure they want to proceed with an action. The sixth heuristic is **recognition rather than recall**. This guideline is created to minimize the users' cognitive load. Users should not have to remember

[1] `https://www.iso.org/obp/ui/#iso:std:iso:9241:-210:ed-2:v1:en`

important information through an action; instead, it should be visible to them. The next heuristic is concerned with the **flexibility and efficiency of use**. This states that different methods for performing an action should be integrated into the system; this way, it is usable for both experienced and inexperienced users. The eighth heuristic is about the **aesthetic and minimalist design**. This heuristic simply states that the design of the user interface should not contain irrelevant or unuseful information. The next heuristic states that a system should **help users recognize, diagnose and recover from errors**. This means that error messages should be expressed in plain and understandable language; they should indicate the problem and provide a solution. Lastly, the heuristic **help and documentation** states that the user should easily be able to access instructions about the use of the system. According to Nielsen, meeting these heuristics in a user interface makes a system usable.

2.3.2 Usability and UX Evaluation in IoT

Even though UX is an important characteristic of computing systems, it has not been widely studied in the world of IoT systems. Many studies focus on the technological aspect of these systems, thus focusing on the first four layers of the IoT architecture proposed by Farooq et al. (2015). Research into the Application Layer, especially human-to-device applications, is limited. Research into technological aspects of IoT sometimes evaluates IoT systems based on the Quality of Service (QoS). Although QoS measures users' requirements, it focuses more on technological requirements like security, energy consumption, and cost (Asghari, Rahmani, and Javadi, 2019). Thus, there is no focus on UX or usability specifically, which is important for consumer IoT. Rowland et al. (2015) argue that consumers are the most challenging market to design IoT for. Unlike in other sectors (e.g., industry), users in the consumer market have a choice as to whether they would use the IoT system. Consumers will subconsciously try to estimate the benefit they would get from using a product compared to the cost involved in acquiring, setting up, and using it. These users also have a very low tolerance for unreliability; the product must deliver what it promised to do.

Shin (2017) proposes a model to evaluate IoT systems based on the Quality of Experience (QoE). QoE can be seen as a more holistic evaluation than UX, as it assesses the overall acceptability of an application or service as perceived by the user, taking into account users' expectations, feelings, perceptions, cognition, and satisfaction towards a system. This research underlines the importance of UX in IoT systems, as it indicates that the quality of a system is more a user-dependent concept than a device-dependent concept, meaning the users' experience, in the end, determines their satisfaction. The study suggests that the quality of a system should be assessed as a combination of technological and user-perceived quality.

The proposed model looks at the content quality (i.e., relevance, reliability, and timeliness of the content provided by a system), system quality (i.e., system performance when delivering content and meeting users' needs), and service quality (i.e., how well a system conforms to users' expectations), as they positively influence the utilitarian and hedonic performance of the IoT systems. In turn, a system's utilitarian and hedonic performance positively influence user satisfaction. Sequentially, user satisfaction should positively affect coolness and affordance of IoT, increasing users' QoE about the IoT. Affordance is a relation between the system and the user that affords the opportunity for the user to perform an action. This QoE model is depicted in Figure 2.2.

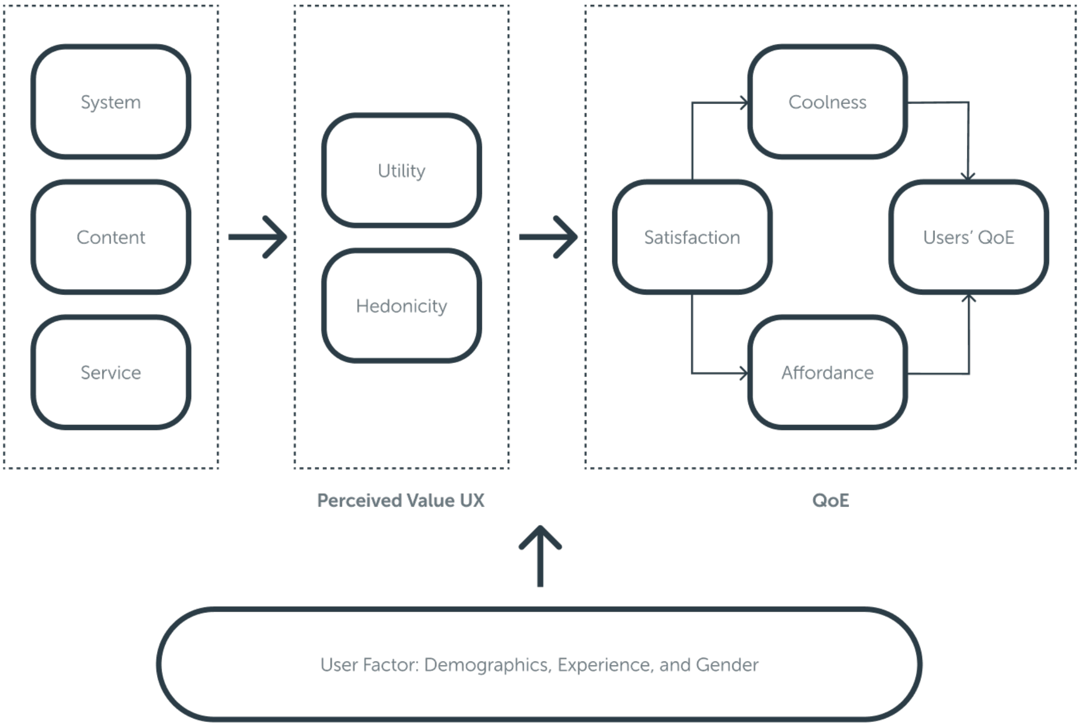


FIGURE 2.2: Proposed model by Shin (2017) to evaluate an IoT system based on QoE.

The author suggests that coolness and affordance could have an increasing effect on the usability of IoT systems since they have a positive effect on the users' perception. These results, however, could not be validated.

Although the proposed model by Shin shows promising results for evaluating the overall quality of a system, it does not provide clear usability criteria for IoT. As there are few other user-centered studies in the field of IoT, generally accepted usability criteria for evaluating IoT systems seem absent. Furthermore, a big part of the interaction in IoT systems is invisible to the user, which makes them different from other IoT systems (Bergman et al., 2018). Moreover, the interconnectivity of the devices in the IoT ecosystem and the presence of both a software and a hardware component in the devices make that the heuristics of Nielsen might not be enough to evaluate the usability of IoT systems (Fauquex et al., 2015). Thomas, Onyimbo, and Logeswaran (2016) claim that integrating usability guidelines in the development of IoT devices would ensure smooth and easy-to-use operational systems for the users and would also provide reliability assurance, which is missing at the moment. According to the authors, requirements must be considered in the design process of an IoT system, to achieve the basic usability criteria for a system. The first of those requirements are data requirements. IoT devices produce large amounts of data that must be transmitted and stored somewhere. Thus the data requirement should be an essential factor because a loss of data could interrupt the system and frustrate the user. Next, there are environmental requirements. This says it is required for the operational capabilities of the IoT system to fall within the environment in which it would be deployed. Considering the environmental factors is necessary to develop reliable systems that take into account the users' environment. When the environment is not considered, this could lead to failure, leading to discomfort and panic among users. The third consideration, as stated by the authors, are functional requirements. Considering these means, it is necessary to give a description of the functionalities of the different features in an IoT system. This provides the user with a workflow, that gives them steps to complete the desired task. Interaction between each feature in the system has to be taken into account precisely because not doing that can create functionality issues for the users. Also, the system's decision-making process can be disturbed when the interaction between the features does not work properly. Lastly, there are user requirements. These can be specified when the users of the system's interface and the environment are clearly defined. These requirements have to meet the physical and cognitive needs of the users. The authors claim that these four guidelines should be considered when designing IoT systems as they should improve the usability of those systems.

2.4 Challenges

The need for usability guidelines is demonstrated by the UX challenges users currently face when operating IoT systems. Challenges found in literature are discussed in this section.

Bergman et al. (2018) conducted a qualitative study amongst IoT-developing companies to explore how they handle UX requirements. They found that companies use agile and iterative development instead of requirements engineering to develop their products. This research also underlines that companies do not understand UX requirements overall and that there is a need to improve UX requirements elicitation and analysis methods.

In this study, various UX challenges are discussed that arise in IoT systems. A very important UX issue that often arises has to do with the interoperability of IoT devices. As described before, the devices are part of a bigger ecosystem and often have to communicate with each other. However, the various devices in the ecosystem are developed by many different companies with different goals. Because there is no standardization, issues with interoperability may arise, which has a decreasing effect on UX. These interoperability issues are an important field of research within the technological area of IoT. Due to the ever-growing complex IoT ecosystem, it is often difficult to identify where exactly the issues arise. This makes it hard to communicate to the user what is going wrong and how it should be solved. Studies are now looking into the possibilities of smart-troubleshooting. This would enable IoT systems to recognize where failures are occurring, and it would enable them to solve these failures themselves. This would increase UX, as it requires no interaction with the user (Caporuscio et al., 2020).

Not only does this lack of standardization cause interoperability issues, but it also results in users having to use different applications for managing the different IoT devices they own (Bergman et al., 2018). Consequently, users have to learn how to operate different apps. The increasing functional complexity of the IoT devices and the amount of them can result in cognitive overload (Nazari Shirehjini and Semsar, 2017). Kubitza et al. (2020) describe their vision of using software that could run independently of the specific devices supporting various brands. As described in Section 1.1, *ActPad* tries to deal with this problem by creating a smart desk platform to interact with IoT devices close to the desk (Delgado Rodriguez et al., 2021).

Other challenges described by Bergman et al. (2018) are connectivity challenges. This occurs when an issue arises in the network layer. Communication between the software and hardware component of the system can not take place when this happens. This could also cause problems with the communication between the device and the IoT network. However, this UX challenge can not be solved by the company because it is caused by external factors.

The last UX challenge discussed in the study is the ease of the IoT system. Companies describe that services are often made too advanced. They argue that installation and management of the system have to be easy. This challenge relates to an issue arising in the development of Consumer Electronic Products, which are household products connected to IoT. It has been shown that those products often have too many functions that are difficult to learn. Users are sometimes not even aware that some functions exist and never use them as a result (Benvenuti et al., 2021).

Chuang, Chen, and Liu (2018) argue that IoT systems rarely provide sufficient feed-forwards and feedback to indicate to the user what their current status is and what actions they are going to perform. As a result, people often do not understand why the system does not work as expected making users question the system's reliability and smartness.

Research by Resnick (2013) described the UX challenges found in ubiquitous computing. As described in Section 2.1, the interaction between these systems and the user is limited. This makes the systems almost invisible. However, it also creates UX challenges. One described challenge occurs when the system misunderstands the user's intentions or when the user has the wrong idea of what a system can do. This also relates to the problem described by Chuang, Chen, and Liu (2018). The research proposes that a process for error detection and error recovery is needed with minimal user interaction. The author of this paper also discusses another challenge in the installation and management of the system. The set-up and modification of the action rules do not only have to be easy (as described before), but they also have to be transparent. In a well-functioning system, the rules do not often have to be modified because data is gathered from the IoT network and the sensors. However, the action rules have to be accessible to the user. The actions taken by the system should also be reversible.

One more UX problem arises in the area of Smart Homes. Systems are currently often designed to serve the needs of individual users. Problems arise when these systems are integrated into places with multiple users, like family homes. Systems that give personal recommendations to individual users might fail in multi-user places. In the worst-case scenario, these systems might even cause the family members to get isolated from one another. Multi-user systems thus have to focus on multi-user recommendations (Eggen, Hoven, and Terken, 2017).

The last UX issues, which is extensively discussed in research, are the privacy and security issues that arise in IoT systems. IoT is seen as a vulnerable point for cyber attacks due to weak security protocols and policies. This increases the chances of a data breach. The reason for these issues is that the interconnectivity of the IoT network enables anonymous and untrusted users to access it. Several models and tools are already designed to account for this issue; however, users are still skeptical about the security of their IoT devices (Tawalbeh et al., 2020).

Chapter 3

Prestudy: Identifying UX Challenges

In this section, the prestudy is described. This study was conducted to confirm the UX challenges users face when interacting with an IoT device as found in literature (See Section 2.4) and discover more challenges if they were present. First, the study design of this study is discussed, which is followed by the results of this study.

3.1 Study Design

In this section, the study design of the prestudy is discussed. This concerns the chosen methods, the recruited participants, the materials, the procedure, and the measures.

3.1.1 Mixed Methods

Two methods were chosen to gather both qualitative and quantitative data about UX challenges in IoT systems.

The first of these methods was usability testing. In this usability test, the researcher watched a participant perform predefined tasks with the system in a specified test environment. The goal here was to discover the studied system's usability problems and find a way to reduce those problems, which is defined as formative usability testing. However, task performance was also measured, which is defined as summative usability testing. This task performance data was collected to be later compared to the task performance data in the main study. It was measured whether participants were able to perform tasks without intervention and how much effort it took to perform the tasks. A combination of formative and summative usability testing was thus applied in current research (Lewis, 2014). During the tests, the participants were asked to think aloud. This gave insight into the thought process of the participants.

The other method that was used were interviews. Interviews were conducted after the participants had finished the predefined tasks from usability testing. Interviews are useful for gathering qualitative data about users' experiences and challenges. The advantage of interviews is that they are very flexible; they allow the interviewer to explore subjective aspects and perceptions of the interviewee. Interviews can be conducted in one of three ways: unstructured, semi-structured, and structured. Semi-structured interviews were conducted for this study, meaning the questions for the interview were predefined. During the evaluation, the semi-structured interview methodology allowed the interviewer to dig deeper into the responses from the participants and ask additional questions. The semi-structured methodology is thus a mixture of the structured and the unstructured methodology (Blandford, Furniss, and Makri, 2016).

3.1.2 Independent Variable

The independent variable is the graphical user interface design of the studied inlite system, which is discussed in more detail in section 3.1.5. For the prestudy, the existing app design of this IoT system was tested to gather the usability and UX challenges users face when interacting with the system. After applying the design criteria to the design, the adapted design was tested in the main study, and the results of both studies were compared.

3.1.3 Dependent Variables

The first dependent variable was the task performance of the participants for each usability task.

The second dependent variable consisted of the results of the System Usability Scale (SUS) (Brooke et al., 1996), which gave a quantitative insight into the system's usability. The scale is widely used because it is technology agnostic and it is quick. It is easy for both experimenter and participants, and because it provides a single score that is easily understood (Bangor, Kortum, and Miller, 2008). Research into questionnaire's the psychometric properties (reliability, validity, and sensitivity) has always been positive, meaning SUS is an appropriate questionnaire for measuring usability. The questionnaire consists of ten items with an alternating positive and negative tone. The participant indicates on a 5-point Likert Scale how much they agree to an item. This gives a score on a range from 0 (poor usability) to 100 (excellent usability) (Lewis, 2018). It is important to note that the SUS score is not a percentage. The average SUS score is 68; scoring 68 means the system scores higher than 50% of all tested systems and applications. A score above 80.3 means the system is in the top 10% of all tested systems and applications. The SUS scores can also be converted to letter grades (Sauro, 2011). The SUS can be found in Appendix A.

