

UTRECHT UNIVERSITY

MSC HUMAN COMPUTER INTERACTION

The Understanding of Ambiguity as a Design Resource

Suggesting concrete design tactics for creating ambiguous design and testing its effects on users' reflection, user engagement and system usability.

Student

S.F. HERBES

s.f.herbes@students.uu.nl

Supervisor

P.W. WOŹNIAK

pawel.wozniak@chalmers.se

Keywords: ambiguity, design tactics, reflection, user engagement, design research

March 6, 2022

Abstract

Researchers have only recently started seeing ambiguity as something positive in the field of Human Computer Interaction (HCI). When used in the right way, ambiguous design can create more engagement or provoke thoughts in the user. Currently there is a gap in research between the understanding of ambiguity in theory and its applied practice. In this research we undertake this gap with a systematic review on current research in ambiguous design. Via this review we answered the question: “how is the concept of ambiguity understood and used as a design resource in the field of HCI?”.

Papers for the analysis were retrieved from ACM and selected using PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses). The final corpus of our review consists of fifty-five papers. We analyzed these papers with open coding using the software Atlas.ti. Of our corpus, in nine papers ambiguity was discussed and in forty-five papers ambiguity was applied in practice.

The review shows that ambiguity can be used in design, as well as in methods for creating design. Additionally, we show different goals and effects on users that can be achieved with ambiguous design, such as curiosity, social engagement, play, immersion, awareness, understanding, reflection, and creativity. We also found that balance is an important subject in creating ambiguous design, as too ambiguous design can create confusion in users, while design that is not ambiguous enough might not evoke the preferred effects. Yet, creating this balance is not an easy task, as it is difficult to know in advance how ambiguity will be perceived. Additionally, we discuss the different types of ambiguity and design tactics that are suggested in the papers. The suggested design tactics appeared to be abstract, and were not mentioned in the majority of the papers in our corpus. To make design tactics easier to use in practice, we created more concrete design tactics based on ambiguous designs that were presented in our corpus. In total, we present ten concrete design along with examples of ambiguous designs that apply these tactics: absence of purpose, absence of obvious connection, absence of information, absence of object, connection of external situations, abstract representation, data physicalization, ambiguous data, and uncomfortable design.

To test the findings in our literature review, we created a prototype that included different concrete design tactics. These were tested in four ways, with four conditions: Condition A did not include the design tactics, condition B included the absence of information, condition C included the absence of (original) context, and condition D included a combination of absence of obvious connection and abstract presentation. In all conditions, participants filled in a questionnaire that included questions regarding the user’s reflection, user engagement, system usability and perceived creepiness of the design. Additionally, participants were invited to an interview to get insights on additional thought processes. In total 82 participants were recruited (age: $M = 25.62$ years, $SD = 6.82$ years.). Although no significance was found, the results indicate a more diverse opinion about the system’s usability and the perceived creepiness of the device in ambiguous conditions, and interestingly an indication of lower user engagement in the ambiguous conditions. The qualitative results also resembled a variety of opinions, as participants either found the device confusing, experienced a learning curve, or thought it was easy to use. Interestingly, the majority of the participants were positive about the interaction with the device, even when they found it confusing. In future research this instance can be further studied. Other suggestions for future research are to study the right balance of ambiguity and create a standardized way to analyze the ambiguity of a design.

Acknowledgements

Before I present this thesis, I would like to thank some people. To start, I would like to thank my supervisor Dr. Paweł W. Woźniak, for his expertise, motivation and overall patience. This is my second thesis I did under your supervision and again you provided me with insightful feedback. Moreover, you motivated me into doing other valuable activities such as volunteering in the mobile HCI conference of 2020, and were able to connect me to interesting people in your network. I really appreciate this. I would also like to thank my second supervisor Christof van Nimwegen for taking the time to read and grade my thesis.

Furthermore, I would like to thank Jelle Algra for creating a 3D-model of my prototype in the software *Blender*. This saved me a lot of time trying to learn the software program myself.

Additionally I would like to thank all everyone who participated in this research, even the one that put my prototype on fire.

I would also like to thank my parents for supporting me emotionally and financially. You gave me the space to do this thesis on my own pace, without having to worry about other things.

Lastly, a special thanks to Katrine Nordeide Kuipers who read and gave feedback to a part of my thesis, Marit Bentvelzen who spent time with me at campus during the summer holiday while I was working on my prototype, and Thomas Jorna whom I studied with in the last couple of months; this really motivated me. And of course thanks to all other friends and fellow student who supported me through the process of my thesis. It was a difficult and sometimes lonely time to write a thesis, due to the pandemic, but you were always there for support.

Contents

1	Introduction	5
1.1	Why study ambiguity?	5
1.2	Goal	6
1.3	Outline	6
2	Related Work	7
2.1	Method	7
2.1.1	Approach and collection	7
2.1.2	Selection of papers	8
2.1.3	Analysis	10
2.2	Corpus Overview	11
2.2.1	Publication timeline	11
2.2.2	Fields of application and discussion of ambiguity	11
2.3	Systematic Literature Review	12
2.3.1	Ambiguity as a method	12
2.3.2	Goals and effects of ambiguous design	13
2.3.3	Finding the right balance	17
2.3.4	Different types of ambiguity	18
2.3.5	Suggested design tactics and strategies	21
2.3.6	Concrete design tactics	21
2.4	Summary and contribution	27
2.5	Our study	27
3	Design	29
3.1	Forming the prototype	29
3.1.1	Chosen design tactics	29
3.1.2	Data to reflect on	30
3.1.3	Physical design	30
3.2	Design process and review	31
3.2.1	Review	32
3.3	Implementation	32
3.3.1	Recovery of the prototype	33
4	Method	34
4.1	Experimental study design	34
4.1.1	Conditions	34
4.1.2	Experiment tasks	34
4.2	Measurements	36
4.2.1	Self reflection	36
4.2.2	User engagement	36
4.2.3	Usability	36
4.2.4	Creepiness	36
4.3	Procedure	37
4.4	Participants	37
4.5	Potential issues to take into account	37

5	Results	39
5.1	Quantitative results	39
5.1.1	Self reflection in technology	39
5.1.2	User engagement	40
5.1.3	Usability	41
5.1.4	Perceived creepiness	42
5.1.5	Summary of quantitative results	43
5.2	Qualitative results	43
5.2.1	Overall impression	43
5.2.2	Feedback	44
5.2.3	Intention to use	46
6	Discussion	48
6.1	Design adaptations	49
6.2	Limitations	50
6.3	Future Research	51
7	Conclusion	53
A	Corpus overview	59
B	Additional resources	61
C	Questionnaire	62
C.1	Informed consent	62
C.2	Participant number	62
C.3	Given instructions and tasks	62
C.3.1	Introduction (all conditions)	62
C.3.2	Instructions (all conditions)	63
C.3.3	Additional instructions (only for conditions A, C, D)	63
C.3.4	History (all conditions)	63
C.3.5	History (only conditions A, B, D)	63
C.3.6	Tasks	63
C.4	Questions	64
C.4.1	Technology Supported Reflection Inventory	64
C.4.2	Self-Reflection and Insight Scale	64
C.4.3	User Engagement Scale - Short Version	65
C.4.4	System Usability Scale	65
C.4.5	Perceived Creepiness of Technology Scale	66
C.5	Demographics	66
C.6	Updates on results	66
D	Qualitative data	67

1 Introduction

At the central train station in Utrecht, at the end of the outside passage road, a digital display provides passengers with daily prompts from nature (Figure 1). Different news websites (e.g., [van Huut, 2018](#)) have discussed the artwork, Arrivals/Departures, by Marcus Coates. The artwork raises one main question: what is that display doing there? It intrigues passengers, and it might even slow them down during rush hour. This piece of art is a clear example of ambiguous design.

Generally, ambiguity is not welcomed in Human Computer Interaction (HCI) related systems, since HCI design should be effective, efficient and usable and ambiguity is seen as the opposite of usability and efficiency ([Rogers et al., 2011](#)). Without a doubt there are systems that should not be ambiguous, in particular systems that are critical for safety or made to be efficient. However, when designing for reflection, ambiguity can play an important role. [Gaver et al. \(2003\)](#) have opened the discussion on ambiguous design in research, by analysing digital art objects and showing that the ambiguity can create engagement or provoke thoughts in the user. They suggest different types of ambiguity one could design for along with design tactics for each type. Researchers have built further on the topic of ambiguity and purposely started creating it in designs to provoke the user.

Ambiguous design can be seen as design where art and HCI meet. This makes it difficult for researchers to create clarity on the subject while keeping objectivity in tact. Indeed, reviews on ambiguity have already been made ([Boehner and Hancock, 2006](#); [Gaver et al., 2003](#); [Sengers and Gaver, 2006](#)), however these only touch the surface of research on ambiguous design. A systematic review on current studies on ambiguous design is still missing.

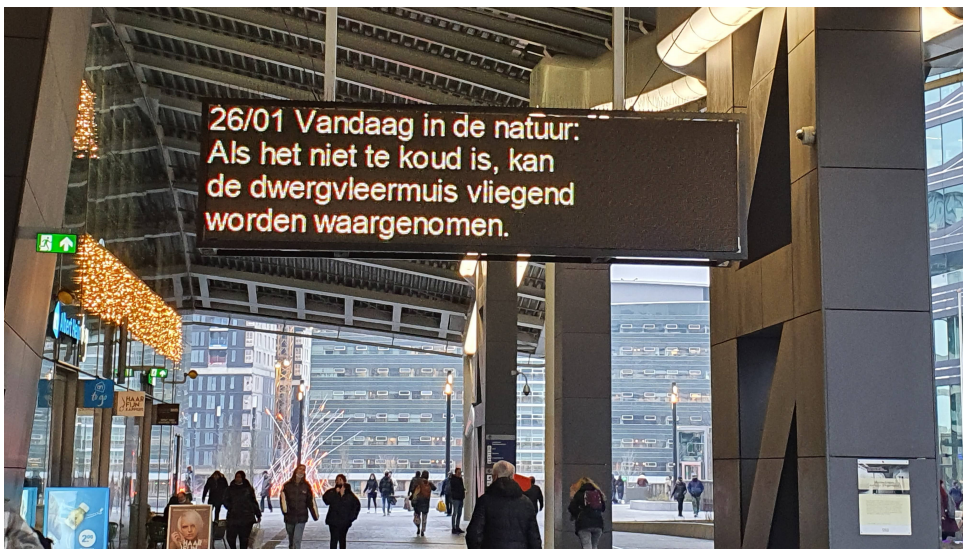


Figure 1: The artwork Arrivals/Departures by Marcus Coates at Utrecht Central station in the Netherlands. The artwork shows a daily prompt from nature in Dutch: “*Today in nature: if it is not too cold, the pipistrelle can be observed flying.*”

1.1 Why study ambiguity?

Personal informatics have become more popular, using technologies users hope to achieve their personal goals and in some cases even try to change their behaviors.

However, a problem that is faced when designing these technologies, is that researchers might

quickly jump to persuasive technologies, ie. modernist technologies (Brynjarsdottir et al., 2012). In modernist technologies, a solution to a proposed problem is to narrow the vision, where only limited aspects of a large and complex problem are tackled, which makes the problem more calculable and easier to tackle. A problem that is often faced in this technique is that it is assumed that all factors have been taken into account, while it is merely a simplification of reality. This makes the proposed solution susceptible to breakdown in the ‘real’ world (Brynjarsdottir et al., 2012). To make users achieve their goals or behavior change, a more reflective approach is needed. Bentvelzen et al. (2021) created a framework that shows design resources for the support of reflection in systems (temporal perspective: past, future, memories, slowness; conversation: with others, with technology; comparison: absolute reference, social reference; discovery: ambiguity, provocation, reframing). If these design resources are effectively used in a system, it would suggest that the system is effectively supporting reflection. Currently there is still a lack of research on these design resources which makes it difficult to judge their effects on reflection. This lack of research on the topic can make it easier for researchers to fall back to familiar modernist technologies, despite its critique. Creating clarity on the topic of ambiguous design can offer support on research on reflection and guide researchers into slowly stepping away from modernist technologies in situations where they are not ideal.

1.2 Goal

The goal of this research is to create a better understanding of ambiguity and its implementation as a design resource. To do this, we formulate the following research question: “*how is the concept of ambiguity understood and used as a design resource in the field of HCI?*” To answer this research question, we perform a systematic literature review that examines contemporary understanding of ambiguity and studies that have created ambiguous design. Based on the results of the systematic review, we create an ambiguous design ourselves. Additionally, we study how users react to and reflect on the ambiguous design.

1.3 Outline

In Chapter 2 we present the related work in the form of a systematic literature review, we present the method of building the corpus, the overview of the corpus and the results, i.e. the systematic review of the corpus. In the systematic review we suggest a list of concrete design tactics that are based on ambiguous designs in our corpus. In Chapter 3 we discuss the ambiguous design we created based on the concrete design tactics. In Chapter 4 we describe the procedure of the experiment in which we test the prototype. In Chapter 5 we discuss the quantitative and qualitative results of the study. Finally, in Chapter 6, we provide a discussion on the results, including the contributions and limitations of our study and direction for future research. In Chapter 7 we present a final conclusion.

2 Related Work

In this chapter we discuss the literature about ambiguity in design through a systematic literature review. First, we describe the method we used to create the corpus (Section 2.1). Second, we provide an overview of our corpus (Section 2.2). Third, we present the results of the systematic literature review that also includes an overview of the concrete design tactics we created (Section 2.3).

2.1 Method

To create the final corpus we first collected data from ACM Digital Library. Then, we filtered this data in five different phases. Finally, we analyzed the final corpus to write a systematic literature review.