Another questionnaire was used to assess UX, namely the AttrakDiff (Hassenzahl, Burmester, and Koller, 2003), which was the third dependent variable. AttrakDiff is a questionnaire that can be used to understand how users rate the usability and the design of an interactive product[1]. The questionnaire is used often and measures both the pragmatic and hedonic quality of systems, hence assessing the user experience. The questionnaire consists of 28 7-point bipolar items in the form of opposite words. The 28 items measure four UX dimensions, namely (1) *pragmatic quality*, (2) *hedonic quality - identity*, (3) hedonic *quality - stimulation*, and (4) *attractiveness* (Walsh et al., 2014). The dimension *pragmatic quality* measures usability aspects like ease and effectiveness of use. The second and third dimensions together form the *hedonic quality* dimension. The *hedonic quality* measures aspects not directly related to the tasks the user wants to accomplish; examples are originality and beauty. The *hedonic quality* is split into *identity* and *stimulation*. The *identity* focuses on the human need to be perceived by others in a certain way. This dimension measures how well a product supports the communication of a desired identity. The third dimension, *stimulation*, focuses on the human need for personal development, which can also be supported by a system. The fourth dimension, *attractiveness*, measures how attractive the user interface is to the user. It is assumed that the *pragmatic quality* and the *hedonic quality* influence the perceived *attractiveness* of a user interface

[1] http://www.attrakdiff.de/index-en.html

(Schrepp, Held, and Laugwitz, 2006). Each dimension gets a score between -3 and 3. The higher the score, the better the system scores on each dimension. The AttrakDiff questionnaire is shown in Appendix B.

The last dependent variable was formed by the results of the interviews.

3.1.4 Participants

A total of ten participants were gathered through convenience sampling (five female, five male). Criteria for participating were that the participant was at least 18 years old and that they had never used the in-lite system before. The participants were between 20 and 29 years old (M = 24.8, SD=1.44), and they all were Dutch native speakers. When asked to rate their tech-savviness on a scale of one to ten, the participants all gave a score between six and nine (M = 7.6, SD=0.97). Besides, seven participants had a tech-related job or followed a tech-related study. Four participants indicated to have had experience with IoT systems before. Lastly, five participants were iOS users, and five were Android users.

3.1.5 Materials

Survey

A survey was created using Qualtrics XM. This survey consisted of several pages. The first was an information page; this page was meant to inform the participant about the research. This was done to prepare the participant for the upcoming 30 minutes and to put them at ease. The second page consisted of a consent form. Participants had to tick the boxes they agreed to and sign the form with their names. They were told that their name would be pseudonymized for publication and was only needed for consent. Furthermore, if the participant did not tick all the boxes, they were excluded from the experiment. The following page was a demographics questionnaire. These demographics could later be used to comment on the generalizability of the results of the study. This part of the survey is shown in Appendix C. After the demographics questionnaire, a page was shown in the survey that told the participant to turn to the researcher for usability testing. The page requested the participant not to close the survey yet, because they would need it after the tasks. The SUS and the AttrakDiff were integrated into the last two pages of the survey; these are shown in Appendix A and Appendix B, respectively.

in-lite System

The IoT system that was used was a smart lighting system from in-lite. This is a system that is specifically designed for lighting up gardens. The user can buy smart lights and place them in their garden. These smart lights are powered by smart hubs. A smart hub has three lighting zones to which multiple smart lights can be connected. The whole zone can be controlled, or the lights can be controlled separately. Controlling means turning a zone or lights on and off or adjusting their brightness. It is also possible to add accessories to the ecosystem, for example, a motion detector. Controlling the hubs, the lights, and the accessories is done through the app. A setup was created containing several elements of the system. This setup is shown in Figure 3.1. For the current study, two smart hubs were used. One of these smart hubs had lights connected to only two lighting zones. One of these zones had one smart light connected to it, and the other zone had three smart lights connected to it. The second smart hub had one light connected to one lighting zone. In the current research, one accessory was used, namely a motion detector. A routine can be set such that one zone or multiple zones turn on when motion is detected. A magnet had to be used to reset the smart lights in the system.

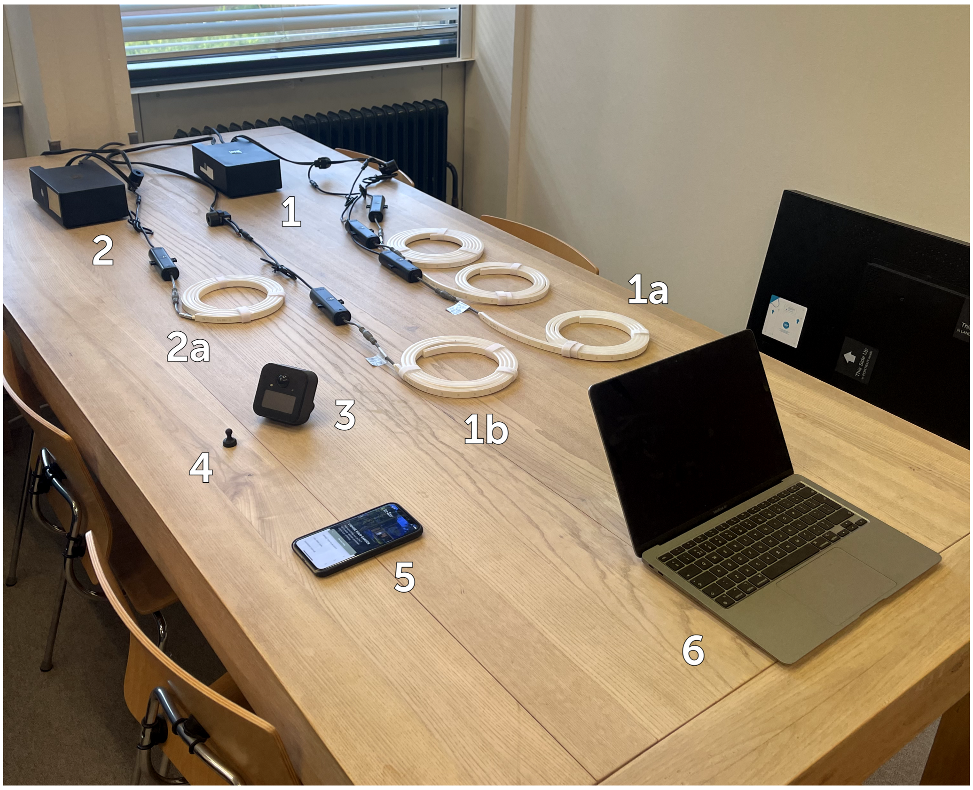


FIGURE 3.1: The in-lite set-up during the study[1].

Tasks

The participants were given a set of twelve tasks. These tasks were given in an imaginary scenario. The scenario and the tasks are shown in Figure 3.2. Some tasks consisted of more actions than others. Task four consisted of making the smart hub discoverable and connecting it afterward. Task seven was making sure all lights were found and resetting them afterward. Lastly, task eleven consisted of finding the place where to add the motion detector and connecting it afterward.

In a pilot study with two participants, ten tasks were given. A list of these tasks is given in Appendix D. It was found that one task could be excluded because it would not help to discover UX challenges relating to IoT systems. The pilot study also revealed that adding three more tasks would give a deeper insight into the UX challenges in the system.

[1]The image shows (1) the first hub with (1a) three lights connected to one lighting zone and (1b) one light connected to another lighting zone. It shows (2) the second smart hub with (2a) one light connected to one lighting zone. It also shows (3) the motion detector and (4) the magnet, which should be used to reset the smart lights. It also shows (5) the app which was used to operate the system. The laptop (6) was used by the researcher to make notes during the usability test and interview.

Imagine you just moved houses, and you decide to use the in-lite system in the garden of your new house, as you did for your old one. The system consists of lights that can be connected to smart hubs. The devices in the system are all operated using an app. Other accessories can also be added to the system, such as a motion detector.

To light up the garden from your previous house you used this light strip, connected to this smart hub. You moved this hub and light strip to your new house. However, the garden from your new house is much bigger, so you decided to get another hub and some lights. You bought this hub and four light strips. A hub contains three lighting zones and as you can see, your old hub is using only one zone. Your new hub is using two lighting zones, three light strips are together installed on one lighting zones and the last light strip is installed on another lighting zone.

1. Your first task is to create a new garden in the app and to add the new smart hub to this garden.
2. Your next task is to set up the first lighting zone of this smart hub, which is the zone that contains three light strips. You will use these strips to light up the swimming pool you have in your new backyard.
3. You now want to set up the second lighting zone of this hub. You will use the strip in this zone to light up your terrace.
4. Now you want to add your old smart hub to the garden, which you have to make discoverable in order to find it.
5. Set up the lighting zone from this smart hub, you will use it to light up the path in your garden.
6. Next, you want to synchronize the garden.
7. You want to make sure all the lights are found and make the ones that are not found discoverable.
8. Adjust the brightness of one of the lights in the zone of your swimming pool.
9. Now adjust the brightness of all the lighting strips in that zone at once.
10. You now want to add a motion detector to your system. Find the manual for this device in the app.
11. Now add the motion detector to your garden.
12. Lastly, you want to set a routine such that the light that lights up your path responds to movement.

FIGURE 3.2: Scenario and tasks for usability testing.

The final tasks were all carefully defined to give insight into the system's usability. These tasks give insight into the usability because they cover a large interaction space of the system. They introduce the participants to the most features the system has to offer. Features of the app that are not explored by the participants are features that are not relevant for the interaction with the devices in the ecosystem. Examples of such features are the setting of account details or the option to share a garden with other users. Furthermore, each task also tests whether some of Nielsen's heuristics were considered in the design of the system (see Section 2.3.1). Table 3.1 shows which heuristic each task was testing.

Task	Nielsen's heuristic
1.	Visibility of the system status
	Match between system and the real world
2.	Visibility of the system status
	Match between system and the real world
3.	Visibility of the system status
	Match between system and the real world
4.	Visibility of the system status
	Consistency and standards
	Match between system and the real world
5.	Visibility of the system status
	Match between system and the real world
	Consistency and standards
6.	User control and freedom
7.	Visibility of the system status
	Consistency and standards
	User control and freedom
8.	Visibility of the system status
	Match between system and the real world
9.	Visibility of the system status
	Match between system and the real world
10.	Help and documentation
11.	Visibility of the system status
	Consistency and standards
	Match between system and the real world
	User control and freedom
12.	Visibility of the system status

TABLE 3.1: Nielsen's heuristics tested by each task.

Interview Questions

Interview questions were predefined and asked to get a deeper insight into the participants' experiences and opinions about the in-lite system. A total of 11 questions were carefully created whilst keeping Nielsen's heuristics and the related works in mind. The goal of the interviews was to find out whether the challenges as described in Section 2.4 were also found in this system and whether more challenges could be defined. The questions asked are shown in Figure 3.3. The first question was asked to gather insight into the user's prior experience with IoT systems, which could influence their task performance. The second question would give an insight into the users' overall opinion of the system. Questions three and four give insight into the perceptions formed by the participants towards various features of the system, which is an important aspect of user experience (see Section 2.3.1). Questions five and six indicate whether the participant found using the system easy and effective, which thus gives qualitative insight into the usability of the system. Questions seven, eight, and nine also give insight into a combination of usability and user experience of the system. For example, when a part of the system is unclear to the user, it makes the system less easy and effective to use, and it evokes a feeling of confusion. Question ten is asked to test to find out whether Nielsen's heuristic *aesthetic and minimalist design* is met. The last question is asked to get a qualitative insight into the *hedonic quality - identity* of the system because it would indicate whether the user wants to be associated with the system or not.

1. Have you used a similar/IoT-system before? And if so, was using that similar to using this system?
2. How would you describe the overall experience you just had with the system?
3. What do you like about the app of the system? And about the device?
4. What do you think could be improved about the app of the system? And about the device?
5. What tasks did you find easy to perform?
6. What tasks did you find harder to perform?
7. Was there anything that confused you when using the system, and if yes, what was it?
8. Was there anything that caused frustration when using the system, and if yes, what was it?
9. Did you feel rushed when using the system and if yes, why?
10. Do you like the way the app looks? Is it appealing to you?
11. Would you recommend using this system to anyone?

FIGURE 3.3: Interview questions.

Other Materials

A smartphone was needed to run the app of the system. It was considered to let the participants use their own phones; however, this would mean that participants had to download the app and create an account. It was decided that this would burden the participants too much. An iPhone X was used to run the in-lite app. The decision was made to only use an iOS device and not an Android device because there were slight differences in the app for these devices. Also, the iPhone enabled the researcher to easily record the screen during the tasks. The iPhone is also shown in Figure 3.1.

Furthermore, a device was needed to record the usability tests and the interviews; this was done with an iPhone as well. Participants were asked to bring their own laptops so they could fill out the survey on there.

3.1.6 Procedure

When the participant arrived, they were thanked for their time and effort. An introduction was given about the research, and afterward the participant was asked to read the information page in the survey and to fill out the consent form and the demographics questionnaire. When they were finished, the participant was asked to put their laptop aside but not close the survey yet.

Subsequently, usability testing started. The audio recording was started, and the screen recording of the iPhone X was activated. The participant was given the iPhone X, and the first task was cited to them. The next task was cited after the participant had finished the previous one. The order of tasks was the same for every participant. If the participant was stuck, a hint was given to help the participant along.

After the participant had finished all tasks, which took about 10 minutes, they were asked to return to the survey. The audio recording and screen recording was stopped. Then, the participant was asked to fill out the SUS first and the AttrakDiff second.

After the participant had completed the survey, an interview was conducted. The audio recording was started again, and the interview questions were asked in a predefined order. However, the semi-structured nature of the interview allowed the researcher to deviate from the predefined questions and ask additional ones. Conducting the interview took about 10 minutes as well.

Lastly, the participants were thanked for their time and participation again. They were also told that they could send a message if they had additional questions. A visualization of the procedure is shown in Figure 3.4.

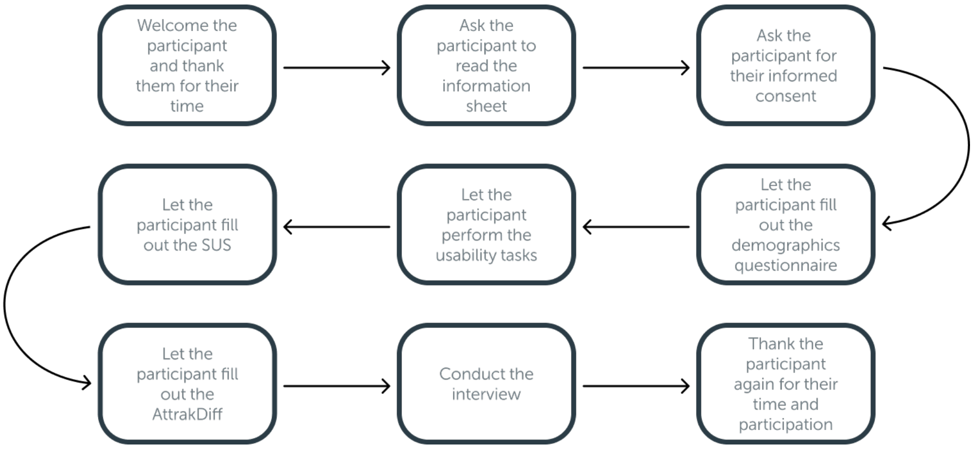


FIGURE 3.4: Procedure of the study.

3.1.7 Measures

During the study, several measures were done. Through the demographics survey described in Section 3.1.5, properties of the participants were measured, namely their age, gender, occupation or study, experience with in-lite or other IoT systems, tech-savviness, color blindness, and the mobile operating system they use. The full questionnaire can be found in Appendix C. With the specified usability tasks, quantitative data was gathered about the performance of the participants. It was measured whether participants were able to perform each task without support and the screen recording gave insight into the time and steps the participants took to finish each task. The SUS and the AttrakDiff gave numerical data about the usability of (different aspects) of the in-lite system. Finally, the interviews allowed for qualitative measures of the experiences and opinions of the participants about the system.

3.2 Quantitative Results

In this section, the task performance results are described, and the results of the SUS and the AttrakDiff questionnaire are given.