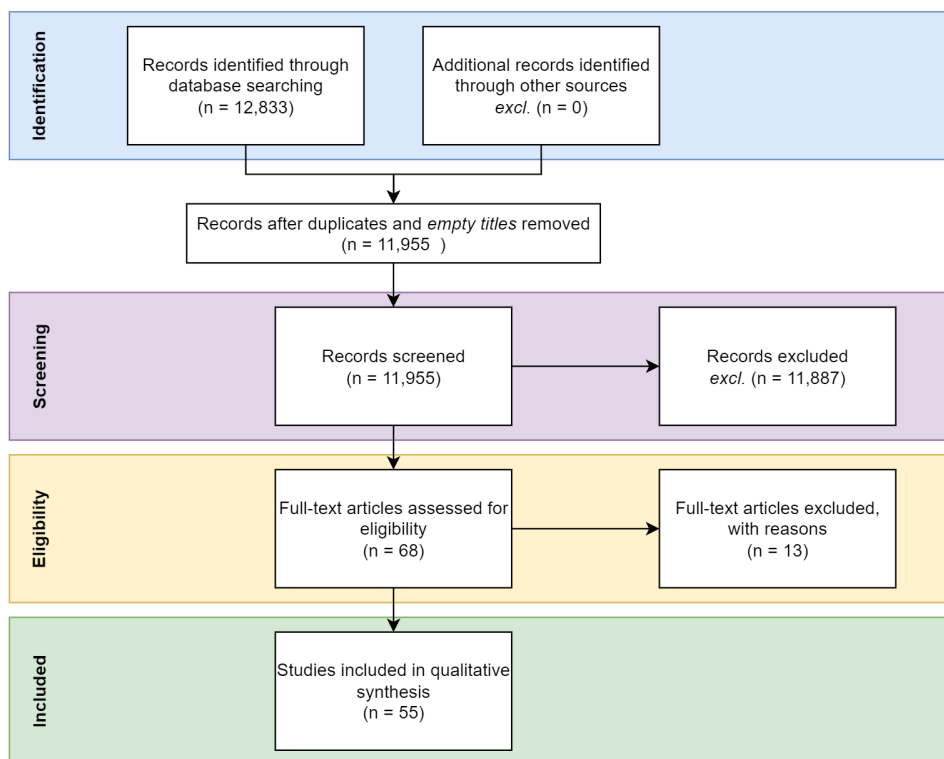


Figure 2: Overview of the selection process using PRISMA. We aim to select papers that go deeper into the topic of ambiguity as a design resource in HCI.

2.1.1 Approach and collection

The main goal for the systematic review is to create a better understanding of ambiguity. The papers were retrieved from the ACM Guide to Computing Literature from ACM Digital Library. This is the most comprehensive bibliographic database that is focused exclusively on the field of computing. We used a keyword-based search to collect all papers that included ambiguity in their content. To reduce the amount of papers about computer software and increase the focus on

designs for humans, from the ACM Computer Classification System two classifications were applied: “Human Computer Interaction” and “Human-Centered Computing”. Under these classifications, all papers using “ambigu*” (ambiguity, ambiguous) as a keyword and with a publication date before January 2021 were collected. We applied the keyword-search to different article aspects such as; title, abstract and full text. This created a large dataset with a high degree of noise, as the word ambiguity can be used for purposes other than a design resource. The advantage of such a large collection of papers is that it provided us with more certainty of having a complete final corpus.

2.1.2 Selection of papers

To keep structure in the selection process we used PRISMA (Preferred Reporting Items for systematic Reviews and Meta-Analyses (Figure 2). PRISMA is relevant to all types of systematic reviews as it reduces bias and increases transparency during the data collection (Moher et al., 2009). We collected the papers using the following query:

[All: ambigu*] AND [Publication Date: (01/01/1950 TO 12/31/2021)] AND [CCS 2012: Human-centered computing] OR [CCS 2012: Human computer interaction (HCI)]

This initial collection resulted in a total of 12,883 papers. As mentioned in the previous section, this collection contains a high degree of noise. To tackle this problem, we filtered the collection in five phases: 1) removing duplicates, 2) based on the title, 3) based on the research type, 4) based on the meaning of the keyword (ambiguity), 5) based on the content and collaborative filtering. The first four phases were done by one researcher and the fifth phase was done by three researchers to minimize the risk of bias.

Definition of ambiguity as a design resource. Before we started filtering the collection, we established a broad definition of ambiguity as a design resource. Clarity on this definition is important, because ambiguity can be used to describe different concepts, artifacts and can also be something other than a design resource. The definition on ambiguity, as presented below, was mainly based on the research of Gaver et al. (2003). During our selection process this definition was merely used as a guideline.

Definition of ambiguity. In the context of HCI and interaction design, ambiguity should be defined as a (design) resource used for adding more than one possible meaning to the design of a system, with the intention to evoke something from the user. In particular ambiguity can be used to 1) suggest issues and perspectives, 2) evoke a reaction of skepticism or belief, 3) let the user ask questions about the truth of a situation, 4) make the user interpret the design in their own way, 5) have the user freely interacting with the system, 6) lead the user to consider new beliefs and values, 7) encourage engagement with the system.

Gaver et al. (2003)

Phase 1: merge duplicates, remove blanks. In our collection we found that a number of references either were the same, or did not have a title and thus were papers we would not be able to find. To reduce this noise in the collection, we merged duplicate references (888) and removed references without a title (40). This resulted in a total of 11,955 references .

Phase 2: filter based on title. In this phase we filtered the references based on the title. We started by filtering based on the title because a title generally gives a clear indication of what a paper has to offer. Additionally, reading more than the title alone would take an enormous amount of our time, which would not outweigh the advantages, since our collection contained 11,955

references in this phase. However, we do acknowledge that titles can be unclear. To take this into account, we decided to include unclear titles for the next filtering phase to make sure that potentially relevant articles were not excluded from the collection.

It was important to keep our focus on ambiguity as a design resource as this was our main goal for the systematic review. Furthermore, we wanted to include papers that go deeper in on the topic of ambiguity, by either analyzing it, applying it in practice, or analyzing other aspects which ambiguity can be a part of – i.e. different things that it can evoke from users. Moreover, we focused on designing for ambiguity in HCI, which means that focus on computer technologies alone did not suffice for our collection. These motives led us to the following criteria for removing papers based on the title:

- Titles that did not meet our broad definition of ambiguity as stated earlier.
- Titles that did not discuss ambiguity either on a meta level or in a system design.
- Titles that focus on computer programming, software development and other related technical studies that did not include users.

After this filtering phase, we removed 11,330 references from our collection, which resulted in a collection of 625 references.

Phase 3: filter based on research type. In the third phase we filtered the papers based on the research type. We started by downloading all papers from the ACM library; if these were either locked or not available, we tried accessing them through other websites via Google Scholar. In this selection we only wanted to include full papers, and thus excluded other types of papers such as posters, columns, demonstrations and notes. Full papers would give us more in depth knowledge on the topic of ambiguity, with extensive explanations than other types of papers. We needed this to create an in-depth review of ambiguity in HCI. In total we removed 199 papers, which resulted in a collection of 426 papers.

Phase 4: filter based on meaning of keyword. In the fourth phase we filtered our collection in a more thorough manner where the meaning of the keyword (ambiguity, ambiguous) was evaluated. We searched the keyword in the remaining papers and studied the paragraph containing that keyword. Only reading a part of a paper would suffice for this phase, as the intention of the use of the word ambiguity can generally be read off from a sentence or paragraph. Again, in case of doubt, the paper was kept in the collection to ensure that potential relevant papers would not be excluded.

Ambiguity can be used in a number of different ways, for instance as a figurative speech where an element is unclear and described as ambiguous. For our systematic review we are primarily focusing on design that has been deliberately made ambiguous, or later found to be ambiguous and is further discussed. With these motives in mind we created the following exclusion criteria:

- Papers that used ambiguity as a figurative language where it would stand for unclarity.
- Papers that used ambiguity in a way that did not refer to a designed system or did not further investigate the topic of ambiguity on a meta level.

In total we removed 127 papers from our collection, which resulted in a collection of 299 papers.

Phase 5: filtering based on content and collaborative filtering. In the fifth and last phase, the papers from our collection were filtered based on the content. All remaining papers were fully read by one researcher, and labeled based on its value for the systematic review. For the

systematic review we wanted to include papers that would give us more in-depth information on the use of ambiguity as a resource design for HCI. We wanted to collect a number of designs that were made ambiguous, to be able to analyze them and potentially find patterns. These considerations led us to the following criteria to be labeled (low, medium, high) and papers that did not meet the criteria were removed:

- Low: papers that discuss ambiguity in a negative manner and focused on usability related studies.
- Medium: papers that mention ambiguity just once, but use it to describe a designed system.
- Medium/high: papers that use or describe ambiguity as a design resource.
- High: papers that discuss ambiguity as a design resource on a higher or deeper level.

In total, 137 papers were removed from the collection, leaving 162 papers to be evaluated. For finalizing the corpus two other researchers were engaged in the selection process, to minimize the risk of bias. These researchers also labeled the remaining 162 papers, based on the criteria mentioned above. Communication and labeling of the papers was done via Google Spreadsheets. We finally decided that papers labeled as low value should be excluded from the final corpus. In these papers ambiguity was mainly used in studies about usability, where ambiguity was seen as a negative thing that should be resolved, and not as a design resource that can be used to evoke positive things in users. This resulted in an exclusion of 94 papers, and a remaining 68 papers in our corpus.

Extra papers excluded. During the writing process of the literature review, we found 13 papers not to be useful after all, as they did not provide the reader with their own insights or were different papers in which the same design was discussed. This finally resulted in a **corpus of 55 papers**.

2.1.3 Analysis

After establishing the final corpus, we imported the papers in Atlas.ti, a software that supports research and qualitative data analysis. Along Atlas.ti, a spreadsheet was used with the full corpus added in rows and the main questions are presented in columns. The spreadsheet provided the possibility to add additional notes on the questions to support the coding process. In Atlas.ti we analyzed the papers using open and closed coding. For open coding we created leading questions that have been used for other literature review studies and adjusted them to fit our topic of ambiguity (Baumer et al., 2014; Doherty and Doherty, 2019; Terzimehić et al., 2019). These questions were there to serve as a guidance during the analysis and keep a general structure.

- In which domain was ambiguity supported?
- How was ambiguity defined?
- If the paper describes an interactive system, how is ambiguity applied as a design resource?
- What goal does the author want to reach using ambiguity?
- How was ambiguity evaluated? (i.e. discussion, review, application)
- What were the results of using ambiguity?
- What are the (other) co-aspects of ambiguity? (i.e. stated benefits)

- If ambiguity is discussed on a higher or deeper level, how is it explained?

The codes created for open coding were in turn used for closed coding. This way we could check whether the codes included all the related papers. In addition, closed coding was used to see whether ambiguity was implemented according to the types and tactics originally suggested by Gaver et al. (2003) and whether new suggestions based on this analysis were given.

2.2 Corpus Overview

In this section we provide an overview of the corpus. We give an overview of the years in which the papers were published to create an impression of our corpus. We also present an overview of the fields of application and the way in which ambiguity is discussed.

2.2.1 Publication timeline

The complete corpus consists of 55 papers and are published between 2003 and 2020 (figure 3). A clear pattern in the timeline cannot be observed as the number of publications fluctuates per year.

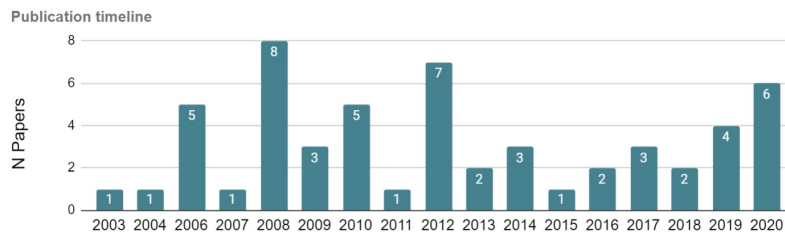


Figure 3: Timeline of years in which papers from the corpus were published.

2.2.2 Fields of application and discussion of ambiguity

In our corpus we included papers that review ambiguity in different ways: 1) with the creation of ambiguous design, 2) through discussion on ambiguity, 3) by discussing ambiguity in a discussion on something else (Figure 4). Furthermore, the fields of applications in which ambiguity is discussed varies widely and can be divided in different categories: art/culture, communication, daily life, design research, education/work, health/well-being, home and play (Table 9 in Appendix A).

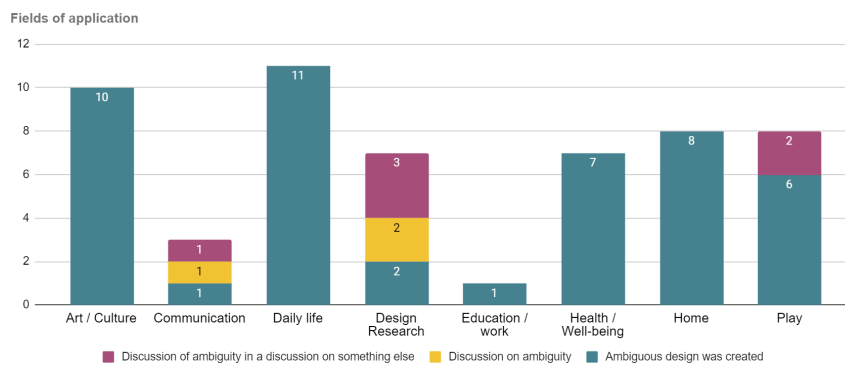


Figure 4: Fields of application with the way in which ambiguity is reviewed in the papers.

2.3 Systematic Literature Review

Ambiguous design is design that is open to interpretation of the user. Ambiguity can be a vague concept that is hard to grasp. In this section we give clarity on this concept by explaining how ambiguous design is used and created on the basis of our corpus.