3.2.1 Task Performance

Task 1

The first task was done correctly by all ten participants. Some interesting remarks were made by the participants during the task. When searching for smart hubs in the app, only the correct smart hub appeared in the list. The other one did not appear because it had to be reset first. There were seven of the ten participants that wondered whether the right smart hub appeared in the list; this would be confirmed when connecting to the found smart hub. When connecting to the smart hub, the reset button started to flash shortly and quickly, and the app asked the participant whether they saw this. Only one of the ten participants saw the button blinking; the other nine tried again by going back to the list of smart hubs and connecting again. After the second try, they saw the reset button blinking and proceeded to the next step.

Task 2

The second task was done correctly by all ten participants.

Task 3

The third task was done correctly by nine participants. However, one of the participants thought they had to save their settings of the first lighting zone before setting up the second one. By doing this, they finished the setup process of this hub and they could set up another zone. The participants thus had to restart the setup process.

Task 4

The fourth task was connecting to the second smart hub. This one had to be reset first because, in the scenario, the participant had used it before in another garden. Only one of the participants realized they had to reset the hub. When telling the other nine participants that they had to reset it, six of them did this correctly. The other three needed some help resetting the hub. All participants connected to the smart hub correctly afterward.

Task 5

The fifth task was done correctly by all ten participants.

Task 6

The sixth task was done correctly by all eight participants. However, none of the participants understood why this was needed or what the result of doing this was (*note: only eight participants participated in this task because it was added after the pilot*).

Task 7

There were two steps in this task. The first was realizing that the lights had to be reset, and the other step was resetting them. None of the participants realized that the lights had to be reset. When told that the lights had to be reset, none of the participants knew how to. When told that they had to use the magnet, six of the ten participants reset the lights correctly.

Task 8

This task was done correctly by six of the eight participants. The other two participants took a long time to figure out how to perform this task (*note: only eight participants participated in this task because it was added after the pilot*).

Task 9

This task was done correctly by all ten participants.

Task 10

This task was done correctly by all ten participants.

Task 11

This task consisted of two steps. The first was finding where to connect the motion detector in the app. The second step was connecting it. The first step was done correctly by seven of the ten participants. The other three participants took some time to find the correct place to connect the motion detector. The second step was done correctly by six of the ten participants. The other four participants did not take the correct steps that were displayed in the app to connect the motion detector.

Task 12

This step was done correctly by all ten participants.

3.2.2 SUS

The SUS score given by the participants was relatively high ($M = 77.50, SD = 10.99$). However, it did not meet the top 10% of tested systems yet. The data was also tested for normality with a Shapiro-Wilk test. The data distribution would later be important when comparing the untouched version of the system to the improved version of the system. An alpha level of 0.05 was assumed. It was concluded that the data was normally distributed, namely $W(10) = 0.98$, $p = 0.98$, so $p > \alpha$.

3.2.3 AttrakDiff

In Figure 3.5 the results from the AttrakDiff questionnaire are shown. The mean for *hedonic quality - stimulation* is the lowest (M = 0.47, SD = 0.12), followed by the *pragmatic quality* (M = 1.10, SD = 0.45). The mean for *hedonic quality - identity* and *attractiveness* are quite close together (M = 1.36, SD = 0.52; M = 1.44, SD = 0.12 respectively). The AttrakDiff results are particularly interesting for later use to compare those results to the AttrakDiff results that are gathered after testing the improved app. It could be said, however, that AttrakDiff scores are already quite high. The graph shows a line right from the center, which is a positive score.

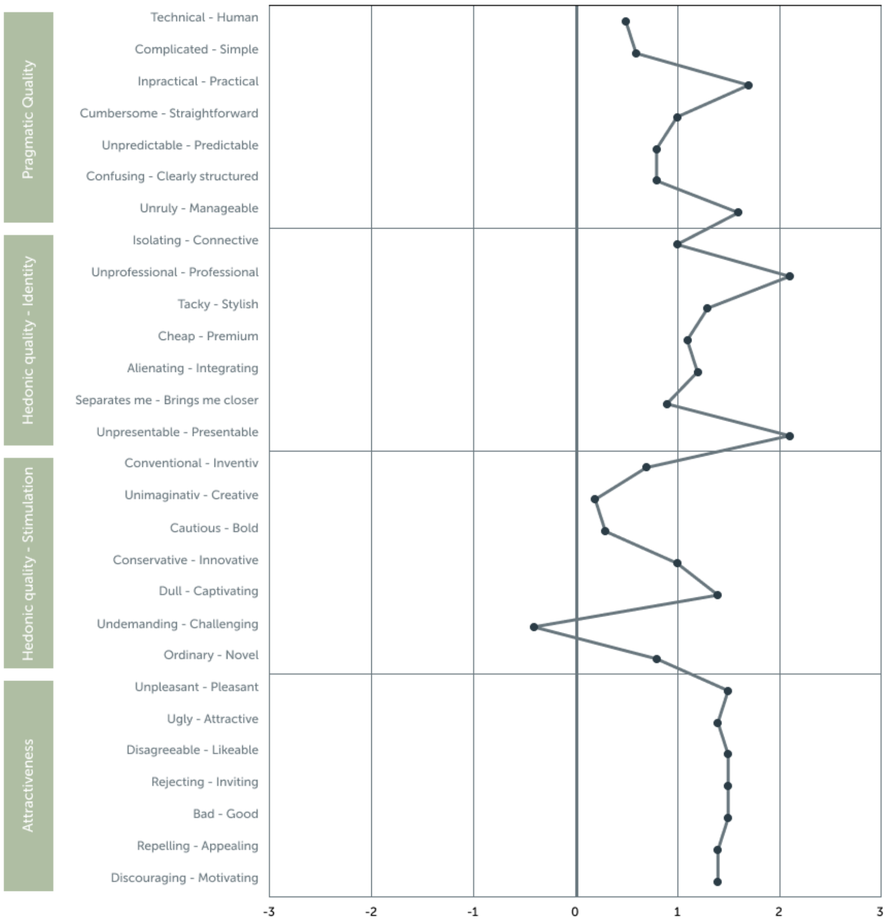


FIGURE 3.5: AttrakDiff results of the prestudy.

3.3 Qualitative Results

The results of the interview were analyzed through thematic analysis. This is a method for analyzing qualitative data. Through this method, patterns can be identified, analyzed, and interpreted. Codes are given to interesting features of the data which are relevant to the research question. Codes together can form themes (Clarke, Braun, and Hayfield, 2015). The analysis was done via ATLAS.ti version 8.4.3. Three interviews of the nine interviews were randomly chosen to generate codes. After that, the rest of the interviews were coded with the codes that were generated on the first three interviews. These codes were later categorized into four themes. Each theme is discussed below.

3.3.1 Ease of Use

This theme describes which elements in the app made using the system easy and efficient. The general feeling the participants had after using the system was that it was intuitive and easy to use, especially after the onboarding. So after connecting the devices, the participants felt that it was easy to control them:

> *The most easy part was after the set up of the lights. After that it is very easy and pleasant to control the lights. (P1)*

The users also felt that it was easy to distinguish the hubs during the onboarding. The reset button of the selected hub would start blinking, so the user could see which one was selected. Users also felt that when the devices were made discoverable, connecting to them worked quickly and easy:

> *Connecting works very easy, which is not always the case. It is very easy to see whether you have the right hub. That is nice about the system, the hardware works and connecting to the software works as well.*

Users felt that interaction with part of the devices worked intuitively and as expected. During the experiments, users were asked to connect to a smart hub which was not discoverable yet. Although users felt that there should be more explanation as to why the hub was not discoverable (see Section 3.3.3), part of them found that process of resetting the hub worked intuitively:

> *Resetting the hub worked very quickly and in one try. You just had to press the button for a long time like with most phones and a PlayStation. If you have to reset those devices, you usually have to press the reset button for a few seconds. I think that is some sort of safety, so that was clear and intuitive for me.*

3.3.2 Design

This theme describes what participants felt about the way the app of the system was designed. This concerned both aesthetics and the structure of the app. With the structure, it is meant where the different buttons, functions, and actions can be found in the app. Generally, the users felt that the app was structured in an intuitive way, meaning they could find the functionalities without having to think too much about finding them. Even though the app contained a lot of functionalities, the users felt that it was structured and clear:

> *I experienced using the system as good, intuitive and easy. When I saw this system I thought using it was going to be hard, but actually it was quite simple. I think a manual would probably distract from how easy it is to use the system (...) It is intuitive where you have to click in the app. For example to operate a certain light, you just have to tap that light in the app. (P5)*

> *The app is very clean, with a lot of functionalities. However, I do not feel like it is overwhelming. It is simple, yet functional. (P9)*

To ease the interaction between the user and the devices, the app uses pictures of the devices to indicate to users which device they should focus on. Users felt that this was an easy way to distinguish the devices and that because of this, they knew which button to press or which device to look at.

> *I found the pictures very helpful. You think 'I have to press a button', but you do not have to look for the button yourself, you just see in the app where the button is. That is nice. (P5)*

However, some users felt like the pictures were not always in line with the action that had to be taken. An example of this is when connecting with the motion detector. The reset button on the motion detector had to be pressed for 5 seconds and then once shortly afterward; however, the picture only displayed that the reset button had to be pressed once. This led to mistakes because users did not always read the textual explanation and only looked at the image.

> *I do not really read what the app says in the texts, so I pressed the button of the motion detector quickly and did not see I have to press it for 5 seconds, because the picture only shows a hand and the button. So I thought 'just press once and then it turns on'. (P5)*

Overall, the simplicity of the design combined with the high functionality was seen as a positive feature of the app. The simplicity of the app and the colors used were seen as aesthetically pleasing and appealing. Users felt that there could be some improvements in the consistency between the pictures and textual explanations that are displayed. However, users felt there were also actions for which explanations were missing and should still be added. This will be described in Section 3.3.3,

3.3.3 Improving UX

A recurrent theme in the qualitative data describes features of the app that have a negative effect on the UX of the IoT system; this especially concerns features and flows in the app that slows the user down in fulfilling an action and thus reaching their goal. These features would need to be improved in order to improve the UX of the system. Many participants felt that the app was unclear because there was a lack of feedback about the progress of a task or the state of the devices. This sometimes resulted in a perceived sense of confusion among the users, which is not beneficial for a good UX. An example of the lack of feedback about the progress of the app occurs during the synchronization. This process takes about 2 minutes, but users felt that it was unclear what was happening exactly and how far the system was in the process of synchronizing:

> *I thought the synchronization took redundantly long. You saw the lights flashing a bit, but that was done in like 10 seconds. So I have no idea what it was doing for the rest of the 1 minute and 50 seconds. It was also annoying that you were not able to do anything else in the app during the process. (P8)*

A lack of feedback about the state of the devices was found in the process of selecting the independent lights in the app. Some users wanted to adapt the names of the independent lights to make a clearer distinction between them, but this process was perceived as cumbersome and slowed down the user flow. However, when selecting a light in the app to adapt it, it was not clear which light was selected exactly. Participants suggest that it would be easier if the selected light would give a pulse at the moment of selection. This way, the user does not have to use cumbersome ways to find out which light is selected:

> *Identifying and naming an individual light was hard. To do that I had to go all the way to the settings, in which I gave each light a number, then I had to go back to the home-screen to see which number light was on and after that I had to go back to settings to name them. That whole process, instead of when you click on a light, that it starts blinking or give a pulse. (P8)*

More unclarity was perceived because users felt the app did not explain certain tasks clearly. This feeling was perceived especially when problems arose, for example, when devices could not be found and could thus not be connected to. They did not give enough support to the user to operate the devices when something went wrong. Users had to have prior knowledge about the system or figure out themselves how to solve the problems in order to perform some tasks:

> *The happy flow of the app is good, quick and nice. But as soon as there is a problem it becomes harder and confusing to use, you feel like 'what should I do now' and 'what is going on, why is it not working', and the app does not really help out with that.*

The clearest example of such a task was the resetting of the lights. Not all lights had to be reset, but part of the lights was not discoverable and thus had to be reset. All participants struggled with this task. Users felt that it was not clear that they had to perform this task and how it should be performed because this was not explained anywhere. Besides that, the method of resetting, for which a magnet had to be used, was perceived as not intuitive and unpredictable:

> *I would have never came up with the fact that I had to use the magnet for resetting the lights. I looked for a button first, I did see the icon of the magnet on the light, but I did not realize what that was for. For that you should grab the manual. Or it could be indicated in the app, something like: 'The lights should be reset, get your magnet'. (P9)*

All in all, the users thus felt that the app was lacking in support for interaction with the devices when problems arose, i.e., it should offer more help with troubleshooting. There was not enough support when the users felt stuck and when they wanted to know which steps to take next.

Another feature of the system that resulted in a perceived sense of confusion was the differentiating method of interaction for each device. During the test, the users were asked to connect and interact with three different devices, namely the smart hubs, the smart lights, and the motion detector. For each of these devices, there was a different way of making them discoverable. This resulted in confusion for the users because there was not enough consistency:

> *This magnet is a disadvantage, you can lose that easily, which is inconvenient. If you have to reset the lights you have to remember where you left the thing. It is also weird to have a button for the smart hub and a magnet for the lights. (P5)*

Users felt that there were also some points in which the system did not respond fluently or correctly to the input they gave in the app, so there was a mismatch between the app and the system. An example of this was when using the brightness slider of the lights in the app, the lights would respond jerky. Another such experience was described by P7 when interacting with the motion detector:

> *When connecting the motion sensor, the app asked me 'do you see the red light on the sensor blinking', it blinked so I clicked yes. But still it could not be found and I had to restart the process again. (P7)*

Lastly, there was a point at the system in which the participants felt rushed to perform an action. When connecting to the smart hub, the light on the smart hub was blinking so the user could see the right smart hub was selected. However, the light was only blinking for a few seconds, so after the participants had read the instructions in the app, the blinking had already stopped. They mostly had to try again and look immediately; then, they would see the blinking button. The participants thus felt that the system did not take into account that they had to read the text in the app:

> *I: Did you feel rushed when using the system?*
>
> *P3: Yes, the lights of the hub only went on for a short period of time, so you had to look very quickly. I can imagine that when you are not that close to the hub, this could give you even more of a rushed feeling.*

3.3.4 Utility

The last theme that was extensively discussed was the utility of the system; this concerns the usefulness of the system and whether it does what it is made to do. Generally, users were positive about the utility of the system. They felt the system did exactly what it was made for, namely lighting up gardens:

> *First of all, it is a fun system, it is a nice way to light up your garden and it is also nice that they app works the way it should. (P7)*

Users felt that the three different devices they had to interact with interoperated well with each other. There were no connection issues and the hardware and software worked in good harmony together. Users felt that it was a useful feature that devices could be operated all together but also could be operated individually:

> *I find it easy and nice that you can decide to operate each thing individually or altogether. You can, after installing something individually, operate the whole zone. So if you want one very bright and the other not, you can do that at the beginning of the evening. Later in the evening you can put them brighter all together. Nice. (P4)*

Generally, users also found it useful that they could operate the lights from a distance and that they could just use their phone for this. They liked that they would not need an extra device to operate the system:

> *If you compare it to Christmas lights for example, it is nice that you can just operate these lights from a distance. It is better than having a separate computer screen on which you have to do all these weird things. (P5)*

All in all, users were very positive about the utility of the system. There was, however, one participant that made an interesting remark about the utility of IoT systems in general:

> *I am not a great fan of IoT stuff. I do not think my washing machine should know what my shower is doing. I do not think everything should always be on the internet. If you can do this with an offline-system, without internet, it would be just as fine for me. Just with a thing in the wall that is in direct contact with your lights (...) What is the problem of just walking up to a device and operating it from there? Why would you need to be able to control the lights from a 20 kilometers distance? That is my whole issue with IoT. I do not need to put my dishwasher on when I am on a holiday. (P6)*

Chapter 4

Design Criteria for IoT

In this chapter, design criteria for the user interface of IoT systems are derived based on the results of the prestudy as described in Chapter 3. It is also discussed how these criteria were applied to the IoT lighting system.

4.1 Derived Design Criteria

Design criteria for an IoT ecosystem can be derived based on the results of the prestudy. The design criteria were mainly based on the interviews that were conducted with the participants and the themes that were derived from them. The task performance was also taken into consideration while formulating the design criteria. These design criteria should be considered on top of Nielsen's heuristics as described in Section 2.3.1. These are the heuristics that should always be used as a rule of thumb for designing the user interface of usable systems. When these systems become IoT, it is proposed that the following criteria should be considered as well.

4.1.1 Troubleshooting

The first important feature that should be added to the user interface of IoT devices is troubleshooting. This is important when the devices in the ecosystem do not behave as desired. Task performance for tasks four and seven in the prestudy (see Section 3.2.1) suggested that participants struggled to connect to the devices in the ecosystem that did not behave as the user expected. This also relates to the theme *Improving UX* as described in Section 3.3.3. When analyzing the quotes related to this theme, it can be concluded that users would highly benefit from troubleshooting. Users often felt uncertain about which action to take when devices could not be connected to, due to unknown problems. Evidently, the system can not always automatically know that there are problems arising because the problem could be that a device can not be connected. Therefore, problems with the hardware can not always be noticed by the system. However, the software should offer help to solve the problems that could arise in the system. The system's user interface should thus check with the user whether the interaction with the devices is going as desired and should give the user an option to start troubleshooting. This relates to Nielsen's heuristic *help users recognize, diagnose and recover from errors*. The difference between the heuristic and the current design guideline is that help should also be offered when the system does not detect an error itself.

4.1.2 Minimize User Interaction

Participants were confused by the synchronization process due to a lack of feedback from the system about what was going on. Participants had to make a choice whether they wanted to synchronize their garden or not. As discussed in Section 3.2.1, all participants chose the right action. However, they all uttered their confusion about making a choice. By automating this process and therefore minimizing user interaction, the sense of confusion could be diminished because there is no choice to be made by the user. This would also be in line with the related works that are discussed in Section 2.4. Research by Benvenuti et al., 2021 suggested that IoT systems are often made too advanced and minimizing user interaction could make a system easier to use. Other research by Chuang, Chen, and Liu, 2018 suggested that set-up has to be easy and transparent.

4.1.3 Clear Distinction Between Devices

It is not always the case that an IoT ecosystem contains different devices. When it does, however, a clear distinction should be made in the user interface between the different devices in the ecosystem. There are two use cases in which the distinction is important. The first is when the system asks the user to perform certain tasks with one of the devices. It should then be clear to the user which device they should operate. This is discussed in the theme *Design* (see Section 3.3.2). Participants felt that the current system clearly indicates which device should be operated by adding pictures of the devices for a certain task.

Another use case in which the distinction of the different devices is important is when the user selects the devices themselves. It should be clear to the user which device is selected at that moment. If this is not the case, the user can get confused about which device they are operating, and this also increases the chance of errors being made by the user. The participants of the experiment described that they missed this feature when selecting individual lights in the app, as described in the theme *Improving UX* (see Section 3.3.3). The system did not give feedback about which light they had selected, which led to confusion. The participants also indicated that the flashing reset button when selecting a smart hub was enough feedback for them to know which hub they had selected, which was discussed in the theme *Ease of Use* (see Section 3.3.1).

4.1.4 Consistent Interaction With Devices

In line with the aforementioned criterion is the consistency in interaction with the different devices in the ecosystem. The three different devices that had to be operated all had to be made discoverable in varying ways. The hub had to be reset by holding the reset button for a few seconds, the lights had to be reset by holding a magnet on a magnetic area and the motion detector had to be reset by holding the reset button for five seconds and pressing it once again. The qualitative data showed that the different methods for interaction result in confusion for the participants, as discussed in the theme *Improving UX* (see Section 3.3.3). Preferably, the method of interaction with the devices is intuitive, meaning the user can use the knowledge they gained from interacting with other devices for the interaction with the current device. Preferably the user can interact with the device non-consciously and without much effort.

4.1.5 Interaction With Multiple Devices at Once

This criterion builds on the previous two criteria that concern the multiplicity of devices in an IoT ecosystem. The user interface of the IoT system must take into account that the user sometimes has to interact with multiple devices at once. For example, the user has to read the text in the graphical user interface of the system before being able to see whether a light on one of the devices might be blinking. Therefore, the system has to give the user enough time to read the text and look at the other device. Not doing this might lead to confusion and mistakes for the user as well as a delay in the user flow (see Section 3.2.1 and Section 3.3.3).

4.1.6 Utility of IoT

The last important criterion to keep in mind while designing an IoT system is whether the system's utility is increased by connecting this particular system to the internet. In some cases, a system works just as well without being connected to the internet. This could partly take away concerns of the user about security and privacy that might arise when the system is connected to the internet (see Section 2.4). Consequently, without a connection to the internet, the system might also be less complex. These two features are beneficial for the UX of the system. One of the participants uttered their doubts about the utility of IoT systems in the interview (see Section 3.3.4).

Related to this criterion is the necessity of the different functions in an IoT system. While designing the IoT system, it is important to keep the system's functionalities limited. The system should do what it is supposed to do, but it should not have irrelevant additional features. Additional features could defeat the purpose of the system. This was also confirmed by the participants in the qualitative results (see Section 3.3.2).

4.2 Implementation of the Design Criteria

To investigate whether the derived design criteria would improve the UX of the IoT system, they were implemented in the design of the in-lite app. For most screen that had to be adapted based on the derived design criteria, different versions were created that integrated the design criteria in different ways. In the end, the chosen designs were the ones that were most consistent with each other and the rest of the app and the ones that followed Nielsen's heuristics best. It was also considered which design was least invasive to develop in the app. After the design criteria were implemented on the different screens, a prototype was created using Figma. This prototype was tested in the main study to get an insight into the effect of the design criteria on the UX and usability of the IoT system. The screens that were integrated into the prototype and thus tested in the main study, are described in the following. In the following, the figures show the old designs of the app compared to the new design, in which new elements are outlined in red.

4.2.1 Troubleshooting

Troubleshooting has to be implemented in a system in order to help the user when a system does not behave as expected. Smart-troubleshooting is a recently investigated topic, as described in Section 2.4. This feature would enable the system to recognize issues itself and solve them without intervention of the user. However, this smart technology still faces challenges that should be overcome before it can be implemented in different IoT systems. Another way of troubleshooting is by checking with the user whether the system is behaving as expected and providing them with solutions when this is not the case. This way of troubleshooting was implemented and tested in the in-lite system. It is first important to identify what could go wrong in the system and then describe the solutions that would solve these issues. In the current system, it was shown that users struggled when devices were not found. Users should thus be guided toward a troubleshooting page when they are not able to find the devices. In the in-lite system, this should be integrated into the process of connecting to smart hubs and smart lights. The troubleshooting page that was created for making a smart hub discoverable is shown in Figure 4.1. A button that guides the user to this page was added on the page when no smart hubs are found, as shown in Figure 4.2. This page was shown after the system has been searching for hubs for 120 seconds and has not found one in that time. The same button was also added below the list of discovered smart hubs. In the case that the desired hub is not found, the user could make that one discoverable by clicking the button and following the troubleshooting steps; this page is shown in Figure 4.3. Lastly, a pop-up and a button were added on the page that was shown while the system was looking for a smart hub. The button and pop-up appeared after the system had been searching for a hub for 60 seconds. This page is shown in Figure 4.4. When comparing the old design (Figure 4.4a) and the new design (Figure 4.4b), it is shown that a similar pop-up was already in the old design. However, it was found that the solution described in this pop-up is not the one that generally solves issues when connecting the smart hub. The new pop-up showed the solution that generally works, and if it did not, the user could try other solutions by clicking the button and following the steps on the troubleshooting page of Figure 4.1. A similar troubleshooting page was also created for the smart lights in the case that the user has not found all their smart lights; this is shown in Figure 4.5. Thus a button to this page was also added to the list of found smart lights. This is shown in Figure 4.10.

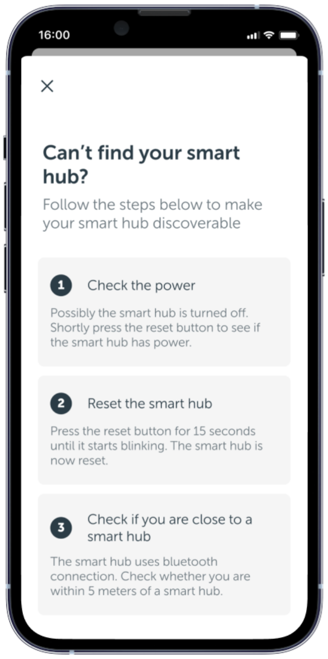
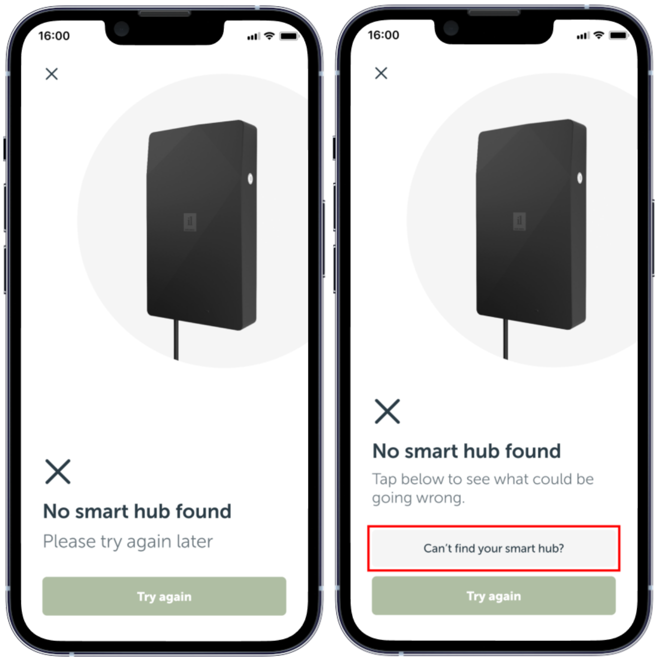
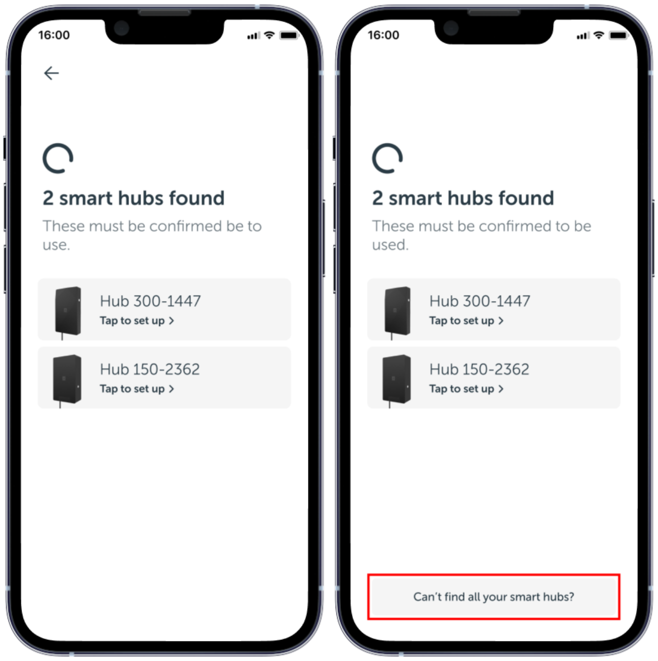


FIGURE 4.1: Troubleshooting to make a smart hub discoverable.



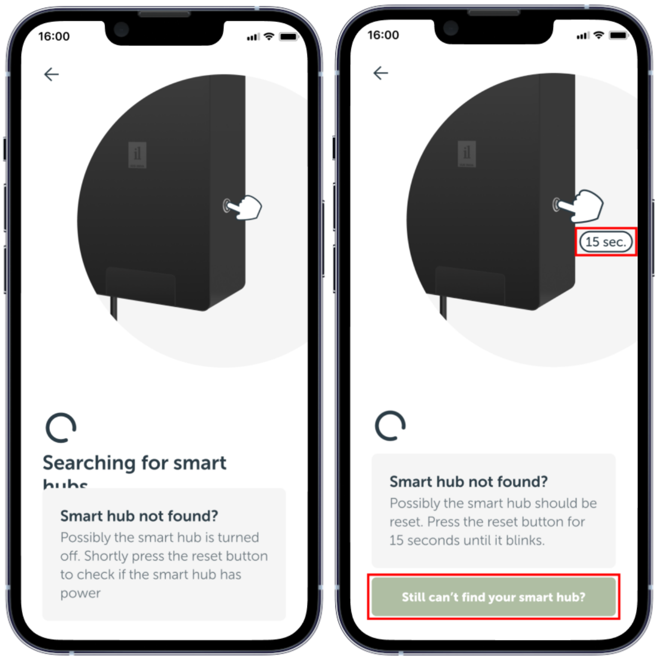
(A) Old design (B) New design

FIGURE 4.2: The screen that is shown when the system can not find a smart hub.



(A) Old design (B) New design

FIGURE 4.3: The screen that is shown when the system has found some smart hubs.



(A) Old design (B) New design

FIGURE 4.4: The screen that is shown after the system has been searching for hubs for 60 seconds.

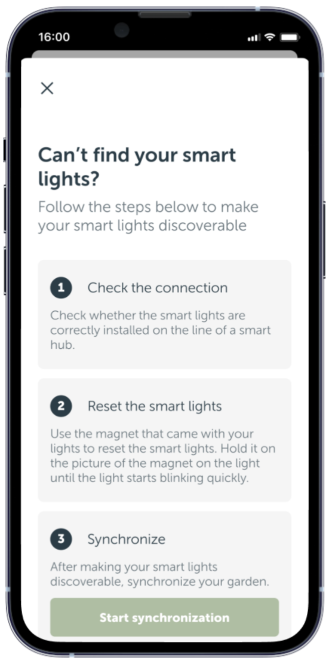


FIGURE 4.5: Troubleshooting to make smart lights discoverable

4.2.2 Minimize User Interaction

User interaction should be minimized where that is possible, meaning users should not be asked to make decisions about a process when the system can do this itself. In the current system, this design criterion was integrated into the synchronization process. Before integrating this criterion, the system would ask the user whether they wanted to synchronize their garden, which is necessary when there are smart lights connected to the system. This user flow is shown in Figure 4.6. The screens that belonged to the first part of this flow is shown in Figure 4.7. In the renewed design, the system looked for smart lights itself while the user was setting up the smart hubs and light zones. The new user flow is shown in Figure 4.8. If the system found smart lights in the system, it synchronized automatically when the setup of the smart hubs and light zones was finished; this screen is shown in Figure 4.9a. If it did not find smart lights, the system asked the user whether there are smart lights installed; this is shown in Figure 4.9b. If there were smart lights found, the user could use the first button and was sent to the troubleshooting page in Figure 4.5. If the user indicated that they did not have smart lights installed with the second button, synchronization was skipped.