Copyright clarifications. We provide examples of the ambiguous designs found in the corpus along with photos of their work. Important to note is that we do not hold copyright for the photos that are used as an example. However, because it provides support and a clear image of ambiguous design we included them in the thesis. In the caption of the concerned figures, we provide a reference to the original work. If this work will be used for publication, the original researchers should be contacted and asked for permission.

2.3.1 Ambiguity as a method

Ambiguity is not only created in design, it can also be purposely created in a method to come up with new designs. In this section we describe three techniques that make use of ambiguity: cultural probes, design documentaries and invisible design. The latter two both make use of film.

Probes. A technique that is often used in methods and that uses ambiguity is the cultural probe (Jeng et al., 2011; Karyda et al., 2020; Lockton et al., 2020; Noortman et al., 2019; Sanches et al., 2019; Sengers et al., 2008; Wright and McCarthy, 2008). The cultural probe is developed by Gaver et al. (1999). Cultural probes are packages that can include any sort of artifact (i.e. camera, postcard, map) along with evocative tasks that are given to participants in order to elicit different types of qualitative information from them (Gaver et al., 1999). In a review of empathy, cultural probes were seen as a technique that can elicit empathy through dialogues and can be used to support empathic engagement between designers and users (Wright and McCarthy, 2008). Cultural probes count many adaptations such as design probes, object probes, self-probes, fictional probes, which we will shortly describe – more adaptations exist, but were not found in the corpus. In design probes participants are asked to do something with a design or prototype. Jeng et al. (2011) used a design probe that included tasks to find a purpose for a design prototype by asking the participant to make photos of the device and give answers to the questions about the device. In object probes, participants are asked to show objects they already have, such as their most cherished object, which can give researchers insight into people’s social relationship with and through (personal) objects (Karyda et al., 2020). Self-probes can be used as a tool to support self-inquiry and autoethnography (Lockton et al., 2020). A fictional probe combines design fiction and cultural probes, where elements of a near future world and a prototype that feels as a professional product are created for participants to enact in and with (Noortman et al., 2019).

Film. A technique often used in HCI is the creation of personas and scenarios. In a design study by Philips the creation of personas through photo and text did not seem believable to people (Raijmakers et al., 2006). To tackle this issue, Raijmakers et al. (2006) created design documentaries from these personas. Design documentaries aim to inform and inspire design by getting access to the inconsistency of everyday life and understanding it in its own turn. Ambiguity plays an important role in design documentaries, since it is also present in daily life, along with paradoxes. Another technique that uses film and builds on the personas and scenarios is invisible design (Briggs et al., 2012). Using this technique, a scenario about a new design is presented through film, but the design itself is not shown, hence, invisible design. By using this technique, the films were meant to be ambiguous which would create space for participants to generate new ideas about the designs. Additionally the ambiguity of the film was less prone to becoming dated, which could happen



Figure 5: MStoryG, created by [Jorge et al. \(2013\)](#), which is an airport split-flap display that is connected to Twitter. It is used as an Exquisite Corpse, which is a collective storytelling technique, to which passersby can contribute.

quickly with a prototype. In [Briggs et al. \(2012\)](#) their study, many participants were willing to play with the ambiguity in invisible design. However, for the older participants the ambiguity provoked frustration, although these changed to a more positive thought about the underlying idea later on.

2.3.2 Goals and effects of ambiguous design

Researchers have used ambiguity in order to evoke different effects in the users. In this section we discuss the main effects as found in the corpus, which are: curiosity, social engagement, play, immersion, awareness, understanding, reflection and creativity.

Curiosity. Ambiguity has been found to evoke curiosity in users ([Benford et al., 2012](#); [Jorge et al., 2013](#); [Li et al., 2018, 2020](#); [Rogers et al., 2010](#); [Ryding, 2020](#); [Seitinger et al., 2010](#); [Seok et al., 2014](#); [Wensveen et al., 2004](#)) and stimulate curiosity-driven engagement ([Back et al., 2017](#); [Jorge et al., 2013](#); [Li et al., 2020](#); [Mueller et al., 2012](#); [Sanches et al., 2019](#); [Seitinger et al., 2010](#)). [Jorge et al. \(2013\)](#) created MStoryG, an airport split-flap display that was used in a context other than the airport (Figure 5). This created ambiguity and made passersby curious what it was doing there, this in turn made them interact with the design.

Social engagement. The engagement created by ambiguity and curiosity in the users was also found to be social-oriented ([Cosley et al., 2008](#); [Dagan et al., 2019](#); [Devendorf et al., 2016](#); [Li et al., 2020](#); [Nagargoje et al., 2012](#); [Noortman et al., 2019](#); [Sokoler and Svensson, 2007](#)). The threshold to social engagement can be high for some, especially when one feels lonely ([Sokoler and Svensson, 2007](#)). Ambiguity can help lead people towards social engagement in an indirect way. [Sokoler and Svensson \(2007\)](#) designed an extended TV remote, PresenceRemote, that would make it possible for elderly to get in contact with other people present in their building. In PresenceRemote people from the same building can be added as buddies. It displays what channel is currently watched the most and if their added buddies were also watching TV. It also facilitates a button to get in contact with the added buddies. Because of the stigmatization of loneliness, the aim was to not force people explicitly into social contact by leaving room for different intentions: “with the PR

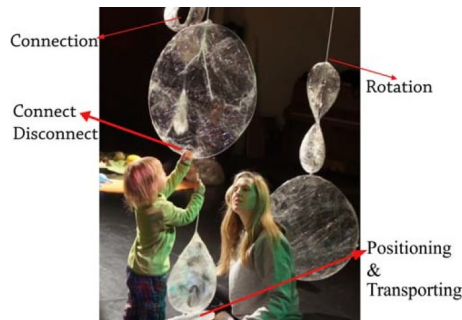


Figure 6: [Patel et al. \(2018\)](#) created Suspended Discs that are flexible, detachable, flat, lightweight, structures that can be hung up and detached by magnets.

we deliberately allow people to leave it unarticulated whether they in fact are watching TV as an excuse to meet others” ([Sokoler and Svensson, 2007](#)).

Play. In discussion about designing for play researchers concluded that ambiguity plays an important role ([Back et al., 2017](#); [Pichlmair et al., 2017](#)). Researchers have created ambiguity in design with the goal to evoke play in its users ([Creighton, 2010](#); [Gaver et al., 2006](#); [Håkansson and Gaye, 2008](#); [Li et al., 2020](#); [Lindley et al., 2009](#); [Patel et al., 2018](#)). When thinking of play, one might directly think of game designs. Indeed, in our corpus we find papers that developed games in which ambiguity was incorporated ([Creighton, 2010](#); [Li et al., 2020](#); [Patel et al., 2018](#); [Mendes et al., 2012](#)). [Patel et al. \(2018\)](#) created a variety of childrens’ games, and interestingly their ambiguous design, ‘Suspended Discs’, was found to be the most popular out of all the designs (Figure 6). The ambiguity ‘freed’ the design and made it possible for participants to use their own imagination. With this in mind, [Creighton \(2010\)](#) used ambiguity to create more free play in their game.

The importance of stimulating play was also found in papers where the focus was not on creating games, but on creating interaction design for museums ([Cosley et al., 2008](#)), for homes ([Desjardins et al., 2020](#); [Lindley et al., 2009](#); [Wensveen et al., 2004](#)), for public spaces ([Håkansson and Gaye, 2008](#); [Seitinger et al., 2010](#)) or wearable design ([Wright et al., 2008](#)). In the study by [Seitinger et al. \(2010\)](#), ‘Light Bodies’ were created (Figure 7). During a performance, people were handed Light Bodies and they quickly started playing with the devices. In later experiments, [Seitinger et al. \(2010\)](#) encouraged their play more by adding ambiguity as they installed random light patterns to the Light Bodies and placed them around a piano. This provoked people into playing with the device by tapping, sharking and re-orienting it. More studies from the corpus have confirmed that



Figure 7: Light bodies, created by [Seitinger et al. \(2010\)](#), are blocks of color changing lights that react differently to sound and can be carried anywhere. Here we see the light bodies placed on stage next to dancers’ personal objects.



Figure 8: SKIN is an interactive dance piece created by [Fdili Alaoui \(2019\)](#). The figure shows two dancers wearing sensors with a video playing in the background. The sensors capture muscle activity, heartbeat and touch data, and when triggered, the video and sound are either frozen or changed in frequency.

play was either evoked or stimulated in the user through ambiguity ([Cosley et al., 2008](#); [Creighton, 2010](#); [Håkansson and Gaye, 2008](#); [Li et al., 2020](#); [Lindley et al., 2009](#); [Mendes et al., 2012](#); [Patel et al., 2018](#); [Seitinger et al., 2010](#); [Wensveen et al., 2004](#)).

Immersion. During an interaction with a device, a user can become fully immersed, which can prompt the state of experience to continue. Ambiguous design has been found to create a more immersive experience ([Benford et al., 2012](#); [Dagan et al., 2019](#); [Fdili Alaoui, 2019](#); [Papworth, 2010](#); [Vidyardhi et al., 2012](#)). In an interactive dance piece, SKIN, ambiguity is created between dancers and videos that are shown behind the dancers (Figure 8). The dancers wear sensors that capture muscle activity, heartbeat and touch data, when triggered, the video and sound is either frozen or changed in frequency. This ambiguity made the dance piece into a more immersive and poetic experience.

Awareness. In a review about awareness, ambiguity is seen as a key requirement for design for creating awareness ([Hassenzahl et al., 2012](#)). The main advantage of ambiguity is that it diminishes the feeling of being monitored which is inherent to most approaches to creating awareness. This indirect approach which ambiguity makes possible seems to be a common thread, as we previously discussed how social engagement can be created with the use of ambiguity without putting emphasis on it. In practice researchers have made ambiguous designs that created awareness in users ([Fdili Alaoui, 2019](#); [Jeng et al., 2011](#); [Mendes et al., 2012](#); [Sengers et al., 2008](#)). In order to create environmental awareness, ‘Play With Fire’ was made. In Play With Fire, a forest is projected on a screen on which users can create fire, this fire then spreads and users have the task to keep the fire under control. This design was made to be ambiguous, but the researcher did not further explain the ambiguity

Create own understanding. Offering multiple interpretations to users is inherent to ambiguous design when following the definition of ambiguity as described in section 2.1.2. A next step is for users to truly interpret design differently and create different meanings or experiences – ie. understanding. According to [Pichlmair et al. \(2017\)](#), ambiguity is key to allowing users



Figure 9: The Lega, created by [Laaksolahti et al. \(2010\)](#), is a hand held device that lets visitors to an art gallery share their experience via physical traces such as light and vibration. By touching or moving the devices, users can create expressions that are left at the location.

to appropriate a game and by doing so, players can create their own experience, expanding on the original design. Researchers have strived for to create design that makes users create their own understanding of it ([Desjardins et al., 2020](#); [Devendorf et al., 2016](#); [Gaver et al., 2006, 2009](#); [Håkansson and Gaye, 2008](#); [Höök, 2006](#); [Laaksolahti et al., 2010](#); [Rasmussen et al., 2013](#); [Rogers et al., 2010](#); [Sanches et al., 2019](#); [Seevinck and Edmonds, 2009](#)). [Sanches et al. \(2019\)](#) notes that once these understandings are made, it is difficult to create new ones. A design that was made ambiguous and offers users to create their own meaning is ‘Lega’ (Figure 9). Lega is a device that lights up and vibrates and is created to be used in a museum. Users can leave light and vibrations behind for other users at different places within the museum. The meaning of these traces are left open for the users, which leads them to creating their own interpretation of the traces.

Reflection. Researchers have stated in discussion about ambiguity, that ambiguity can support reflection in design ([Mols et al., 2016](#)). [Mols et al. \(2016\)](#) stated that ambiguous visualization of data can be a content trigger for reflection that provides direction. This theory has been applied in practice where ambiguity was created in design to evoke reflection. In qualitative research it is shown that users were reflective during or after interaction with ambiguous design ([Bae et al., 2020](#); [Karyda et al., 2020](#); [Lee et al., 2015](#); [Li et al., 2020](#); [Merrill and Cheshire, 2017](#); [Normark and Tholander, 2014](#); [Noortman et al., 2019](#); [Núñez-Pacheco and Loke, 2014](#); [Rasmussen et al., 2013](#); [Rogers et al., 2010](#); [Sengers et al., 2008](#); [Wright et al., 2008](#)). A playful design, InsideOut, shows users intestines in an ambiguous way by transforming the video based on the environmental data (Figure 10). This created self-reflection in users, but did not necessarily lead to a proper understanding of the body. This shows that ambiguous design should not be implemented to teach factual information per se, but rather to make users interpret information in their own way without the judgment of it being right or wrong. In other cases design was made ambiguous in order to evoke reflection, but failed to do so ([Cosley et al., 2008](#); [Leahu et al., 2008b](#); [Normark and Tholander, 2014](#)). An explanation on why ambiguous design did not work as hoped for was that it did not meet the expectation the users had of the design, which made them frustrated or confused. We will further discuss this later in this section.

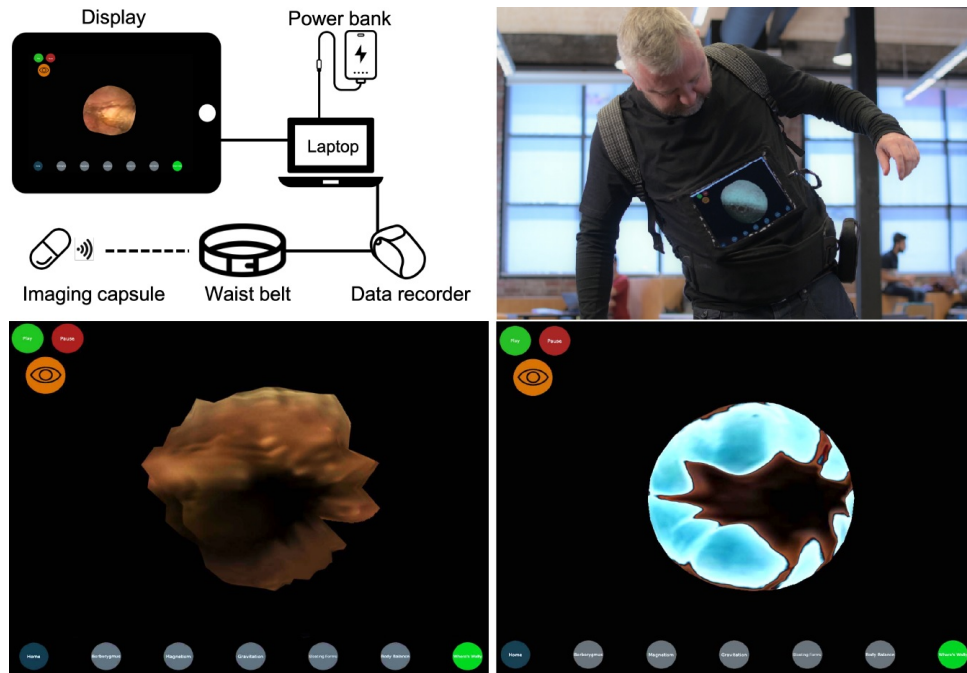


Figure 10: InsideOut, created by [Li et al. \(2020\)](#), is a system that shows your intestines by using a SmartCapsule Endoscopy System and a tablet (1,2). On the tablet the user can enable a variety of modes. Two modes, ‘gravitation’ (3) and ‘magnetism’ (4), embrace ambiguity and environmental data; these modes transform the video based on the surrounding magnetic field’s strength and gravitational acceleration respectively.

Creativity. When design is ambiguous, a more open design is created, which leaves room for creativity. Ambiguous design has proven to stimulate creativity in users ([Bae et al., 2020](#); [Briggs et al., 2012](#); [Håkansson and Gaye, 2008](#); [Jeng et al., 2011](#); [Leahu et al., 2008a](#); [Lockton et al., 2020](#); [Patel et al., 2018](#); [Raijmakers et al., 2006](#); [Seevinck and Edmonds, 2009](#)). Especially for the use of ambiguity in methods, as discussed previously, an important goal can be to evoke creativity in order for participants to contribute more in sessions ([Briggs et al., 2012](#); [Jeng et al., 2011](#); [Lockton et al., 2020](#); [Raijmakers et al., 2006](#)). Creativity can also be needed in more mundane settings, think of brainstorm sessions or even research itself. An example that support this, is the ‘Spinneret’, a mind mapping tool that suggests ambiguous connections ([Bae et al., 2020](#)). In their research, the enabled creativity made it possible for users to create more diverse mind maps.

2.3.3 Finding the right balance

In the previous sections we presented the advantages ambiguity can bring. However, as [Gaver et al. \(2003\)](#) once wrote: “Ambiguity is not a virtue in itself, nor should it be used as an excuse for poor design. Many ambiguous systems are merely confusing, frustrating, or meaningless”. In this section we will discuss how ambiguity can create confusion and what research is done for finding the right balance.

Confusion. Indeed, ambiguous design can be difficult to create and a number of studies have shown that it can create confusion in users ([Cosley et al., 2008](#); [Gaver et al., 2009](#); [Leahu et al.,](#)

2008b; Normark and Tholander, 2014; Papworth, 2010). Interestingly, of these studies, some also show that what might be confusing for one user, could be eye-opening for others (Cosley et al., 2008; Leahu et al., 2008b; Normark and Tholander, 2014; Papworth, 2010). Additionally, when design is perceived as confusing in an early stage of interaction, in time it can still turn into the intended result. For instance, in iSpooks, an audio focused story game, the created ambiguity was found to be confusing, but it still led users to the intended immersive experience (Papworth, 2010). Of course confusion can still lead to frustration or annoyance in users and this is not something we should diminish. Finding the right balance of ambiguity and clarity within design is important, making design too ambiguous can create confusion and frustration, but the other downside is that too much clarity in design might not evoke the topics discussed in previous sections.

Balance. Finding the right balance is a popular topic of discussion in research on ambiguous design (Cosley et al., 2008; Fdili Alaoui, 2019; Gaver et al., 2006, 2009; Håkansson and Gaye, 2008; Höök, 2006; Noortman et al., 2019; Normark and Tholander, 2014; Sanches et al., 2019; Wright et al., 2008). This right balance has been acknowledged as *evocative balance*, and is defined as an experience of an interaction that can be made with a familiar experience, while at the same time be ambiguous enough to allow for new interpretations (Sanches et al., 2019). Creating balance is not an easy task to fulfill as it is difficult to know in advance how ambiguity will be perceived (Fdili Alaoui, 2019).

2.3.4 Different types of ambiguity

The first recognition of ambiguity as a design resource, according to our systematic review, has been by Gaver et al. (2003). In their review they study four different ‘contemporary’ art pieces (from 1999-2001). Based on this, they distinguish three types of ambiguity that can be created: ambiguity of relationship, context and information. Furthermore, they suggest design tactics for creating the specific types of ambiguity. In a later paper, ambiguity of semantics is acknowledged in addition to the other three ambiguity types (Gaver et al., 2006). The different types of ambiguity serve as a guideline for creating ambiguous design (Gaver et al., 2003). Additionally, multiple design types can be represented in design (e.g. designs by Jorge et al., 2013; Papworth, 2010).

In this section we discuss the different types of ambiguity acknowledged by Gaver et al. (2003). We connect these to other discussions on and studies of ambiguous design and bring light to possible other types of ambiguity. Furthermore, we introduce the topic of design tactics, but we present a more in depth discussion of the design tactics in the next sections (Sections 2.3.5, 2.3.6)

Ambiguity of Information. Ambiguity can arise from the way information is presented. Ambiguity of information can make people project their expectations into an interpretation of incomplete information (Gaver et al., 2003). Gaver et al. (2003) suggest four tactics for enhancing ambiguity of information: 1) use imprecise representations, 2) over-interpret data, 3) expose inconsistencies, 4) cast doubt on sources. Inspired by Gaver et al.’s review, Boehner and Hancock (2006) analyzed ambiguity as a design resource for designing space for stories in personal communication systems. They suggest a variety of design tactics that would help with creating ambiguity. Although Boehner and Hancock (2006) do not explicitly label ambiguity of information, they suggest design tactics that deal with conveying information – 1) design for opposites, 2) design for convention building, 3) design for something else, 4) design for extremes, 5) design for over-interpretation. These design tactics however, have not been applied in ambiguous designs from our corpus. We discuss this further in the next sections that present the design tactics.

In different studies ambiguity of information was created in designs (Devendorf et al., 2016; Jorge et al., 2013; Lee et al., 2015; Mueller et al., 2012; Normark and Tholander, 2014; Papworth, 2010). As an example of ambiguity of information, Jogging over a Distance is a system that makes it



Figure 11: Jogging over a Distance, created by [Mueller et al. \(2012\)](#), seen in action. Two women are running at different places while being connected via headsets. If they do not meet their optimal heartbeat, the connection will fade away.

possible for joggers from different fitness levels to exercise together over a distance (Figure 11). When joggers have an optimal heart rate during their exercise, they will be able to communicate clearly. However, if one of the joggers does not match the optimal heart rate the audio for communication will fade. By representing exertion in low resolution (the audio), ambiguity of information is created. The researchers however, recognized ambiguity of information as a design tactic and not a design type. This indicates an issue of documentation in studies on ambiguous design. Instead, the use of imprecise representation can be seen as a design tactics, that creates an ambiguity of information.

Ambiguity of context. Ambiguity of context arises when things are understood in different contexts that suggest different meanings ([Gaver et al., 2003](#)). [Gaver et al. \(2003\)](#) suggest three tactics for creating ambiguity of context: 1) implicate incompatible contexts, 2) add incongruous functions, 3) block expected functionality. In various studies ambiguity of context was created in designs ([Jorge et al., 2013](#); [Karyda et al., 2020](#); [Leahu et al., 2008b](#); [Rasmussen et al., 2013](#)). In MStoryG, an example shown previously (Figure 5), ambiguity of context is created. A split-flap display is designed to perform a single well-defined task in a specific place – i.e. at the airporting to show arrivals and departures. By removing the split-flap from its original context, a sense of ambiguity is derived from its new purpose.

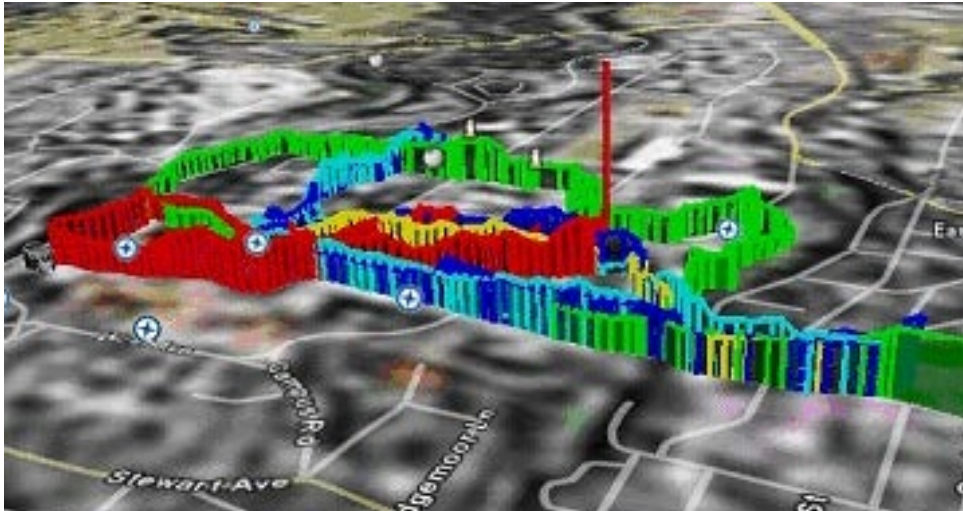


Figure 12: Arousal Map created by [Leahu et al. \(2008a\)](#). The maps visualize paths taken by users along with their arousal data. Each colored wall presents the path of one user and the height of the wall presents the arousal data. Their design was inspired by research on biomapping ([Nold, 2004](#)).

Ambiguity of relationship. Ambiguity of relationship, or ambiguity of interaction ([Jeng et al., 2011](#)), creates the condition for personal projection of imagination and value to a design ([Gaver et al., 2003](#)). [Gaver et al. \(2003\)](#) suggest three tactics to provoke ambiguity of relationship: 1) introduce disturbing side effects, 2) point out things without explaining why, 3) offer unaccustomed roles. In different studies ambiguity of relationship was created in designs ([Benford et al., 2012](#); [Gaver et al., 2006](#); [Papworth, 2010](#)). ISpooks, mentioned previously (Section 2.3.3), created ambiguity of relationship in characters of a story ([Papworth, 2010](#)). In this story, ghosts were created to be more sympathetic and have human traits instead of being ‘typical’ screaming, homicidal banshees. The intention of their design was to keep the user slightly off balance and unsure of the nature of the potential threat.

Ambiguity of semantics. Ambiguity of semantics was later acknowledged, in addition to the three ambiguity types presented above ([Gaver et al., 2006](#)). Semantic ambiguity refers to the semantic mapping between a system and its environment. The mapping itself is specified, but an ambiguity around its implications is created to undermine a simple interpretation. [Gaver et al. \(2006\)](#) suggests a design tactic that embodies representations of external situations, but subverts the authority of these representations. [Leahu et al. \(2008a\)](#) created semantic ambiguity in their design, Arousal Map (Figure 12). The Arousal Map represents the measurement of arousal of users while they walk different paths in a city. In general maps have one simple interpretation, however, by presenting this data in a map viewers are invited to interpret the data themselves.

Confusion about ambiguity types. As mentioned above, documentation issues arise in studies on ambiguity, as happened in the research by [Mueller et al. \(2012\)](#). It is important to not confuse ambiguity types with design tactics for creating ambiguity. The ambiguity types are merely different classifications of ambiguities that can arise in design, while design tactics can help with creating or emphasizing ambiguity. It might not always be clear what type of ambiguity is found in the design and only in the papers mentioned above the design types were mentioned, meaning

that a large part of our corpus did not discuss this. We question whether this is an issue when the design is already created, since the different types of ambiguity are there to serve as a guideline for designing. Nonetheless, it is important to recognize the theorized design types and not to confuse them with design tactics.

2.3.5 Suggested design tactics and strategies

In discussion about ambiguity a variety of design tactics are suggested. In the previous section we show design tactics created by [Gaver et al. \(2003\)](#) and [Boehner and Hancock \(2006\)](#) that are connected to a specific design type. In addition to design tactics, strategies were also suggested for designing to create open interpretation by 1) clearly specify usability, without leaving interpretation open 2) support a space of interpretation around a given topic, 3) stimulate new interpretations by purposefully blocking expected ones, 4) gradually unfold new opportunities for interpretation over the course of interaction, 5) make space for user reinterpretation by downplaying the system's authority and 6) thwart any consistent interpretation ([Sengers and Gaver, 2006](#)).

Ambiguous designs that we studied in the papers of our corpus do not directly refer to these tactics and strategies. An explanation for this is that the tactics and strategies are written on a higher level of abstraction. This higher level of abstraction can make it more difficult for researchers to directly implement the design tactics or strategies in practice.