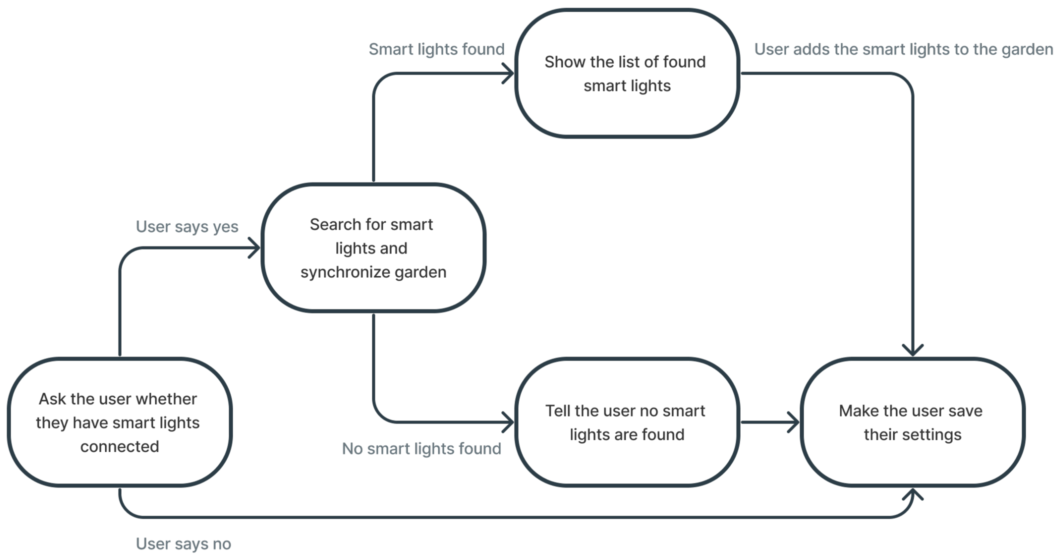


FIGURE 4.6: The old user flow for synchronization.

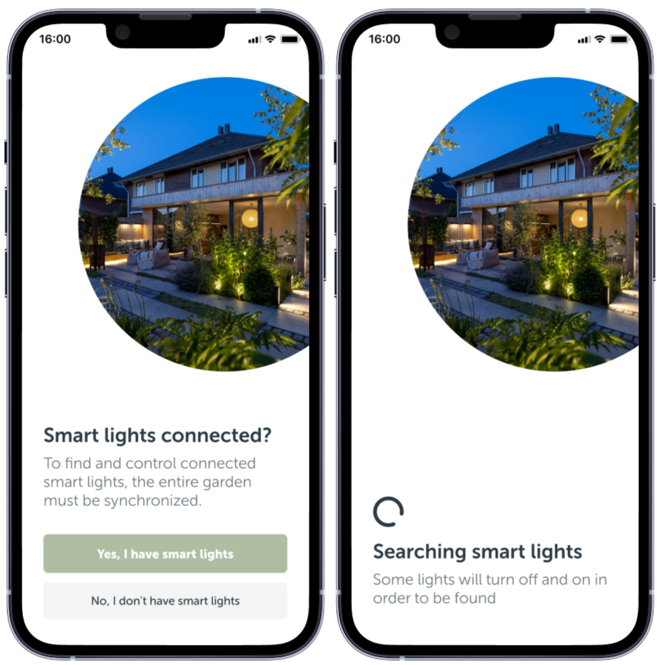


FIGURE 4.7: Screens that belong to the old flow for synchronization.

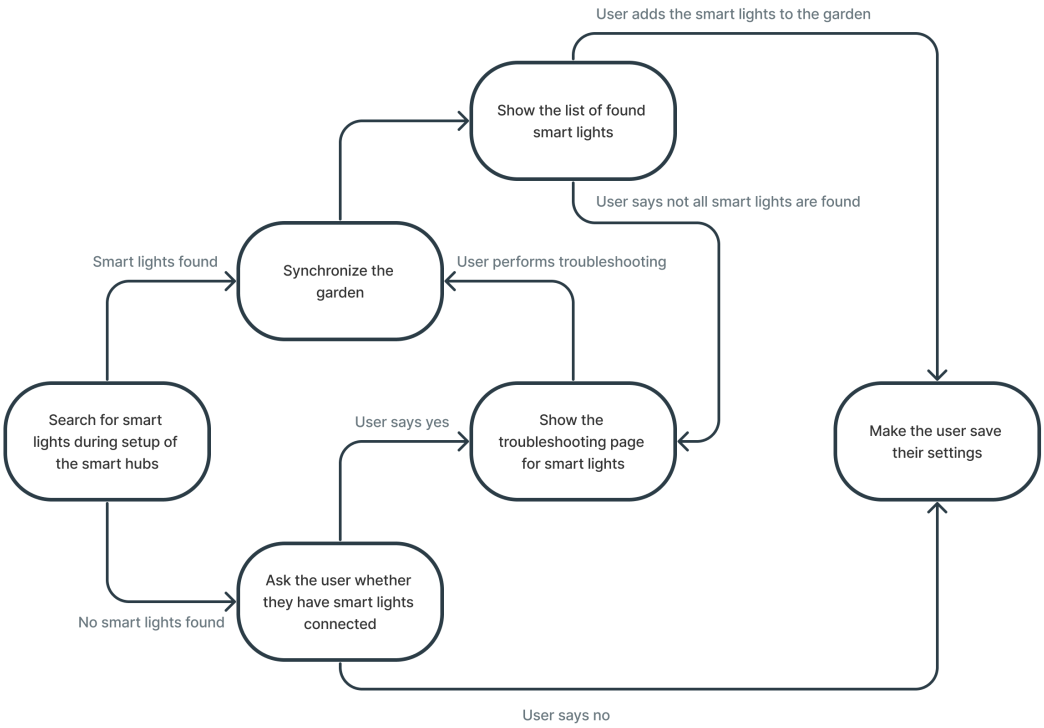
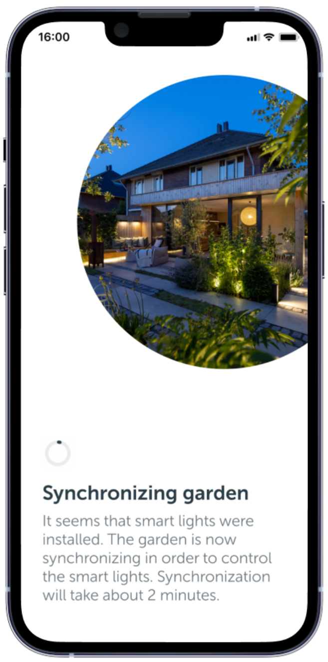
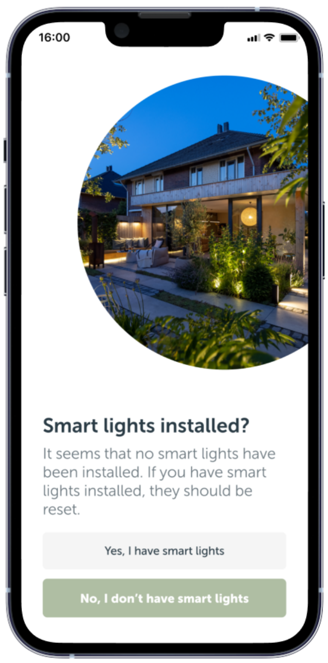


FIGURE 4.8: The new user flow for synchronization.



(A) Synchronization

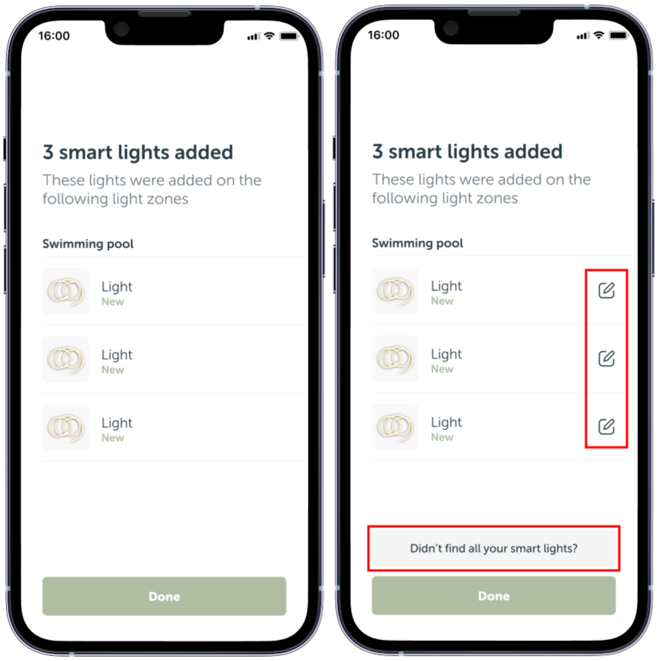


(B) No smart lights found

FIGURE 4.9: Screens that belong to the new flow for synchronization.

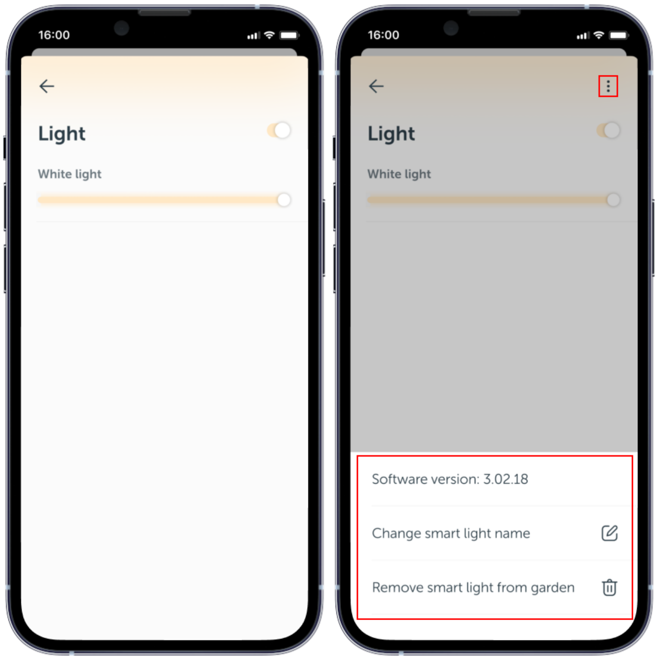
4.2.3 Clear Distinction Between Devices

As described before, the user should be able to distinguish the different devices that are present in the ecosystem. In other words, it should be clear to the user which device in the graphical user interface maps to which real device in the ecosystem. As described in Section 3.3.2, the user could already distinguish between the devices in the current system clearly, due to the images that were used in the app. Due to these images, the user knew which button to press or which device to look at. However, it was currently not clear to the user which light they are interacting with when selecting one, for example, to change its name or to delete it from the ecosystem. However, when selecting a hub to set up, the users confirmed that it was clear to them which one was selected because the reset button of the selected hub started flashing. To solve the problem of distinguishing the smart lights, a similar method could thus be adopted. Preferably, the selected smart light blinked once to indicate that it was selected. This, however, was not part of the graphical user interface design and was thus not adapted. However, to make it possible for the user to clearly distinguish between the lights in the app, it was made easier for the user to change the name of a light. Before, this was only possible in settings. In the new design, the user was able to change the name of the lights that were found after synchronization by clicking the edit button (see Figure 4.10). The same functionality was added where the user had to adjust the brightness of a single light. The user could press the name of the light or the three dots in the top right corner to adjust the name of a light (see Figure 4.11). This way, the user could turn the selected light on and off themselves, which would inform them about which light was selected.



(A) Old design (B) New design

FIGURE 4.10: The screen that shows what lights are found after synchronization.



(A) Old design (B) New design

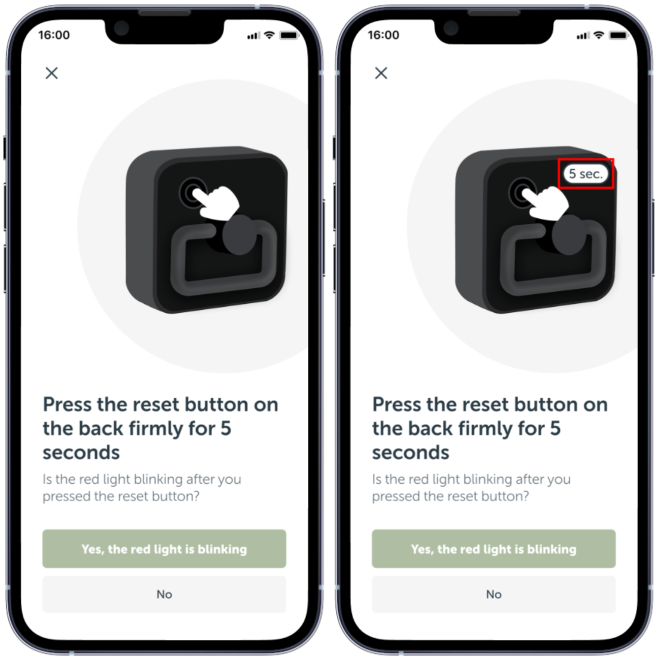
FIGURE 4.11: The screen in which the brightness of an individual light can be adjusted.

4.2.4 Consistent Interaction With Devices

For the current research, this criterion could not be tested. However, to make it easier for the user to know how to interact with the devices to connect them, the images that were used for this were made more detailed. It is shown in Figure 4.4 that this was done on the page when hubs are searched. A duration label is added to the page to indicate clearly to the user how long they have to press the button in order to reset the hub. A similar label was used on the screen that shows how to connect to the motion detector. This screen is shown in Figure 4.12.

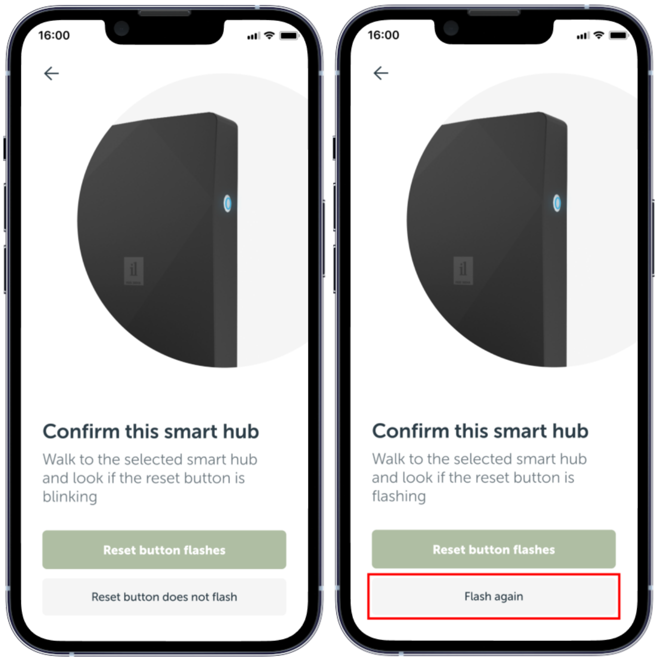
4.2.5 Interaction With Multiple Devices at Once

As described before, when a system has a separate user interface on a separate device, the system needs to take into account that the user is interacting with multiple devices at once. In the prestudy, this was a problem when the user had to connect to a smart hub. They had to read the instructions on the app and afterward look whether a light on the hub was blinking; however, the user did not have enough time to do both tasks because the light did not blink long enough. To be sure the light blinked, the user had to return to the list of found hubs and click the hub they wanted to connect to again. Ideally, the button would blink for a longer period. However, this was not part of the graphical user interface design. Therefore, a button was added to enable to user to make the light blink again; this is shown in Figure 4.13. This way, trying again took only one click instead of multiple.



(A) Old design (B) New design

FIGURE 4.12: Screen explaining how to connect to the motion detector.



(A) Old design (B) New design

FIGURE 4.13: Screen when connecting to a smart hub.

4.2.6 Utility of IoT

This design criterion could not be tested in the current study. This would defeat the purpose of the current study as the study is about increasing the usability and UX of IoT systems as opposed to finding out whether the usability and UX of the systems are improved when they are not connected to the internet. The general feeling about the current system was that it did not have irrelevant features that defeated the purpose of the system; therefore, no new functions were added. Users were only supported in their performance of certain tasks.

Chapter 5

Main Study: Overcoming UX Challenges

In this section, the main study is described. The study is conducted to test the design implications that are suggested in Chapter 4. First, the study design is discussed, followed by the results of the study.

5.1 Study Design

In this section, the study design of the main study is discussed. The study design of the main study was similar to the study design of the prestudy. The main difference was that a prototype of the in-lite app was used with the implemented design criteria instead of a fully developed version. This adjustment has implications for the study design, which is discussed in the following. Furthermore, the conditions were kept the same as in the prestudy to test the effects that the design criteria have on the usability and UX experienced by the participants.

5.1.1 Methods

The methods used in the prestudy were also used in the current study, namely usability testing and interviews. However, to let the users perform the usability tasks, a design prototype was used (see Section 4.2). This prototype was not actually connected to the devices in the IoT ecosystem. Therefore, the Wizard of Oz method had to be adopted. For the method, the researcher steps in to operate the system and perform the actions that it can not operate itself. The participants should get the impression that the system is acting autonomously and that it is responding to their input. This is often done to test out early aspects of a design that have not been implemented yet, as is the case for the current study (Riek, 2012).

As described in Section 3.1.2, the results of the prestudy and the main study were compared. For the main study, in which the new design was tested, new participants were gathered. A between-subject design was adopted. The reason for adopting this design as opposed to a within-subject design is that with the latter, a learning effect could be created. This would influence task performance, as the flow for performing most tasks is not different for both designs. Therefore, tasks would be performed based on prior knowledge, which would make using the system for a second time easier and could thus improve the usability and UX the participant experiences.

5.1.2 Independent and Dependent Variables

The independent variable and dependent variables are described in Section 3.1.2 and 3.1.3, respectively. From now on, the prestudy is referred to as the control condition, in which the original app was tested. The main study, in which the prototype is tested, is referred to as the experimental condition.

5.1.3 Participants

Again, ten participants were gathered through convenience sampling (seven female, three male). The criteria for participating were the same as in the prestudy. Participants were between 19 and 25 years old ($M = 22.7, SD = 2.06$). Nine of the participants were Dutch native speakers, and one of them spoke English. When asked to rate their tech-savviness on a scale of one to ten, the participants gave a score between four and ten ($M = 7.6, SD = 1.65$). Besides, seven participants had a tech-related job or followed a tech-related study. Five of the ten participants indicated to have had experience with IoT systems before. Lastly, five participants were iOS users, and five were Android users. Table 5.2 shows an overview of all participants in both conditions.

Condition	Participant ID	Age	Gender	Tech-Savviness	IoT-Experience
Control	1	25	M	6	No
Control	2	24	F	8	Yes
Control	3	25	F	6	No
Control	4	24	F	7	No
Control	5	24	F	8	No
Control	6	27	M	8	No
Control	7	26	F	8	Yes
Control	8	26	M	8	Yes
Control	9	29	M	9	Yes
Control	10	20	M	8	No
Experimental	11	23	F	7	No
Experimental	12	25	M	9	Yes
Experimental	13	20	F	4	No
Experimental	14	19	F	8	Yes
Experimental	15	24	F	8	No
Experimental	16	23	F	7	Yes
Experimental	17	23	F	7	Yes
Experimental	18	24	F	10	No
Experimental	19	21	M	7	Yes
Experimental	20	25	M	9	No

TABLE 5.1: A list of all participants.

5.1.4 Materials

The materials for the main study were the same as the materials used in the prestudy (see Section 3.1.5). However, for the main study, the in-lite app was not used. Instead, the participants interacted with the prototype that was created in Figma, which is described in Section 4.2. This prototype ran on the same iPhone X that was used in the prestudy.

5.1.5 Procedure and Measures

The procedure and the measures for the experimental condition were identical to the procedure and measures for the control condition (see 3.1.6 and 3.1.7, respectively).

5.2 Quantitative Results

In this section, the results for task performance, the SUS, and the AttrakDiff questionnaire of the main study are presented. These results are also compared to the results from the prestudy.

5.2.1 Task Performance

Task 1

The first task was done correctly by all ten participants. Again, participants did not see the light blinking on the first try. Eight of the ten participants thus used the *flash again* button, which is shown in Figure 4.13b, and connected to the hub afterward.

Task 2

The second task was again done correctly by all ten participants.

Task 3

The third task was done correctly by all ten participants.

Task 4

The fourth task was done correctly by all ten participants. Users all used the button *Can't find all your smart hubs?*, which is shown in Figure 4.3b. The participants all followed the steps from Figure 4.1 correctly.

Task 5

The fifth task was done correctly by all ten participants.

Task 6

The sixth task was done correctly by all ten participants. This time, no remarks were made by the participants that indicated they did not understand what was going on.

Task 7

The first part of this task was done correctly by all ten participants. They all used the button *Didn't find all your smart lights?*, which is shown in Figure 4.10b. The second step was performing the reset of the smart lights. Five users correctly followed the troubleshooting steps that are shown in Figure 4.5. However, five of the ten participants did not follow step 2 correctly and did not keep the magnet on the light long enough.

Task 8

This task was done correctly by all ten participants.

Task 9

This task was done correctly by all ten participants.

Task 10

This task was done correctly by all ten participants.

Task 11

This task again consisted of two steps. The first was finding out where to connect the motion detector in the app. This was done correctly by seven of the ten participants. The other three participants took some time to find the correct place to connect the motion detector. The second step was actually connecting the motion detector; this was done correctly by all of the ten participants.

Task 12

This task was done correctly by all of the ten participants.

Overview

The graph in Figure 5.1 shows the task performance in the control condition compared to the experimental condition. Tasks for which no mistakes were made during the prestudy or the main study were left out. The graph shows that participants in the experimental condition generally had a higher task performance than in the prestudy. The only deviation from this result is seen in the second part of task seven. More participants made a mistake during the reset of the lights in the experimental condition than in the control condition.

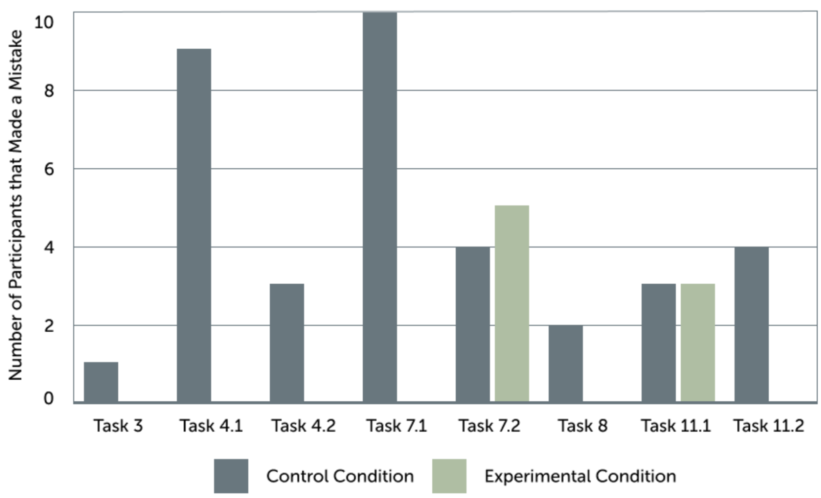


FIGURE 5.1: Task performance from the tasks in which mistakes were made during the prestudy and the main study.

5.2.2 SUS

As discussed in Section 3.2.2, the data gathered with the SUS in the control condition was normally distributed. A Shapiro-Wilk test was used on the SUS data from the experimental condition, and an alpha level of 0.05 was assumed again. It was concluded that this data was also normally distributed, namely $W(10) = 0.97, p = 0.92$, so $p > \alpha$. Therefore, an independent t-test was run for which an alpha level of 0.05 was assumed as well. It was found the mean SUS in the experimental condition ($M = 86.25, SD = 7.10$) was significantly higher than the mean SUS in the control condition ($M = 77.50, SD = 10.99$), namely $t(18) = 2.115, p = 0.049$. The means are shown in a bar graph in Figure 5.2. The SUS score given by each participant is shown in Table 5.2.

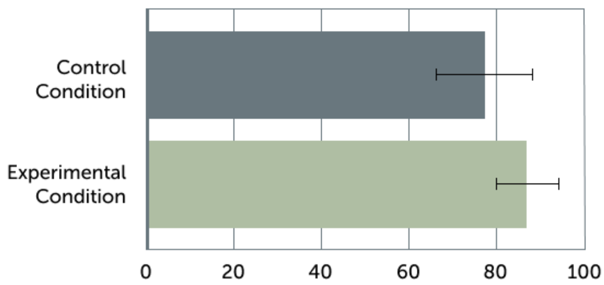


FIGURE 5.2: The mean SUS score given in each condition.

It could be noted that in the control condition, there were four participants with prior IoT experience, while in the experimental condition, there were five participants with prior IoT experience. One could argue that this could influence the mean SUS score because experience with a system could make it more usable for a user. However, when comparing the mean SUS score from the participants without prior IoT experience in the control condition ($M = 77.50, SD = 10.0$) to the mean SUS score from the participants with prior IoT experience in the control condition ($M = 77.50, SD = 13.99$), it was found that there was no difference in the means of the groups. The same was done for the mean SUS score from the participants without prior IoT experience in the experimental condition ($M = 88.50, SD = 5.76$) and the participants with prior IoT experience in the experimental condition ($M = 84.00, SD = 8.22$). Here it was also found that there was no significant difference between the means of the groups ($t(8) = 1.003, p = 0.345$). It was even found that the participants without prior IoT experience in the latter condition gave a higher mean SUS score than the participants with prior IoT experience.

The participants in the experimental condition rated themselves with a mean tech-savviness of 7.6, which is the same score the participants in the control condition gave themselves. It could thus be said that this feature of the populations did not affect the mean SUS score given by the groups.

Participant ID	SUS score
1	75
2	90
3	75
4	82.5
5	95
6	70
7	82.5
8	57.5
9	80
10	67.5

Participant ID	SUS score
11	85
12	75
13	82.5
14	90
15	95
16	80
17	80
18	87.5
19	90
20	97.5

TABLE 5.2: SUS scores given by the participants in the control condition (on the left) and in the experimental condition (on the right).

5.2.3 AttrakDiff

In Figure 5.3, the results from the AttrakDiff questionnaire for the control and experimental condition are shown. It is shown that in the first, second, and fourth dimension, the prototype generally scores better than the original version of the app. The design prototype seems to generally score lower in the third dimension, which is *hedonic quality - stimulation*.

The mean for each dimension for both conditions is shown in Figure 5.4. Those results were tested for significance. To do this, it was established what mean score participants gave for each of the four dimensions. These scores were tested for normality with a Shapiro-Wilk test. It was found that the scores given by the participants for each dimension in each condition were normally distributed. The means for each dimension were then compared between conditions with an independent t-test. It was found that the mean for *pragmatic quality* in the experimental condition ($M = 1.51, SD = 0.50$) was not significantly higher than the mean for *pragmatic quality* in the control condition ($M = 1.00, SD = 0.87$), namely $t(18) = 1.621, p = 0.122$). For *hedonic quality - identity*, the mean in the experimental condition ($M = 1.64, SD = 0.63$) was also not significantly higher than the mean in the control condition ($M = 1.39, SD = 0.87$), with $t(18) = 0.757, p = 0.459$. For *hedonic quality - stimulation*, the mean in the control condition ($M = 0.57, SD = 0.80$) was not significantly higher than the mean in the experimental condition ($M = 0.33, SD = 1.01$), with $t(18) = 0.593, p = 0.561$. Lastly, the mean for *attractiveness* in the experimental condition ($M = 1.90, SD = 0.57$) was not significantly higher than the mean in the control condition ($M = 1.46, SD = 1.01$), with $t(18) = 1.205, p = 0.244$. The results gathered with the AttrakDiff questionnaire are thus not significantly different between the conditions. However, the results for the dimensions are generally higher in the experimental condition.

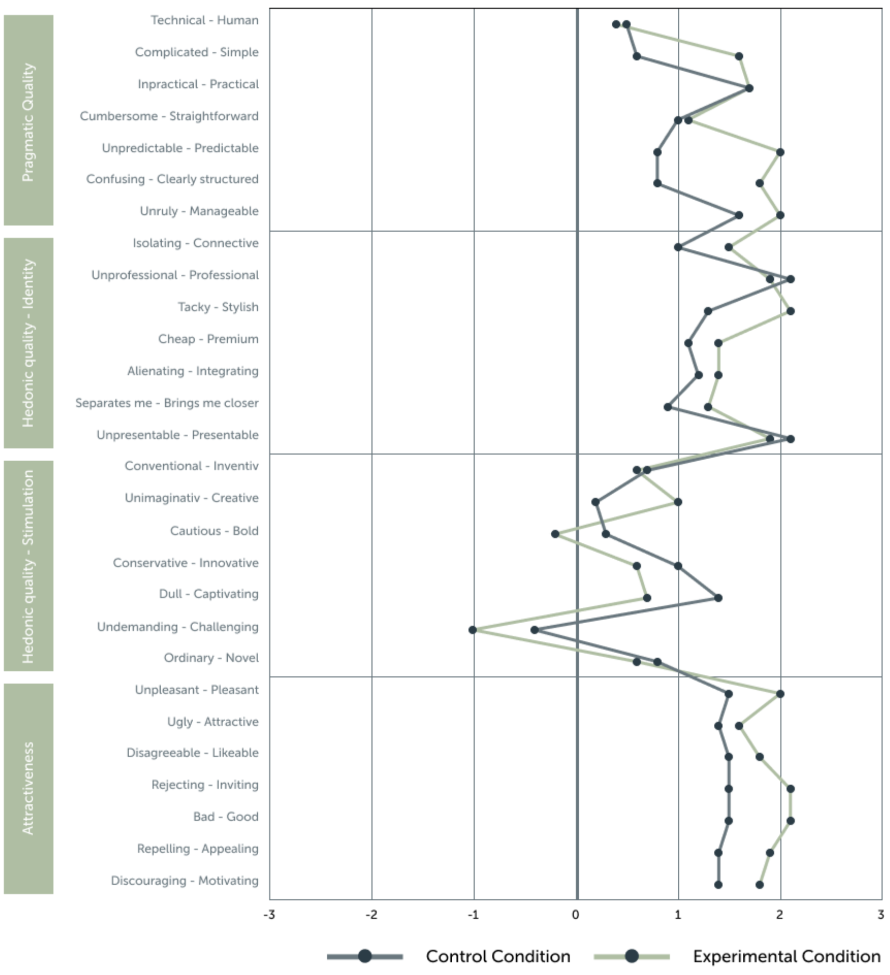


FIGURE 5.3: AttrakDiff results for the control condition and experimental condition.

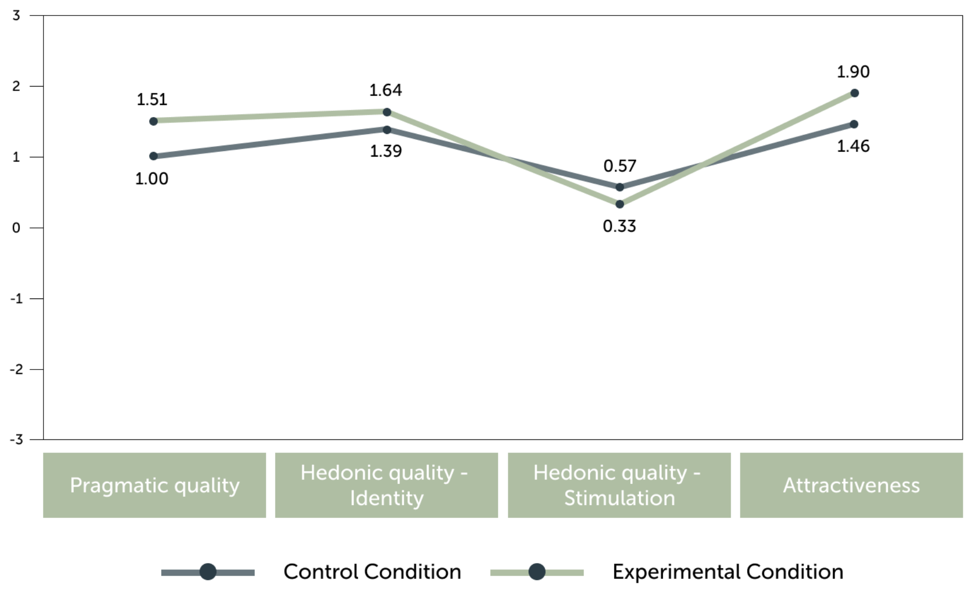


FIGURE 5.4: The means for each AttrakDiff dimension for both conditions.