2.3.6 Concrete design tactics

In this section we present design tactics that are based on ambiguous designs from the papers in our corpus (Table 10 in Appendix A). These design tactics are distinctive from the suggested design tactics and strategies in two ways. First, the design tactics are based on how ambiguous design was created in practice. Second, because the design tactics are based on the created ambiguity in designs, they are presented on a lower level of abstraction. This makes it easier to create ambiguity in the future and to document what aspects of the design are ambiguous and how the ambiguity is created. In total we present ten design tactics: absence of purpose, absence of obvious connection, absence of information, absence of object, connection of external situations, abstract representation, data physicalization, ambiguous data, and uncomfortable design.

Create an absence. Ambiguity can arise by creating an absence of something in the design. Closely related to this is the concept of non-finito products, which was introduced in HCI by [Seok et al. \(2014\)](#) and draws on studies about ambiguous design. Non-finito products in HCI are intentionally unfinished products that foster new creations by end-users. Although a non-finito product seems to be a suitable name for a design in which an absence is created, we did not find the term in other papers from our corpus.

Absence of purpose. An absence of purpose creates room for users to come up with their own interpretation of the design's purpose. Different studies have left out the purpose of a design in order to make it ambiguous ([Jeng et al., 2011](#); [Laaksohlahti et al., 2010](#); [Leahu et al., 2008a](#); [Lockton et al., 2020](#); [Nagargoje et al., 2012](#); [Patel et al., 2018](#); [Ryding, 2020](#); [Sanches et al., 2019](#); [Seitinger et al., 2010](#); [Sokoler and Svensson, 2007](#)). As an example, see the designs that were discussed before, PresenceRemote and Lega (Figure 9). In PresenceRemote, the purpose of the design is held absent in order to tackle their goals of creating social engagement. This in turn can also allow the users to leave their intention of using it unarticulated, whether they in fact are watching TV with the PresenceRemote as an excuse to meet others. In Lega, the purpose of the traces was unknown and left to the user to interpret. In the use of cultural probe, an absence of purpose can also be created to make participants to the study find a purpose for a design ([Jeng et al., 2011](#)).



Figure 13: The alarm clock designed by [Wensveen et al. \(2004\)](#). Each slider can create a different sound. All sliders together form a variety of patterns, (a)symmetrical or (dis)satisfying, depending on the user and their mood.

Absence of obvious connection. Non-obvious connections in design ask for users to make up their own connection. This design tactic is applied in various designs for creating ambiguity ([Bae et al., 2020](#); [Fdili Alaoui, 2019](#); [Vidyardhi et al., 2012](#); [Wensveen et al., 2004](#)). An example that makes use of this design tactic is the alarm clock designed by [Wensveen et al. \(2004\)](#) (Figure 13). This alarm clock has multiple sliders that create different sounds and patterns. This allows for many different sound patterns to be discovered, as the connection between the sliders and the sound are hidden.

Absence of information. An absence of information can take the form of instructions about the design. The absence of information was used in multiple ambiguous designs ([Cosley et al., 2008](#); [Creighton, 2010](#); [Gaver et al., 2009](#); [Noortman et al., 2019](#); [Normark and Tholander, 2014](#); [Patel et al., 2018](#); [Papworth, 2010](#)). As an example, [Creighton \(2010\)](#) created a large step-sequencer, Jogo, designed for children to generate music (Figure 14). Jogo lacks physical representation of the digital content, which makes an absence of information in the form of instructions.



Figure 14: Jogo, created by [Creighton \(2010\)](#), is a large step-sequencer designed for children to generate music. The circular surface consists of a grid with sixteen rows of holes of which each row represents a sixteenth note in a measure of music. When balls are placed in the holes, samples are played, with each colored ball representing a different sample.

Absence of (original) context. The absence of (original) context can be used in different ways to create ambiguity. The absence of original context is applied in different ambiguous designs; either the context can be completely removed, or design itself can have an absence of original context and be placed in a variety of different contexts (Jorge et al., 2013; Leahu et al., 2008b; Rasmussen et al., 2013). An example of design in which the context is removed completely is The Situationist Agent (Leahu et al., 2008b). The Situationist Agents creates collages of images and texts that are presented out of context, which makes them lose their original meaning. Another way to use this tactic is by creating a design with no original context and placing it in a variety of different contexts. A moving bench that did not contain an original context was created by Rasmussen et al. (2013), the coMotion Bench can move up and down in different angles while people sit on it (Figure 15). Multiple interpretations arise from this; in an airport setting people who sat on the bench referred to the movement of the bench as preparation for the flight, while in a concert hall it was seen as an amusement ride.



Figure 15: The coMotion Bench, created by Rasmussen et al. (2013), is a shape-changing bench that can change heights in different ways. When someone sits down on the bench, in time it will provide a small vertical push under the person.

Absence of object. This design tactic was used in a film method, which we explain above (Section 2.3.1). In their film a new design is presented, but the object itself is not shown, to elicit participants' ideas about the design. This way ambiguity about the design is created.

Abstract representation. An abstract representation can create ambiguity. Abstract design can create a more personal and non-prescriptive representation (Rogers et al., 2010). A non-prescriptive representation relates to the suggested design tactic of not specifying how people should relate, which offers room for people to create their own interpretations (Sengers and Gaver, 2006). Different designs were made ambiguous using an abstract representation of data (Desjardins et al., 2020; Fajardo and Moere, 2008; Lee et al., 2015; Höök, 2006; Dagan et al., 2019; Devendorf et al., 2016; Núñez-Pacheco and Loke, 2014; Rogers et al., 2010; Sengers et al., 2008; Sanches et al., 2019). An issue that arises in the documentation of the different papers is that the terms 'abstract' and 'ambiguous' are used alongside each other, without clearly explaining the differences or the relationships between the two (Devendorf et al., 2016; Fajardo and Moere, 2008; Höök, 2006; Li et al., 2018, 2020; Lee et al., 2015; Ståhl and Höök, 2008). We describe abstract design as defined by the dictionary, as simplified design that does not attempt to depict visual reality. Ambiguity can in turn arise in abstract design by leaving room for interpretation.

Using abstract data representation closely relates to the design tactic of absence of obvious connections, as not explaining the connection between the data and presentation of it emphasizes



Figure 16: EloquentRobes, designed by [Núñez-Pacheco and Loke \(2014\)](#), is an interactive art installation that uses heartbeat data to communicate intimate information. The user wears a white robe made of paper and a heartbeat sensor. They sit down in a dark room while their heartbeat data is abstractly presented on them using a projector.

the ambiguity of the design. As an example, the design ExternalEyes is a smartwatch that measures skin conductance and presents this in an abstract way ([Fajardo and Moere, 2008](#)). Higher levels of skin conductance, i.e. stress, are chaotically displayed in lines on the screen. When users do not know the connection between the lines and the data connected to it, the design remains ambiguous.

Abstract representations can be learned. For instance, ExternalEyes (Figure 17) was specifically designed to be decipherable only over a relatively long period. This indicates that once a user makes a connection in the design, it becomes less ambiguous to them. In addition, it seems that once an interpretation of ambiguous design is made, it is difficult for users to create new interpretations ([Merrill and Cheshire, 2017](#); [Sanches et al., 2019](#)). Nonetheless, this occurrence does not make the design itself less ambiguous, although the design will appear less ambiguous to the user after they interpret it.



Figure 17: ExternalEyes, created by [Fajardo and Moere \(2008\)](#), is a watch that measures the amount of stress of a person. It visualizes this in an abstract way by displaying lines that appear more chaotic when the user experiences stress.

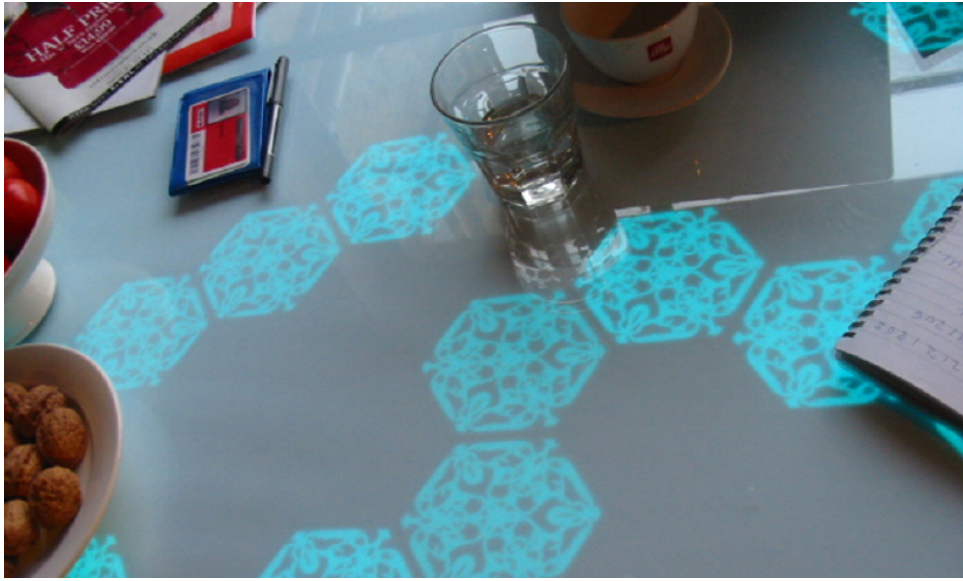


Figure 18: The History Tablecloth is created by (Gaver et al., 2006). The blue hexagons are lit up because something was placed on it 15-30 minutes before the photo was taken.

Data physicalization. Data physicalization can emphasize ambiguity that is already there (Desjardins et al., 2020; Devendorf et al., 2016; Jorge et al., 2013; Karyda et al., 2020; Núñez-Pacheco and Loke, 2014). In a cultural probe study, data physicalization was used where data was coupled to physical objects. The coupling itself was unexpected, which similarly to abstract data representation, shows a connection to the design tactic of absence of obvious connection. Another example where physicalization creates ambiguity is, Eloquent Robes, a design which visualizes data in a pattern of colored dots on a white dress that is worn by the user (Núñez-Pacheco and Loke, 2014) (Figure 16). The dress is made of paper material which is generally not seen as wearable or durable. Using this material Núñez-Pacheco and Loke (2014) aimed to generate a certain ambiguity regarding the nature of the garments as disposable.

Ambiguous data. Not only does the presentation of data create ambiguity, data itself can also be ambiguous by default. For instance, heartbeat data can be interpreted by users in different ways (Merrill and Cheshire, 2017). In EloquentRobes the goal was to interpret heartbeat data and the abstract visualization and physicalization of the design might have emphasized that ambiguity (Núñez-Pacheco and Loke, 2014). Other designs in our corpus that made use of heartbeat data did not address the ambiguity of the data, which can be explained with the fact that they already connected the heartbeat data to a specific task, whether or not the connection was obvious (Fdili Alaoui, 2019; Núñez-Pacheco and Loke, 2014).

Connection of external situations. This design tactic was suggested in research by Gaver et al. (2006) who proposes that ambiguity can be created by creating systems that relate to external situations without indicating a judgment about their meaning. Gaver et al. (2006) implemented this design tactic in the History Tablecloth (Figure 18). When something is placed on the History Tablecloth, the place lights up over a period of hours. This allows users to create their own interpretation.



Figure 19: The context camera is created by [Håkansson and Gaye \(2008\)](#). The different effects are connected to different contexts: shadow is created by movement and the color of shadow connected to frequency of surrounding sound (a), object of movement is zoomed in and transparency of object is connected to level of surrounding sound (b).

We also found this tactic in other designs, although it was presented in different ways ([Håkansson and Gaye, 2008](#); [Li et al., 2020](#); [Mueller et al., 2012](#); [Wright et al., 2008](#)). For instance, by bringing context to the foreground as done in the Context Camera (Figure 19). In this design, sensor data is connected to graphical effects that are applied to images taken by users in real time. The sensors collect data of the environment, while the graphical effects communicate this data to the user, which creates the opportunity for users to play with photography and interpret the photos.

Uncomfortable design. Uncomfortable design creates ambiguity by making the participant question the design itself. [Benford et al. \(2012\)](#) created Breathless which is a swing that moves based on the breath of the user (Figure 20). Users are asked to wear a gas mask that is uncomfortable to wear as it has an overpowering rubbery smell, is hot to wear and restricts visibility of the user. The swing was tested in an amusement park and evoked users to question the object on its comfort and safety.



Figure 20: Breathless, created by [Benford et al. \(2012\)](#), is a swing that moves according to a user's breath. The user puts on a gas mask with embedded breath sensors. When the user inhales, the swing will pull backwards and when the user exhales, the swing will push forwards.

2.4 Summary and contribution

The systematic literature review presents an overview of the understanding of ambiguity in literature. Ambiguity can be created in an artifact, as well as a method. Creating ambiguity in method can be done with probes (Gaver et al., 1999) or specific film making, such as design documentaries (Raijmakers et al., 2006) and invisible design (Briggs et al., 2012). Ambiguity can also evoke a variety of things in the users. For instance, different forms of engagement, such as curiosity (e.g., Jorge et al., 2013), social engagement (e.g., Sokoler and Svensson, 2007), play (e.g., Seiting et al., 2010), immersion (e.g., Fdili Alaoui, 2019). It can evoke awareness (e.g., Hassenzahl et al., 2012), make users create their own understanding of the design (e.g., Laakso et al., 2010) and offer support for reflection (e.g., Mols et al., 2016), or make users more creative (e.g., Bae et al., 2020).

To make sure these goals are met, it is important to find the right balance of ambiguity. When design is too ambiguous, it can create confusion in users (e.g., Cosley et al., 2008). At the same time, it can also evoke a different reactions in the user, where the design is confusing for one user, but eye-opening for others. The other side, if a design is not ambiguous enough, it might not evoke the goals discussed above. Finding this balance can be difficult as it is hard to know in advance how users will perceive the ambiguity (e.g., Fdili Alaoui, 2019).