5.3 Qualitative Results

The results of the interview from the main study were analyzed through thematic analysis, which is described in Section 3.3. The codes generated in the prestudy were used again in the current study. These codes were generated by randomly choosing three of the ten interviews and generating codes on them. These codes were used to code the other seven interviews. Afterward, the codes were categorized into four themes in the prestudy. Because the same codes were used in the current study, the same themes were also formed. The codes and themes are kept the same because this way, the results for each theme can be easily compared between conditions.

5.3.1 Ease of Use

In general, the participants felt that the system worked easily and intuitively. Many participants noted that they found the steps in the step-by-step plan very easy to use and that it helped them to connect devices that did not behave as expected:

> *It is nice that everything is explained in steps, so when you have finished one, and it goes well, you go on to the next one. It is nice that you don't have a huge text to read first. (P11)*

Participants also noted that it was nice that they could indicate themselves when there was a device that was not found:

> *P15: I found it a nice app. You just start with the set-up, which was quite easy. What I like is that it gave the suggestions: 'Did you not find everything?' or 'Is not everything working?', you can then click that and get help. That makes that you go fluently through the system.*
>
> *I: So you found the extra help helpful?*
>
> *P15: Yes, it makes that you don't have to search for something but that the system offers it to you.*

Participants also noted that they felt the app was clearly structured and uncluttered. They felt that they could navigate through the app easily and did not have to search long to find certain features:

> *It was very clear, where you had to be. And also if you were not in the correct place, it was easy to find where you had to be and how to get back and stuff. The design was nice and simple, which I like, with little fuss. (P17)*

Participants noted that it was also nice that all the settings for a light were found in one place. This way, they could distinguish for themselves very easily which light they were operating at the moment:

> *I found it very useful that you could adjust the settings of the lights and immediately operate them as well. This way you could check which light you are adjusting, by turning it on and off.*

These were the most important points noted by the participants that made the system easy to use. In general, participants were positive about their experience when using the system and when being asked what tasks they found the hardest to do during the experiment, participants noted that they had not found many hurdles that limited them in using the system:

> *I found the experience easy and straightforward, it was very clear what you had to do. You connect them, synch everything and then everything just works. After that it is very straightforward how everything works as well. (P12)*

5.3.2 Design

Overall, the statements made in this theme are very similar to the statements given by the participants in the prestudy. This is understandable since the aesthetics of the design have not changed. Things that are added to the design generally have more influence on the ease of use. Participants again mentioned a lot that they liked the fact that the app was clean and uncluttered:

> *Yes, well, I just really liked the app in general, because it did not have too many stimuli in the design, which makes it very clear (...) I think that is important for a system like this because you have to set it up yourself and I think this can be hard for people that don't have great technical skills. (P11)*

Participants also noted that they liked the layout of the app and the way it is structured. They felt like they could find everything easily and did not get stuck. They also noted this about troubleshooting that was added to help the participants through the reset process:

> *I found the app very clean. The manual for when something is not connected or when you have to set something up was step-by-step and clear. You could not really make a lot of mistakes. Also nice that you have three different screens, one for the lights that you have already added, one for routines and one for the settings. I think that is very clear. Also, when you have to operate a light it works exactly the way you would expect. (P16)*

Participants also noted that it was clear to them how they should change the name of a light. They also noted that this would help them to distinguish the lights:

> *I think that it is important to not have too much distraction in such an app because you're not going to do more than operate the lights. It is nice that you can install your own name and I think that it is very important because you make it easier for yourself. Also a light background combined with the green color is nice. (P11)*

Again, participants indicated that the pictures of the devices helped them in using the system and operating and distinguishing the different devices:

> *The pictures have helped me. Otherwise, I would not have known where a button was for example, because the text was not completely clear to me. The pictures really helped. (P13)*

5.3.3 Improving UX

This time, less statements were made within this theme, which means the participants had less suggestions to improve the UX of the system. The main thing that participants noted was that they found it confusing where to add the motion detector to the system. There were two places to do this, the first place to add this was in the settings tab. Some participants felt that this was not the most logical place:

> *I had a hard time connecting the motion sensor, because my first idea was to go to a light and add it there or to add it to a hub. I would not have looked for that in the settings. (P11)*

The other location to add the motion sensor was in the routine tab. However, participants that used this option also noted that they felt that this was not clear to them. The main reason for this was that this tab showed a sun and a moon, confusing the participants that it is only possible to set a time-based routine here, not a motion-based one:

> *The only thing that was not very good was that the description for routines indicated that you could use it for sunrise and sunset, but not that there are more options than that. I would indicate with an extra piece of text: 'or other things'. It is intuitive to click there anyway, but it was a but conflicting in my opinion. (P12)*

Other participants noted that they found it confusing that they should use a magnet to reset the smart lights. However, they also noted that it was clearly described how the magnet should be used, which made it easier to use the magnet:

> *The magnet is a bit confusing, but there are clear instructions about what to do with it, so that makes it easier. (P17)*

Lastly, a participant noted that they felt they had to go into too many screens to adjust settings. This made it a bit cumbersome for them to use:

> *I can not really name something to improve. The only thing is that I find that you have to click a lot to get somewhere, but I don't know if that is really something you can change. But I notice myself thinking: 'now I have to click here.'. I feel like the system is very simple, but because you have to go into a lot of different screens, it seems like it is more than it is. I feel like some things could be solved with a drop-down instead of a new screen. (P15)*

A similar note was made by another participant who felt like they had to do the same step over and over again in the setup for each hub:

> *I found setting up the light zones a bit repetitive. For each zone you had to do the same, I can imagine that with a lot of hubs this process takes a long time. (P19)*

5.3.4 Utility

Lastly, the remarks made in this theme were not different from the ones made in the prestudy. This was as expected because the functionality of the system had not changed. Users were thus again positive about the utility of the system:

> *My boyfriend uses smart lights in his house. At first, I did not really see why using this was necessary. But when you start using it yourself, you realize that it can be quite relaxed that you can operate the lights from your phone. Especially when you want the lights to be brighter one day and less bright on the other, this is a nice option to have. (P11)*

Besides being useful, participants also felt that it was a fun system to use and interesting to discover what it has to offer:

> *I: How would you describe the overall experience that you just had with the system?*
>
> *P20: Very fun, I found it very nice to play with the system.*

Chapter 6

Discussion

This chapter discusses the results of the main study in combination with the prestudy and the related works to answer the research questions. Furthermore, the limitations of the current study and recommendations for future work are discussed.

6.1 Review of Results

The results of the literature review, prestudy, and the main study are discussed and used to answer the research questions.

6.1.1 UX Challenges in IoT Systems

The first research question formulated in Section 1.2 is: "*What are current UX challenges when interacting with IoT systems?*". This research question can be answered based on the literature review and the prestudy.

Research by Bergman et al. (2018) introduced multiple UX challenges in IoT systems, some of which are confirmed by the prestudy. The first was **the lack of standardization amongst IoT systems**, which could cause cognitive overload due to the increasing functional complexity of the systems. Another problem addressed in this study were **interoperability issues**, which arise when devices in an IoT ecosystem are unable to communicate with each other. These problems were also found in the prestudy. Two interoperability issues were simulated, which had to be solved by the participants. The task performance showed that none of the participants knew how to do this. In the interviews, participants indicated that this was because the system did not communicate the issues to the user and because the system did not provide conventional and similar solutions to solve both problems.

This also relates to another issue described in the same study by Bergman et al. (2018), which describes that often **IoT systems are made too advanced for the user**. This problem was also found during the set-up of the system when the users had to decide whether they wanted to synchronize their system or not. Participants indicated that they did not understand what had to happen.

The next UX challenge described in literature which was also found in the prestudy, was introduced by Chuang, Chen, and Liu (2018). It describes that **IoT systems rarely provide sufficient feedback to indicate to the user what their current status is and what actions they are going to perform**. In the prestudy, users indicated during the interviews that they missed feedback from the lights in the system upon selection and also that there was insufficient feedback upon the selection of smart hubs.

Other UX challenges found in the literature review, such as interoperability issues between multiple IoT systems and privacy and security issues, were out of the scope of the current research because they are related to issues occurring on the network layer of IoT systems and thus not related to the user interface design of IoT systems.

To sum up, multiple challenges regarding interaction with IoT systems were found in the literature that were confirmed by the prestudy. This together answers the first research question.

6.1.2 User Interface Design for IoT

By conducting the literature review and the prestudy, challenges were established that influence the usability and UX of IoT systems. Next, design criteria were formulated and implemented in the in-lite system, which is described in Chapter 4. The main study was conducted to find out whether the established challenges were overcome and whether the usability and UX of the system had increased after the implementation of the design criteria.

The first indication of the usability of the system was whether the task performance increased after the implementation of the design criteria. It was found that for the tasks, users generally performed similarly or better in the main study than in the prestudy for all tasks. For tasks where no mistakes were made in the prestudy, no mistakes were also made in the main study. For tasks where mistakes were made during the prestudy, the task performance increased or remained the same in the main study. The only task where less mistakes were made during the prestudy was the resetting of the lights, which was the second part of the second task (See Figure 5.1). This could be due to the fact that users in the main study only had the description on the troubleshooting page in Figure 4.1 to guide them, while users in the prestudy got hinted about how to perform the reset. **Based on the task performance, it thus seems like the usability of the system has increased because less mistakes were made during the tasks after the implementation of the design criteria.**

The SUS was used to gain quantitative insight into the usability of the IoT system before and after implementing the design criteria. It was found that the SUS score given by the participants in the main study was significantly higher than the SUS score given by the participants in the prestudy (See Section 5.2.2). The groups scored similarly on tech-savviness and had a similar mean age; the increased SUS score could thus not be an effect of these characteristics. It could be said that, in general, the group of participants in the main study had more experience with IoT systems. However, it was found that participants with prior experience with IoT did not necessarily score higher on the SUS than the participants without prior experience with IoT. **It could thus be concluded, based on these quantitative results, that the usability of the system after the design criteria were implemented was higher than the usability before the design criteria were implemented.**

To gain quantitative insight into the user experience of the IoT system before and after implementing the design criteria, the AttrakDiff questionnaire was used. It was found that the latter design scored higher on three dimensions of this questionnaire, namely *pragmatic quality*, *hedonic quality - identity* and *attractiveness* (see Figure 5.4). These results were not significant though (see Section 5.2.3). The new design scored non-significantly lower on *hedonic quality - stimulation*, which could be explained. The dimension measures how well the system supports the human need for personal development and thus stimulates the user. It could be said that the new design stimulates the user less as the user has to make less decisions themselves. The system provides the user with more tools to help them solve problems occurring in the system, it provides the user with more feedback, and it is designed to make the user navigate through the system easier. **All in all, it could be said, based on these quantitative results, that the user experience of the new user interface is better than of the old user interface, however, these results are not significant.**

To gain qualitative insight into the usability and user experience of the IoT system, interviews were conducted. Four themes were derived in the prestudy that are discussed in the interviews; the interviews of the main study were also coded based on these themes in order to compare the two effectively (see Section 3.3 and Section 5.3). It was found that in the themes *Design* and *Utility*, the results between both studies were similar. This could be explained by the fact that the aesthetics and the architecture of the app, which were connected to the theme *Design*, did not change after implementing the design criteria. This is clearly shown in the Figures shown in Figure 4. The aesthetics of the old design and the new design are the same. The same holds for the utility of the app; no new functionalities were added that would expand or limit the utility of the app. There were changes in the theme *Ease of Use*. Before the implementation of the design criteria, participants discussed in this theme how they found it easy to connect to the smart hub and how they found it easy to control the lights after set-up. **After the implementation, all users described how they found the troubleshooting step-by-step plan helpful and how they found it easy and useful to be able to change the names of the lights in its control panel**. This thus shows that the users were positive about the new elements added to the design. There were also changes found in the theme *Improving UX*. Before implementation of the design criteria, participants discussed problems that decreased UX. Most of those were concerned with the interaction between the user and the system as a whole. An example was that **the app did not offer enough help to the user when some problem arose or that the user had a hard time identifying devices because the system did not give enough feedback about this**. After the implementation of the design criteria, most of these problems were solved. Problems decreasing UX that were mentioned after implementation were more about the architecture of the app. The most important one was that the users found it confusing where to add the motion detector. Users also found it confusing that they had to use a magnet to reset the lights, which was also mentioned before the implementation of the design criteria. This, however, was not related to the design of the user interface and was thus not adapted in the main study.

Summarizing, **users were more positive about the interaction with the devices in the IoT system after the implementation of the design criteria**. Usability issues that were mentioned after implementation were more specific to the studied system and would not necessarily apply to IoT systems in general.

When combining all quantitative results and qualitative results, it could be concluded that the usability and user experience of the system has increased after the implementation of the design criteria. First of all, users were able to perform the given tasks with less mistakes in the experimental condition. The SUS score given by the participants in the main study was also significantly higher than the SUS score given by the participants in the prestudy. In the experimental condition, three of the four UX dimensions from the AttrakDiff questionnaire were scored higher than in the control condition. The user experience was thus generally better after the implementation of the design criteria, although this result was not significant. Lastly, the qualitative data shows that participants were generally more positive about the system after the implementation of the design criteria than before. They had less suggestions for improvements related to the interaction with the devices in the system, and they felt the system was easier to use.

Therefore, the second research question *"How can graphical user interfaces of IoT* *systems be designed to overcome UX challenges?"* can be answered successfully. It was found that the design criteria suggested in Chapter 4 improved the usability and UX of the showcase IoT system. A graphical user interface of IoT systems should thus be designed by following Nielsen's heuristics and the six suggested criteria. The most important one of those design criteria seems to be the implementation of troubleshooting in the system. During the interviews, participants were most positive about this feature and this feature also increased task performance the most.

6.2 Limitations and Future Work

The study had some limitations, which will be discussed in this section.

The first limitation was that, due to time constraints, a limited amount of participants were recruited. Participants were recruited through convenience sampling and therefore had a limited age range, rated themselves high on tech-savviness, and the majority followed tech-related studies or had tech-related jobs. Therefore, the results of the study can not be generalized to an older or less tech-savvy population. Furthermore, the population sample was small, which resulted in non-significant results for the AttrakDiff questionnaire. A suggestion for future work would thus be to replicate the study with a bigger and more diverse population sample. This could point the results in the direction of significance and could make the results generalizable.

Another limitation that comes with the method of convenience sampling is the relation of the participant to the researcher. Because all participants knew the researcher, they might have had a tendency to tell the researcher what they thought they wanted to hear. Therefore participants might have been generally more positive about using the system and be less critical.

The next limitation is concerned with the fact that for the main study, the Wizard of Oz method had to be adopted, which was not used in the prestudy. In other words, a fully working system was used in the control condition, while in the experimental condition, a prototype was used. Therefore, not all functions worked in their full form in the main study, and participants were not able to discover all functions the app had to offer. This might have influenced the SUS scores and AttrakDiff results. A suggestion for future work would thus be to implement the design criteria in an IoT system and perform usability tests with it. This way, the users could explore the full system and thus get a better view of its usability.