In literature, different types of ambiguities are recognized: ambiguity of information, context, relationship and semantics (e.g., Gaver et al., 2003, 2006). These types can serve as a guideline for creating ambiguous design, but are not often used in practice. Additionally, along these types a variety of design tactics were proposed (Gaver et al., 2003). These design tactics, however, appear to have a high level of abstractness, which explains why they are not directly found in the other studies of our corpus. A higher level of abstraction can make it more difficult for researcher to directly implement a design tactics in practice. Because of this, we propose more concrete design tactics.

By analyzing all ambiguous designs that are studied in the papers of our corpus, we created a variety of design tactics. Five design tactics create a form of absence in the design: absence of purpose, obvious connection, information, original context, and object. These tactics closely relate to the concept of non-finito products, introduced by Seok et al. (2014). Another way of creating ambiguity in design is by connecting external situations to the design (e.g., Gaver et al., 2006). Ambiguity can also be created by giving an abstract representation. An issue that arises in the documentation about this, is that abstractness and ambiguity are not always kept separated in terms of how it is defined (e.g., Fajardo and Moere, 2008). In our understanding is abstract design a simplification of design, and ambiguity can in turn arise in this abstract design. Abstract design was also found to present a learning curve in users, once users make the connection in the abstract design, it becomes less ambiguous to them. Furthermore, the physicalization of data is also a design tactic that can emphasize ambiguity and data itself can also be ambiguous. Lastly, the creation of ambiguous design can create ambiguity in users by making the participant question the design itself (Benford et al., 2012). Important to note is that ambiguous design can be created using multiple design tactics.

2.5 Our study

The systematic literature review shows that there is still a lot to study when it comes to ambiguous design. The review offers more questions than we can study in this research. In this study we focus on the creation of ambiguous design using four concrete design tactics and their effects on users. The concrete design tactics that we implemented are: absence of information, absence of (original) context and a combination of absence of obvious connection and abstract presentation. We chose these tactics because they touch different aspects of a design: information, context and relationship. Ambiguity changes to these aspects can create different types of ambiguity as discussed by Gaver et al. (2003). Additionally these design tactics were implemented in multiple designs, which makes

them more popular than the design tactics that were only used in one ambiguous design (e.g. absence of object).

We test different effects on the ambiguous design: reflection in users, user engagement, system usability and perceived creepiness. We present the hypotheses in detail in the results section (Section 5.1). We test reflection and user engagement, as this effect was highly presented in the systematic review. Furthermore, we test the usability of the design, because users of ambiguous designs can become confused about the design, when this is not the intention (e.g., [Cosley et al., 2008](#)). Lastly, we test the perceived creepiness to test whether users become uneasy by ambiguous design. Testing perceived creepiness builds further on the design tactics of uncomfortable design. Although we do not use this design tactic, it is still interesting to see whether unsettling feelings are also evoked using other design tactics for creating ambiguous design.

In this research we assume that the prototype we created with the design tactics is ambiguous. Of course this is a topic of discussion, as there is currently no standardized way of testing the ambiguity of a system. Nonetheless, using the systematic review as a guidance is a good starting point.

3 Design

In this section we present the design process and our prototype. First we describe our reasoning behind the design of the prototype. Second we present the different versions and iterations that were made during the design process. Third, we present our final prototype, with the design tactics we chose to implement, a list of hardware and a recovery that needed to be done.

3.1 Forming the prototype

In this section we explain how we move from concept to a prototype. We describe the design tactics we chose to implement in the prototype and explain how we implemented this. We also explain what data we chose to use in our prototype and how the prototype should look.

3.1.1 Chosen design tactics

In our prototype and experiment we implemented different design tactics: absence of instructions, no original context, absence of obvious connection and abstract representation. These design tactics are tested separately in the experiment by creating different conditions. We created a base version of the prototype to which we can add the different design tactics. The base version is also tested without design tactics, as a control condition. The base version is intended to have a simple design of which it is not easily readable what it does and how it works exactly. This way we implemented the absence of instruction as one condition that receives less instructions about the device. Furthermore, it is a device that can appear in any given context, since air quality can be measured anywhere, and thus it does not represent an original context. Lastly, we use a combination of abstract representation and absence of obvious connection. In this version, we created an abstract representation of data to present to the participant. This is the only design tactic that changes the prototype directly.

Condition	Design tactic(s)	Implementation
A	No design tactics applied	<ul style="list-style-type: none"> - Clear instructions on how to use the device - Give example of the context in which the device can be used - <i>Base version prototype</i>: visualization that shows the air quality data in numbers
B	Absence of instructions	<ul style="list-style-type: none"> - Limited instructions on how to use the device - Example of the context in which the device can be used - <i>Base version prototype</i>: visualization that shows the air quality data in numbers
C	No original context	<ul style="list-style-type: none"> - Clear instructions on how to use the device - No example of the context in which the device can be used - <i>Base version prototype</i>: visualization that shows the air quality data in numbers
D	Absence of obvious connection and abstract representation	<ul style="list-style-type: none"> - Clear instructions on how to use the device - Example of the context in which the device can be used - <i>Abstract version prototype</i>: visualization that shows the air quality data in the form of a wave

Table 1: The design tactics we tested in the experiment and how we implemented this in the prototype.

3.1.2 Data to reflect on

The majority of papers we reviewed created designs that presented data in an ambiguous way for users to interpret (Table 10 in Appendix A). In our study we test whether ambiguity influences the user's reflection. To give users a specific goal to reflect on, we decided to let them reflect on certain data. Preferably, the data would be personal, as examples in Bentvelzen et al. (2021) also deal with personal data. However, we decided to focus on more general data, since personal data had some downsides in our case. One downside was the fact that the experiment was conducted during the Covid-19 pandemic, meaning that getting in contact with participants was a difficult task itself. If we were to focus on personal data, we also needed to gather participants that already collected that type of personal data, which would have limited our participant collection. Additionally, if we would use live participant's data, i.e. heartbeat, skin conductivity, we could encounter a hygiene issue. Besides, using live personal data can take up more time and be more prone to errors in the design. Altogether, we decided on using air quality data for users to reflect on. We believe that this data is personal enough for users to reflect on, yet general enough for running the experiment smoothly.

In our brainstorm session we decided to create a design that made it possible for users to see a form of history of the data. This way the design would not only present the data, but it would also make it possible for participants to interact with the design. The opportunity for interaction with the design is important, as the majority of the ambiguous designs in our corpus were also interactive.

3.1.3 Physical design

We decided on creating a physical artifact which users can interact with, because the majority of the papers we reviewed also had physical objects for users to interact with. A physical design made it possible for us to easily connect different sensors to an Arduino to test what would work best. The size of the design should not be too large in order to be able to test it out at different places.

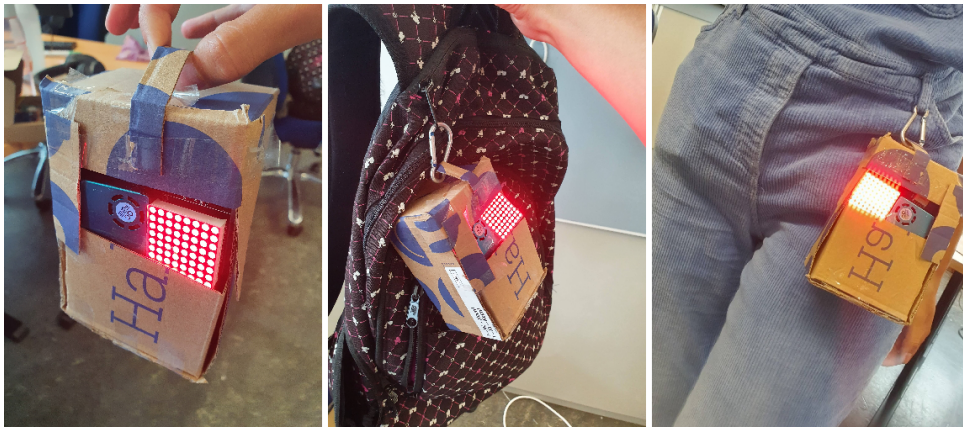


Figure 21: The first prototype, made out of cardboard, consists of a 8x8 red led matrix and air quality sensor that are connected to an Arduino. The hook makes it possible to connect it to different things and take it with you.

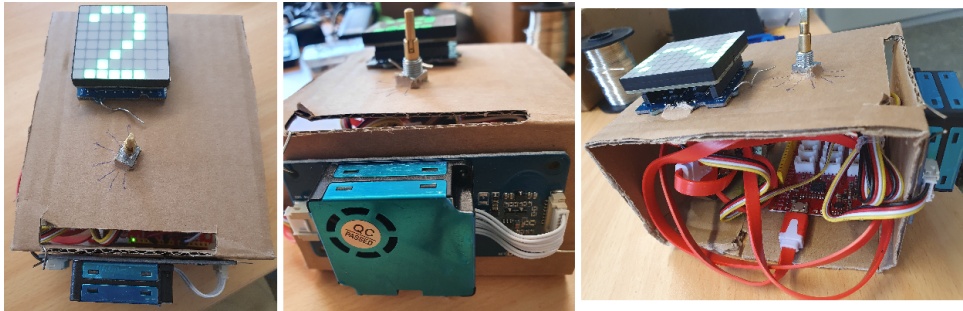


Figure 22: The second prototype, made out of cardboard, consists of an 8x8 rgb matrix, air quality sensor and a potentiometer.

3.2 Design process and review

After we decided to use air quality as data and create a physical object, we started prototyping. First we created a simple prototype with a 8x8 red led matrix and an air quality sensor (Figure 21). This prototype also presents portability, as it is possible to clip it to other things, i.e. a backpack.

In the second prototype we focused on the hardware and how the device would work accordingly (Figure 22). In this prototype we added a potentiometer, with the idea that participants can turn the potentiometer to scroll through the history of the air quality.

In the third prototype we focused on the visual design (Figure 23). This prototype consists of two ‘boxes’ made out of crafting paper. To the first box, all sensors are connected and screwed in place. The second box was wrapped around it, to hide the screws and cables, making it a more clean design. On top of the second box, small squares are glued which all indicate a point of time. We also cut out a pointer in the shape of a leaf and wrapped it around the potentiometer. The leaf-shape is meant as a funny link to air quality, as trees help with cleaning the are.

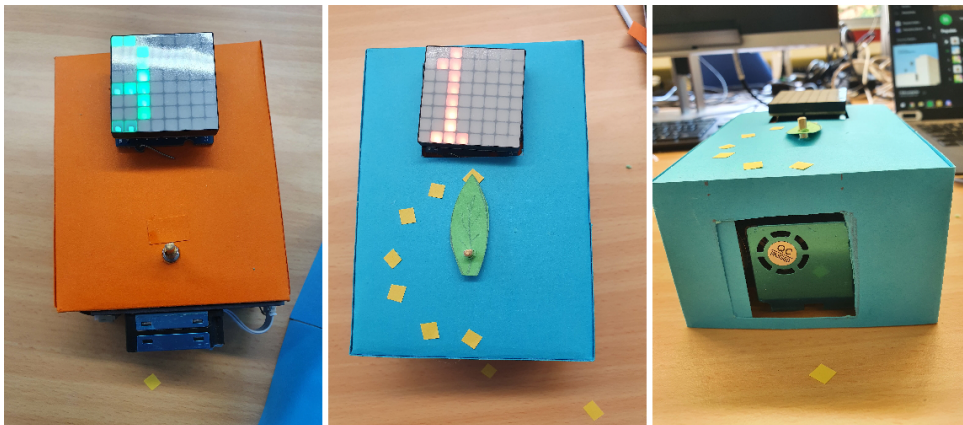


Figure 23: The third prototype, made out of crafting paper. The hardware is connected to an orange box that serves as a base. A blue box is put over it to hide screws and cables.

3.2.1 Review

The third prototype was reviewed by two participants. The participants received full instructions about the device and were asked to give comments on any unclarities. Two unclarities about the device were mentioned. The first one was the inconsistency of the number shown; the live setting showed large scrolling numbers on the screens, while the history settings showed smaller numbers that would fit next to each other without scrolling. We decided to change this for the final prototype and chose to show large numbers with scrolling. This was mainly a convenience choice, as the scrolling numbers were already integrated in the library, which would make coding the Arduino easier. Besides, participants mentioned they did not mind the scrolling numbers. Another unclarity were the indication points for the different settings. One participant mentioned that there was not clear point for the live setting and changing that point to another color would make it easier to understand the device. It was not our intention to be ambiguous about the different settings (live and history), so we created a clear difference in the final design.

3.3 Implementation

The prototype measures air pollution using the Grove Laser PM2.5 Sensor (HM3301). The output of the sensor is shown on an 8x8 Grove RGB LED Matrix w/Driver. The turning knob is a potentiometer with detents with on top a 3D-printed knob-cover shaped like a leaf. The sensor and knob are wired to an arduino SEEED Lotus Corext-M0+ and powered by either a battery or small power bank. Everything is covered by a gray 3D-printed case that leaves an open square for the RGB LED Matrix and the sensor (see Figure 30 in Appendix B for 3D-models). On top, we glued the yellow-orange bead and blue 3D-printed dots on the case forming half a circle.