A fourth limitation is also related to the method of the study. During usability testing, users were asked to think aloud, giving insight into the thought process of the participants. However, participants often forgot to speak their thoughts and had to be reminded multiple times during testing to think aloud. Therefore, interesting thought processes could have been missed. This also could influence the UX the participants experienced, as it could have felt unnatural to speak their thoughts.

The last limitation was that in the system, only adaptions could be made in the graphical user interface. Preferably, there would also be adaptions on the devices in the system to increase its usability. An example of this is described in Section 4.2.3. Ideally, the lights in the system would give feedback upon selection. This could, however, not be integrated because it was not part of the graphical user interface. Future work could thus look into applying the design criteria to all elements of an IoT system and thus improving the usability of the system as a whole.

The last suggestion for future work would be to look into ways of standardizing the development of IoT systems amongst IoT-developing companies. The UX challenges that arise and that are found during the literature review are a result of the lack of standardization and knowledge about UX criteria amongst companies. This results in interoperability issues amongst IoT devices from different companies and cognitive overload for the users because they have to learn new operation methods for each IoT system they use. By standardizing design criteria amongst IoT developing companies, these challenges can be overcome, and the usability of IoT systems can be increased.

Chapter 7

Conclusion

In this chapter, the research is concluded. The goal of this research was to discover UX challenges that users currently face when interacting with IoT systems. Consequently, it was investigated how the graphical user interfaces of these systems could be designed to overcome those challenges. A focus here was on increasing the usability of IoT systems and therefore their UX.

To reach these goals, several steps were taken. First, a literature review was conducted to establish the UX challenges regarding IoT systems that were discovered in other studies. The next step was to conduct a prestudy in which participants performed usability tests with an IoT system. They were asked about their experiences afterward.

The results of this prestudy confirmed the majority of the UX challenges discovered in literature. The next step was to derive design criteria for the graphical user interface of IoT systems and to implement them into the studied IoT system. To investigate the effect of the implementation of these criteria, the main study was conducted. The study design of this study was similar to the design of the prestudy, but with a design prototype of the IoT system. It was found that the UX and usability of the system had increased after implementing the design criteria into its graphical user interface.

As discussed in the literature review, Nielsen's heuristics are important guidelines to follow when designing the user interface of a system. Current research shows that additional guidelines have to be followed when designing usable IoT systems. An interesting next step to take in this field of science would be to look into ways of standardizing the development of IoT systems amongst IoT-developing companies. By standardizing the development of these systems and making the design criteria the norm, the majority of UX challenges discovered in this research could be overcome.

Bibliography

Asghari, Parvaneh, Amir Masoud Rahmani, and Hamid Haj Seyyed Javadi (2019). "Internet of Things applications: A systematic review". In: *Computer Networks* 148, pp. 241–261.

Atzori, Luigi, Antonio Iera, and Giacomo Morabito (2017). "Understanding the Internet of Things: definition, potentials, and societal role of a fast evolving paradigm". In: *Ad Hoc Networks* 56, pp. 122–140.

Bangor, Aaron, Philip T Kortum, and James T Miller (2008). "An empirical evaluation of the system usability scale". In: *Intl. Journal of Human–Computer Interaction* 24.6, pp. 574–594.

Benvenuti, Dario et al. (2021). "Detecting and Explaining Usability Issues of Consumer Electronic Products". In: *IFIP Conference on Human-Computer Interaction*. Springer, pp. 298–319.

Bergman, Johanna et al. (2018). "An exploratory study on how Internet of Things developing companies handle User Experience Requirements". In: *International Working Conference on Requirements Engineering: Foundation for Software Quality*. Springer, pp. 20–36.

Bevan, Nigel (2009). "What is the difference between the purpose of usability and user experience evaluation methods". In: *Proceedings of the Workshop UXEM*. Vol. 9. 1. Citeseer, pp. 1–4.

Blandford, Ann, Dominic Furniss, and Stephann Makri (2016). "Qualitative HCI research: Going behind the scenes". In: *Synthesis lectures on human-centered informatics* 9.1, pp. 1–115.

Brooke, John et al. (1996). "SUS-A quick and dirty usability scale". In: *Usability evaluation in industry* 189.194, pp. 4–7.

Caporuscio, Mauro et al. (2020). "Smart-troubleshooting connected devices: Concept, challenges and opportunities". In: *Future Generation Computer Systems* 111, pp. 681–697.

Choi, Wonyoung et al. (2021). "Smart home and internet of things: A bibliometric study". In: *Journal of Cleaner Production* 301, p. 126908.

Chuang, Yaliang, Lin-Lin Chen, and Yoga Liu (2018). "Design vocabulary for human–IoT systems communication". In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–11.

Clarke, Victoria, Virginia Braun, and Nikki Hayfield (2015). "Thematic analysis". In: *Qualitative psychology: A practical guide to research methods* 222.2015, p. 248.

Delgado Rodriguez, Sarah et al. (2021). "ActPad–A Smart Desk Platform to Enable User Interaction with IoT Devices". In: *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–6.

Eggen, Berry, Elise van den Hoven, and Jacques Terken (2017). "Human-centered design and smart homes: how to study and design for the home experience?" In: *Handbook of smart homes, health care and well-being*. Springer, pp. 83–92.

Fang, Kaiming and Guanhua Yan (2020). "IoTReplay: Troubleshooting COTS IoT Devices with Record and Replay". In: *2020 IEEE/ACM Symposium on Edge Computing (SEC)*. IEEE, pp. 193–205.

Farooq, M Umar et al. (2015). "A review on internet of things (IoT)". In: *International journal of computer applications* 113.1, pp. 1–7.
Fauquex, Milène et al. (2015). "Creating people-aware IoT applications by combining design thinking and user-centered design methods". In: *2015 IEEE 2nd World Forum on Internet of Things (WF-IoT)*. IEEE, pp. 57–62.
Glowniak, Jerry (1998). "History, structure, and function of the Internet". In: *Seminars in nuclear medicine*. Vol. 28. 2. Elsevier, pp. 135–144.
Hassenzahl, Marc, Michael Burmester, and Franz Koller (2003). "AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität". In: *Mensch & computer 2003*. Springer, pp. 187–196.
Kubitza, Thomas et al. (2020). "Developing IoT Systems: It's All About the Software". In: *Computer* 53.4, pp. 58–62.
Lee, In and Kyoochun Lee (2015). "The Internet of Things (IoT): Applications, investments, and challenges for enterprises". In: *Business horizons* 58.4, pp. 431–440.
Lewis, James R (2014). "Usability: lessons learned… and yet to be learned". In: *International Journal of Human-Computer Interaction* 30.9, pp. 663–684.
— (2018). "The system usability scale: past, present, and future". In: *International Journal of Human–Computer Interaction* 34.7, pp. 577–590.
Madakam, Somayya et al. (2015). "Internet of Things (IoT): A literature review". In: *Journal of Computer and Communications* 3.05, p. 164.
Nazari Shirehjini, Ali Asghar and Azin Semsar (2017). "Human interaction with IoT-based smart environments". In: *Multimedia Tools and Applications* 76.11, pp. 13343–13365.
Ngu, Anne H et al. (2016). "IoT middleware: A survey on issues and enabling technologies". In: *IEEE Internet of Things Journal* 4.1, pp. 1–20.
Nielsen, Jakob (2005). *Ten usability heuristics*.
Petrie, Helen and Nigel Bevan (2009). "The Evaluation of Accessibility, Usability, and User Experience." In: *The universal access handbook* 1, pp. 1–16.
Resnick, Marc L (2013). "Ubiquitous computing: UX when there is no UI". In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Vol. 57. 1. SAGE Publications Sage CA: Los Angeles, CA, pp. 1007–1011.
Riek, Laurel D (2012). "Wizard of oz studies in hri: a systematic review and new reporting guidelines". In: *Journal of Human-Robot Interaction* 1.1, pp. 119–136.
Rose, Karen, Scott Eldridge, and Lyman Chapin (2015). "The internet of things: An overview". In: *The internet society (ISOC)* 80, pp. 1–50.
Rowland, Claire et al. (2015). *Designing connected products: UX for the consumer Internet of Things*. " O'Reilly Media, Inc."
Sauro, Jeff (2011). *Measuring usability with the System Usability Scale (SUS)*. URL: https://measuringu.com/sus/ (visited on 08/03/2022).
Schrepp, Martin, Theo Held, and Bettina Laugwitz (2006). "The influence of hedonic quality on the attractiveness of user interfaces of business management software". In: *Interacting with Computers* 18.5, pp. 1055–1069.
Shackel, Brian (2009). "Usability–Context, framework, definition, design and evaluation". In: *Interacting with computers* 21.5-6, pp. 339–346.
Shin, Dong-Hee (2017). "A User-based Model for the Quality of Experience of the Internet of Things". In: *Network* 100, p. 80.
Singh, Sushil Kumar, Shailendra Rathore, and Jong Hyuk Park (2020). "Blockiotintelligence: A blockchain-enabled intelligent IoT architecture with artificial intelligence". In: *Future Generation Computer Systems* 110, pp. 721–743.

Suresh, Priya et al. (2014). “A state of the art review on the Internet of Things (IoT) history, technology and fields of deployment”. In: *2014 International conference on science engineering and management research (ICSEMR)*. IEEE, pp. 1–8.

Tawalbeh, Lo’ai et al. (2020). “IoT Privacy and security: Challenges and solutions”. In: *Applied Sciences* 10.12, p. 4102.

Thomas, Michael Onuoha, Beverly Amunga Onyimbo, and Rajasvaran Logeswaran (2016). “Usability evaluation criteria for internet of things”. In: *Int J Inf Technol Comput Sci* 8, pp. 10–18.

Walsh, Tanja et al. (2014). “Axe UX: Exploring long-term user experience with iScale and AttrakDiff”. In: *Proceedings of the 18th international academic mindtrek conference: Media business, management, content & services*, pp. 32–39.

West, Matthew T (2011). “Ubiquitous computing”. In: *Proceedings of the 39th annual ACM SIGUCCS conference on User services*, pp. 175–182.

Xenofontos, Christos et al. (2021). “Consumer, commercial, and industrial iot (in) security: Attack taxonomy and case studies”. In: *IEEE Internet of Things Journal* 9.1, pp. 199–221.

Appendix A

System Usability Scale (SUS)

In this Appendix the SUS is shown derived from Brooke et al. (1996).

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

Appendix B

AttrakDiff Questionnaire

Below the AttrakDiff word-pairs are shown derived from Walsh et al. (2014). Word-pairs 1 to 7 measure pragmatic quality, word-pairs 8 to 14 measure hedonic quality - identity, word-pairs 15 to 21 measure hedonic quality stimulation and word pairs 22 to 28 measure attractiveness. The participant can indicate in the 7-point scale what they think of the product with regards to a word-pair.

	-3	-2	-1	0	1	2	3
1. Technical - Human							
2. Complicated - Simple							
3. Inpractical - Practical							
4. Cumbersome - Straightforward							
5. Unpredictable - Predictable							
6. Confusing - Clearly Structured							
7. Unruly - Manageable							
8. Isolating - Connective							
9. Unprofessional - Professional							
10. Tacky - Stylish							
11. Cheap - Premium							
12. Alienating - Integrating							
13. Separates me - Brings me closer							
14. Unpresentable - Presentable							
15. Conventional - Inventive							
16. Unimaginativ - Creative							
17. Cautious - Bold							
18. Conservative - Innovative							
19. Dull - Captivating							
20. Undemanding - Challenging							
21. Ordinary - Novel							
22. Unpleasant - Pleasant							
23. Ugly - Attractive							
24. Disagreeable - Likeable							
25. Rejecting - Inviting							
26. Bad - Good							
27. Repelling - Appealing							
28. Discouraging - motivating							

Appendix C

Survey

C.1 Introduction

My name is Isa Buwalda and I am currently working on my graduation project for the master's Human Computer Interaction at Utrecht University. For this research, I am studying the interaction between humans and IoT systems. IoT stands for Internet of Things and it refers to a network of objects or 'things' that are connected to the internet and are able to communicate and exchange data. IoT systems often require minimal human intervention, although users can interact with the devices to set them up, give instructions or access the data for example.

To investigate this interaction, we will use a smart lighting system from in-lite. This system is created to light up gardens. The system is operated by an app created by the company CoffeeIT. In this research, I will ask you to perform some tasks with this system and talk to you about your experiences.

At the start of the research, I will ask you to complete a simple survey to gather some demographics. Afterwards, I will ask you to perform some tasks with the in-lite system while talking me through your thought process . When the tasks are completed, I will ask you to fill out two questionnaires about the system. This will give me an insight into your initial thoughts about the system. When you have finished those, I will ask you some questions to get a deeper insight into your experiences and opinions. Remember that all throughout the research there are no right or wrong answers or ways to perform tasks!

The materials produced during this session may be used for publication but will be fully pseudonymised. An audio recording may be taken of the interview (if you consent to it) and notes made. Taking part in this research is entirely voluntary. You may withdraw from the research at any time for any reason.
If you have any questions, feel free to ask at any time! Or send an email afterwards to i.buwalda@students.uu.nl or isabuwalda@coffeeit.nl.

C.2 Consent Form

Please fill out the following form by checking the boxes and signing the form with your name at the end.

- I confirm that the research project "Investigating Operation Challenges of IoT Systems to Improve the Usability through Intelligent User Interface Design" has been explained to me. I have had the opportunity to ask questions about the project and have had them answered satisfactorily.

- I consent to the material I contribute being used to generate insights for the research project "Investigating Operation Challenges of IoT Systems to Improve the Usability through Intelligent User Interface Design".
- I am aware that the researcher will take an audio recording of the interview. I understand that I can request to stop these recordings. I understand that I can ask for the recordings to be deleted.
- I understand that my participation in this research is voluntary and that I may withdraw from the study at any time.
- I consent to allow the fully pseudonymised data to be used for for future publications and other scholarly means of disseminating the findings from the research project.
- I understand that the data acquired will be securely stored by the researcher, but that appropriately pseudonymised data may in the future be made available to others for research purposes only.
- I confirm that I am 18 years or over.
- I understand that I can request any of the data collected from/by me to be deleted.
- I agree to take part in the above study on "Investigating Operation Challenges of IoT Systems to Improve the Usability through Intelligent User Interface Design".

Please write your name here as a signature of consent to all statements answered above.

C.3 Demographics

1. How old are you?
2. What is your gender?

    Male

    Female

    Non-binary/third gender

    Prefer not to say
3. What is your occupation?
4. If you are a student, what do you study?
5. Have you used the in-lite system before?

    Yes

    No

6. If you used the in-lite system before, how often did you use it?

    1-5 times

    5-10 times

    10+ times

7. Do you have experience with other IoT systems?

    Yes

    No

8. On a scale of 1 to 10, how tech-savvy do you consider yourself to be? (1 being not at all and 10 being very tech-savvy)

9. Do you have color blindness?

    Yes

    No

10. What mobile operating system do you use?

    iOS

    Android

    Other

Appendix D

Pilot Study

In this appendix, the tasks that were used in the pilot study are listed.

1. Link the lights to the smart hubs with the wires.
2. Create a new garden in the app and add the new smart hub to this garden.
3. Set up the first lighting zone of this smart hub.
4. Set up the second lighting zone of this smart hub.
5. Add the other smart hub to the garden by making it discoverable first.
6. Set up the first lighting zone of this smart hub.
7. Reset the smart lights.
8. Adjust the brightness of a lighting zone.
9. Find the manual for the motion detector.
10. Add the motion detector to the garden.
11. Set a routine such that the light that lights up your path responds to movement.