The final prototype was made using a 3D printer (Figure 24). It is similar to the third prototype and only minor changes were made due to practicality. One change we made was the

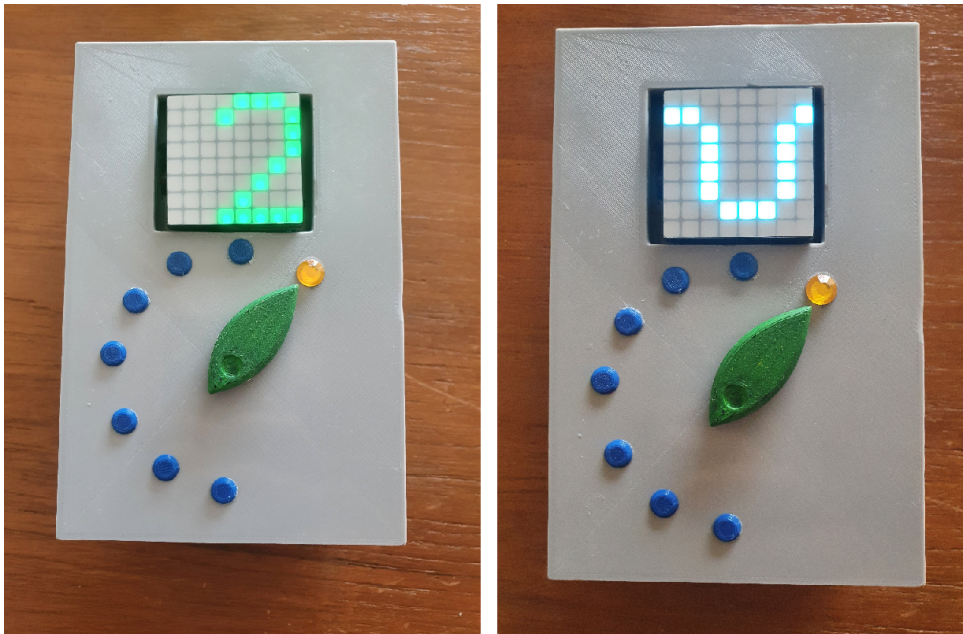


Figure 24: The final prototype was made with a 3D printer. The air quality can be visualized as a number (left) and in an abstract visualization as a wave (right).

size of the device and the holes, due to different material and fitting of the electronics. Another change we made is the way the device opens up: in the third prototype it opened up by pulling the bottom of the device, while the final prototype it can be opened by sliding the bottom open. Furthermore, the colors of the prototype are different compared to the third prototype, because we were limited to the available colors in the 3D printer (gray and blue). Although we could have painted everything, we decided to only paint the leaf-knob. The gray box gave the device a neutral feel to it, while the blue indicators could be seen as a link to air quality. The choice of the yellow-orange indicator happened by accident. We did not manage to paint a blue indicator in another color as the paint would not stick to the small surface. In a box with crafting supplies we found this shiny yellow-orange bead that had the size as the blue indicators. Since orange is a complementary color of blue, this is a nice contrast to the blue indicators. We managed to paint the leaf-knob as its surface is larger. We painted this green to better resemble a leaf.

Although we tested different design tactics, these are not directly seen in the prototype. The prototype itself has two versions, a base which are used for testing conditions A, B and C, and the abstract version which is used in condition D (Figure 24). The base version presents the air quality in numbers. The abstract version presents the air quality in a wave that scrolls faster if the air quality is worse and vice versa. Additionally, in the abstract version the connection between air quality and the pace of the wave is not explained, which creates an absence of obvious connection.

3.3.1 Recovery of the prototype

While conducting the experiment, we encountered an issue. In the experiment, users are allowed to fire matches in order to change the air quality, we further discuss this in Section 4.1.2. Despite the fact we cautioned participants to be careful with the fire and not burn anything or themselves, this still happened. One participant set the prototype on fire, which made the cover melt (Figure 25). Fortunately, the sensor was still intact and all hardware was working accordingly. Since reprinting the entire case would take up too much time, we decided to print a thin piece and glue it to the case of the prototype to cover up the melted part (Figure 25).

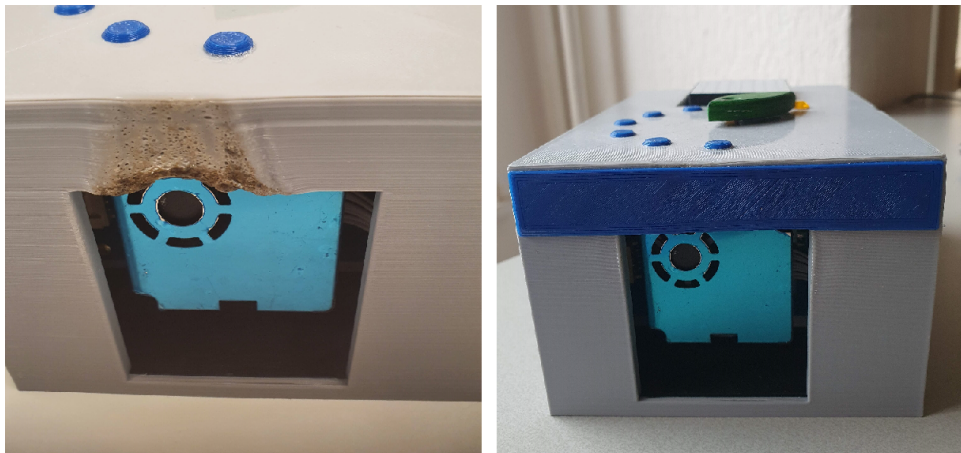


Figure 25: The recovery of the burnt prototype. The prototype was burnt in a way that it melted the plastic (left). We 3D-printed a thin blue piece and glued it on the case to cover up the damage (right).

4 Method

In experiment we study the influence of ambiguous design on the user’s self-reflection, user engagement, the system’s usability and the perceived creepiness of the system. We try to manipulate the ambiguity of the system in three different ways, by using the design tactics: absence of instructions, absence of context and absence of connection combined with abstract visualization. We tested this using between-subjects design in which each tested one condition. Using mixed methods, the participants were asked to fill in a questionnaire and to be interviewed.

In this section we first discuss our experimental study design which includes the different conditions and experiment tasks. Second we describe the measurements that we used that explain the different scales to test self-reflection, user engagement, system usability and creepiness. Third, we describe the procedure of the experiment. Fourth we describe the participants that were recruited. Fifth and final, we describe possible issues that can arise and which we have little control over.

4.1 Experimental study design

We set up the experiment as a between-subjects design. We divided participants in four different groups and each group tested one condition. We used between-subjects design to make sure that there would not be a learning effect, since our literature review showed that it is difficult to create new interpretations, once these are made (Sanches et al., 2019). Another design set-up would simply not have worked as we are using the same device and data to be interpreted by the participant.

4.1.1 Conditions

We created ambiguity by implementing design tactics in our prototype. In total we created four conditions: one control condition and three conditions with different design tactics: absence of information, absence of original context, and a combination of absence of connections and abstract visualization (Table 2). In all conditions the same introduction is given to participants, which informs them about air pollution, PM2.5 and air quality. This way, participants start with the same background information on the data that is presented. Furthermore, all participants receive minimal basic instruction on what the device does, while only conditions A, C, and D receive extra instructions on how the device works. All participants also receive a request to imagine a scenario about the history of the device. However, only conditions A, B, and D receive information about the context about where the device has been.

4.1.2 Experiment tasks

After the participants read the information about the device, they are given two tasks. These tasks are designed to stimulate the user into interacting with the device. The tasks consist of two questions:

1. What is the current air quality?
2. When was the best air quality?

All participants received these tasks and had to give answers in the text box underneath the questions.

Matches. Because air quality can be static, especially in controlled environments – i.e. universities, participants are allowed to explore it in whichever way they like. To stimulate exploration, matches are given to the participant to create smoke and change the air quality.

Information

Conditions

Introduction

You have been given a device called airQualityTime. AirQualityTime shows the PM2.5 value that is generally used for measuring air pollution. PM2.5 are very small particles usually found in smoke and long exposure to high PM2.5 values can be damaging. The table below shows the categorization of the air quality levels per PM2.5 values measured over an hour.

Air quality category	PM2.5 averaged over 1 hr.
Good	< 25
Fair	25-50
Poor	50-100
Very poor	100-300
Extremely poor	> 300

All conditions (A, B, C, D) included an introduction with information about air pollution, air quality and PM2.5.

The airQualityTime device

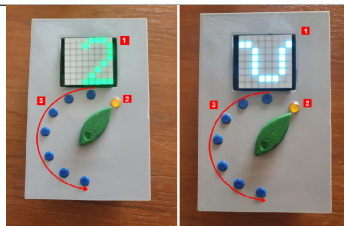
AirQualityTime has two main features:

1. It can display the current air quality
2. It can display the average air quality per hour that was saved over a period of seven hours.

All conditions (A, B, C, D) included basic instruction about the device.

The device has a pixel display that shows the measured PM2.5 value (see 1). Using the knob, which is shaped like a leaf, you can switch to different display views. Turning the leaf to the shining orange dot, the current air quality is displayed (see 2). Turning the leaf to one of the blue dots, the air quality history is displayed (see 3). Each blue dot to the left is one hour backwards in history.

The conditions A, C, and D included extra instruction about the device. It explains how each function works along with the photo below.



The conditions A, C (left) and D (right) included a photo with extra instruction about the device

History

Imagine the following scenario: You have been using airQualityTime and saved the data of different places during the day from 6:00 AM to (and including) 12:00 PM. The device still shows this data in the air quality history.

All conditions (A, B, C, D) included the history scenario.

Imagine you were going on a small trip this morning to explore the city.

- 6:00h - you were in the kitchen at home and ate breakfast there
- 7:00h - you were waiting for the bus at the bus stop going to the city centre
- 8:00h - you arrived at the city centre and walked around for a while
- 9:00h - you were inside in a museum
- 10:00h - you were at the cafeteria of that museum
- 11:00h - you were walking outside in the city centre towards the bus stop
- 12:00h - you were going back home and were waiting at the bus stop

Conditions A, B, D included extra context about where the device was used.

Table 2: The information of the prototype given to the participants, with descriptions of what was included in which condition. Condition A contained all the information.

Participants are, however, not forced to use these matches and can continue the experiment without them.

4.2 Measurements

In accordance with our hypotheses, we use four dependent variables: self-reflection, user engagement, usability and creepiness. Participants in our study had to answer a questionnaire that consists of questions regarding these variables (Appendix C). We present these further in this section.

4.2.1 Self reflection

The self reflection of the participant is tested after the interaction with the device. Participants are presented the Technology Supported Reflection Inventory (TSRI) and the Self-Reflection and Insight Scale (SRIS). The TSRI measures the amount of reflection the device supports. TSRI consists of nine items with a seven point likert scale (Bentvelzen et al., 2021). Because we chose a between-subjects design, different participants interact with different versions, which could influence the self reflection score. To tackle this problem, we use the SRIS focused on self reflection (SRIS-SR) as a baseline, which measures the general self reflection of the participants. The SRIS-SR consists of twelve items using a six-point likert scale (Grant et al., 2002). We test the reflection because the literature review showed that ambiguity can support reflection according to qualitative studies, but no quantitative studies have been done yet.

4.2.2 User engagement

The user engagement with the device is tested using the User Engagement Scale (UES) (O'Brien et al., 2018). We use the short form of the scale which contains twelve items with a five point Likert scale. We test the user engagement since the literature review showed that ambiguity lures people into engaging with a design. However, a quantitative study on this has not been done yet.

4.2.3 Usability

The perceived usability of the system is tested using the System Usability Scale which was created by Brooke (1986). Despite its age, it still remains the most used scale for usability. The scale has proven to be easy for participants, it can be used for small sample sizes ($n > 14$) and is able to test the usability of different technological products, i.e. hardware, software, website etc. (Lewis and Sauro, 2009). The scale consists of ten items with a five point Likert scale. SUS scores above 68 are considered to be above average, nonetheless a system should pass the minimum SUS score 75 to be considered as good usability and above 85 to have excellent usability. We tested the usability of the system, because the literature review showed that ambiguous design does not mean bad design or bad usability, while it also showed cases of design that was perceived as confusing for users. This makes it interesting to test whether ambiguous systems pass the average usability score and if the usability changes when ambiguous design tactics are applied.

4.2.4 Creepiness

Creepiness is tested using the Perceived Creepiness of Technology Scale. This scale consists of 8 items, that are classified into three sub scales ie. implied malice, undesirability and unpredictability (Woźniak et al., 2021). The PCTS is scored on a seven point Likert scale. We used this scale to ensure that our design does not produce the feeling of creepiness in users.

4.3 Procedure

This study was conducted over the course of 11 weeks (23. aug 2021 to 8. nov 2021). This took longer than expected as the researcher got infected with covid-19 and self-isolation was required. It was also not possible to do the experiment online, as it was a physical device.

The experiment was conducted at different locations: the researcher's home, participants' homes, the university and outside. This was also because of covid, as the university was closed for a certain period of time.

Mixed method design. In this study we chose a mixed method design, meaning that we performed a quantitative and qualitative study. In the quantitative study we test the different scales which contributes to ambiguity research as quantitative tests on these aspects in this area have not been done yet. The qualitative study gives participants the opportunity to express other thoughts they have on the design. During the experiment, participants are asked to interact with the prototype and read the information about it. After the interaction they are asked to fill in a questionnaire which contains the questions for the quantitative analysis of self reflection, user engagement, usability and creepiness of the system. When the participants complete the questionnaire, they are invited to a short semi-structured interview. This interview gives us further insights on the thoughts they have on the design. If the participant gives consent, the interview is recorded on a smartphone. During the interview, three open questions were asked:

1. What did you think of the device?
2. If you had this, what would you use it for?
3. Do you have any further questions or comments?

These questions should be open enough to receive any type of answer and not push the participant into a certain direction. Participants were free to speak on whatever came to mind.

4.4 Participants

Participants were recruited based on convenience sampling. We decided to use convenience sampling as it gives us the opportunity to approach a large number of people during COVID times. Following the outcomes of the literature review, a specific type of person for which ambiguity would be most used was not mentioned, which is why we tested the design on all types of people.

In total, 82 participants ($M = 24.4$ y, $SD = 6.0$ y, male = 54, female = 25, non-binary = 1, prefer not to say = 2). Of all participants 30 graduated high school, 42 had a bachelor degree, 10 had a master degree and 1 had a PhD. Participants were divided equally into four groups that were tested on different conditions (Table 3).

4.5 Potential issues to take into account

Despite the fact that we mention the different design tactics we implemented in our design to create ambiguity, other design choices might also lead to ambiguity. Thus, we admit that our control condition might not be fully unambiguous. However, as we try to create more ambiguity in the other designs using the design tactics, we can still test whether more ambiguous designs deliver certain effects.

Furthermore, our literature review showed that it is difficult to predict how users will interpret the ambiguity (Fdili Alaoui, 2019). By interviewing participants after the interaction with the prototype, we can collect more data on how and why the device was interpreted in a certain way.

Version	n	Age	Gender
A (control)	21	min = 18, max = 63 mean = 25.76, median = 24 SD = 2.94	Male = 15 Female = 4 non-binary = 1 prefer not to say = 1
B (no instruction)	20	min = 17, max = 29 mean = 22.6, median = 22 SD = 2.44	Male = 14 Female = 6 non-binary = 0 prefer not to say = 0
C (no context)	20	min = 17, max = 32 mean = 23.3, median = 23 SD = 3.48	Male = 13 Female = 6 non-binary = 0 prefer not to say = 1
D (non-obvious)	21	min = 18, max = 63 mean = 23.75, median = 23.5 SD = 3.08	Male = 12 Female = 9 non-binary = 0 prefer not to say = 0
Total	82	min = 17, max = 32 mean = 25.62, median = 23 SD = 6.82	Male = 54 Female = 25 non-binary = 1 prefer not to say = 2

Table 3: The descriptive statistics of the participants ($N=82$). The participants were equally divided into four groups, of which each participant was tested on one condition (A, B, C, D).

Additionally, the experiment was conducted at different locations. This could possibly have an influence on the interpretations of participants, as the context is different (Rasmussen et al., 2013). By adding the context in the information, we hope to diminish this influence. Otherwise, this is a point of discussion

5 Results

In this chapter we present the quantitative and qualitative results from our study. In the first section we describe the results from the quantitative data analysis. In the second section we describe the results from the qualitative data analysis.

5.1 Quantitative results

In this section we give answer to our hypotheses. The quantitative data we analyzed are the participants' scores of the used scale: TSRI, SRIS, UES, SUS and PCTS. For each analysis we formulated hypotheses that were analyzed using ART ANOVAS.

DV	df1	df2	F-value	p-value
TSRI	3	78	0.73	0.54
UES	3	78	2.15	0.10
SUS	3	78	2.62	0.056
PCTS	3	78	0.26	0.85

Table 4: Results of the ART ANOVA: df1, df2, F-value and p-value per dependent variable (DV).

5.1.1 Self reflection in technology

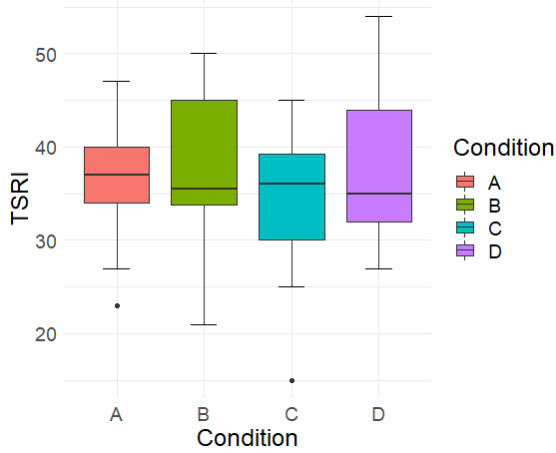
Testing self reflection in technology can depend on the general self reflection of the participant (Bentvelzen et al., 2021). That is why we start by testing the self reflection of participants in general using the SRIS and test whether this correlates with the self reflection in technology via the TSRI. This was tested using Pearson's correlation test. The results were not statistically significant, which means that we did not find a correlation between the TSRI and SRIS. Because of this, we do not use the SRIS scores in the next statistical calculations.

To test whether self reflection in technology is supported by ambiguous design, we formulated the hypothesis presented below. We expect the TSRI score to be higher for the ambiguous conditions, because the systematic review showed that ambiguous design can stimulate reflection in users (e.g., Mols et al., 2016; Lee et al., 2015).

H1: *The use of design tactics for creating ambiguity in systems (condition B,C,D) leads to a higher score of TSRI than systems in which these design tactics are not implemented (condition A).*

H1₀: There is no difference in TSRI score between condition A, B, C and D.

H1_A: There is a difference in TSRI score between condition A, B, C and D.



Condition	N	M TSRI	SD TSRI
A	21	36.8	6.6
B	20	37.9	8.0
C	20	34.3	7.4
D	21	37.9	7.6

Figure 26: A boxplot of the participants' results of the TSRI scores. Each colored box represents one of the conditions. Table 5: Descriptive statistics of the TSRI scores.

ART ANOVA TSRI. To test the hypothesis, an ART ANOVA was performed (Table 4). ART ANOVA revealed that there was not a statistically significant difference in TSRI score between the conditions A, B, C and D ($F(3,78)=0.73$, $p=0.54$). This means that H_{10} can not be rejected and H_{1A} can not be accepted.

Additionally, the standard deviation of the ambiguous conditions B, C and D are higher than condition A (Figure 26, Table 5). Yet, this outcome is not significant and thus we can not draw conclusions on this.

5.1.2 User engagement

To test whether ambiguous design influences the user engagement with the system, the hypothesis presented below was formulated. We test user engagement by using the UES. We expect the UES score to be higher for the ambiguous conditions, because the systematic review showed that ambiguous design can stimulate engagement in users (e.g., Creighton, 2010; Dagan et al., 2019; Seitinger et al., 2010). In the review we describe the user engagement in different forms; curiosity, social engagement, play, and immersion. In this experiment we look at user engagement in general.

H_2 : *The use of design tactics for creating ambiguity in system designs (condition B,C,D) leads to a higher score of UES than systems in which these design tactics are not implemented (condition A).*

H_{20} : There is no difference in UES score between condition A, B, C and D.

H_{2A} : There is a difference in UES score between condition A, B, C and D.

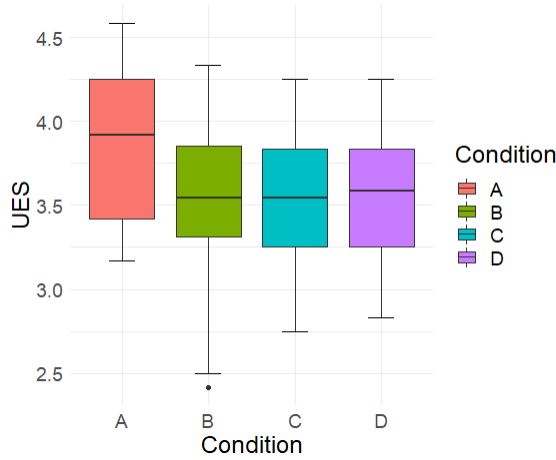


Figure 27: A boxplot of the participants' results of the UES scores. Each colored box represents one of the conditions.

Condition	N	M UES	SD UES
A	21	3.84	0.42
B	20	3.52	0.53
C	20	3.52	0.40
D	21	3.55	0.37

Table 6: Descriptive statistics of the UES scores.

ART ANOVA UES. To test the hypothesis, an ART ANOVA was performed (Table 4). A difference in median is seen between condition A and the other conditions, B, C and D (Figure 27, Table 6). However, the results revealed that there was no statistically significant difference in UES score between the conditions A, B, C and D ($F(3, 78) = 2.15$, $p = 0.10$). This means that H_{20} can not be rejected and H_{2A} can not be accepted.

5.1.3 Usability

To test whether ambiguous design influences the usability of the system, we formulated the hypothesis presented below. Usability is tested using the SUS. According to (Gaver et al., 2003), ambiguous design should not have a low usability. However, in practice, ambiguous design can become more confusing for users in terms of usability (e.g., Cosley et al., 2008). Thus, we expect the usability of the system to differ between the conditions, but we also expect the usability to be sufficient for all conditions.

H3: *The use of design tactics for creating ambiguity in system designs (condition B,C,D) leads to a different SUS score compared to systems in which these design tactics are not implemented (condition A).*

H3₀: There is no difference in SUS score between condition A, B, C and D.

H3_A: There is a difference in SUS score between condition A, B, C and D.

ART ANOVA SUS. The hypothesis is testing by performing an ART ANOVA (Table 4). ART ANOVA revealed that there is weak evidence of a statistically significant difference in SUS score between the condition A, B, C, and D on a significance level of $p < 0.1$ ($F(3,78) = 2.62$, $p = 0.056$). Nonetheless, this means that H_{30} can not be rejected and H_{3A} can not be accepted.

In condition A the SUS score was good ($M > 75$) and in conditions B, C and D the SUS score was above average ($M > 65$) (Figure 28, Table 7). Additionally the standard deviations were higher for the ambiguous conditions B, C and D than for condition A. However, these differences have not proven to be significant.

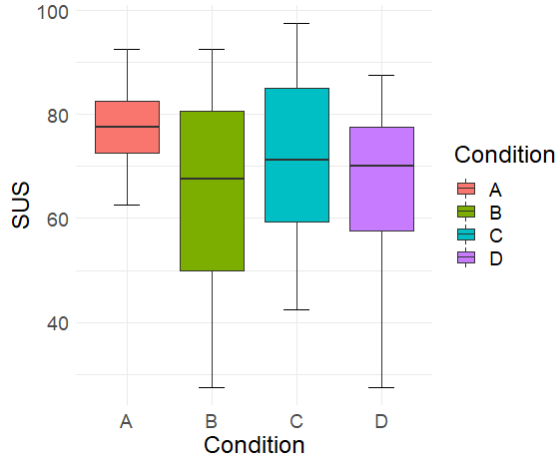


Figure 28: A boxplot of the participants' results of the SUS scores. Each colored box represents one of the conditions.

Condition	N	M SUS	SD SUS
A	21	78.1	7.9
B	20	66.4	17.8
C	20	71.9	15.2
D	21	65.8	16.3

Table 7: Descriptive statistics of the SUS scores.

5.1.4 Perceived creepiness

To test whether ambiguous design is found to be creepy, we formulated the hypothesis presented below. The perceived creepiness is tested using the PCTS. In the systematic review there was one ambiguous design that created an uneasy feeling in users; Breathless, the swing that moves based on the breath of the user (Figure 20 in Section 2.3.6). This design intentionally made their design uncomfortable to evoke unsettling feelings in the users. In our design, we did not intend to implement this design tactics and thus we do not expect to evoke uneasy feelings in the users. Nonetheless, it is still interesting to test whether ambiguity itself has influence on the perceived creepiness in users.

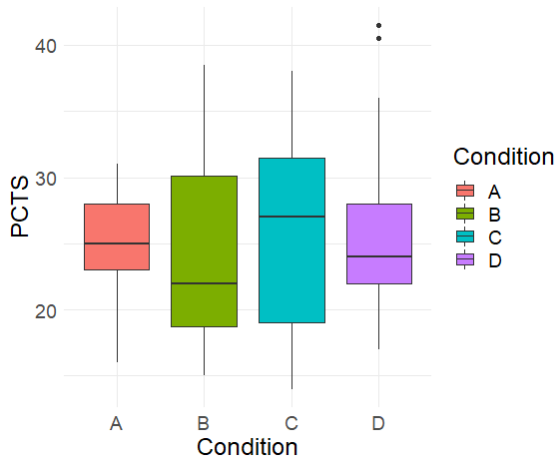
H_4 : *The use of design tactics for creating ambiguity in system designs (condition B,C,D) does not lead to a different score of PCTS compared to systems in which these design tactics are not implemented (condition A).*

H_{4_0} : There is no difference in PCTS score between condition A, B, C and D.

H_{3_A} : There is a difference in PCTS score between condition A, B, C and D.

ART ANOVA PCTS. To test the hypothesis, an ART ANOVA was performed (Table 4). The results revealed that there was no statistically significant difference in PCTS score between the conditions A, B, C and D ($F(3,78) = 0.26$, $p = 0.85$) This means that H_{4_0} can not be rejected and H_{4_A} can not be accepted.

The standard deviations of the PCTS score in the ambiguous conditions B, C and D, are higher than in condition A, but these differences are not significant (Figure 29, Table 8).



Condition	<i>N</i>	<i>M</i> PCTS	<i>SD</i> PCTS
A	21	25.0	4.0
B	20	24.6	7.0
C	20	26.0	7.6
D	21	26.3	7.1

Figure 29: A boxplot of the participants' results of the PCTS scores. Each colored box represents one of the conditions.

Table 8: Descriptive statistics of the PCTS scores.

5.1.5 Summary of quantitative results

In this quantitative study we analyzed whether ambiguous design evokes more self reflection in users, stimulates user engagement, influences the usability of the system and makes users perceive a system as creepy. We did this by implementing different design tactics in the design and having participants fill in a questionnaire consisting of different scales; SRIS, TSRI, UES, SUS, PCTS. The scores were analyzed using ART ANOVAS. For all analysis, no significance was found on a level of $p < 0.05$. Nonetheless, the results of the SUS scores showed a significance of $p < 0.10$, although this is not sufficient support for confirmation of the analysis, it indicates possibilities for further research. Additionally, the designs that make use of the design tactics to create ambiguity, appear to have higher standard deviations for the tested scales TSRI, SUS and PCTS. Again, the results were not significant, but it indicates possibilities for further research.

5.2 Qualitative results

In this section we discuss the qualitative data retrieved from the interviews. In total, we interviewed all (82) participants, and recorded 81 of the interviews with a total duration of 2:42:05 (hr:min:s). The interview of one participant was written down during the interview. The audio recordings were transcribed by hand and analyzed in Atlas.ti. We created a word cloud to get insights of the major themes. Next, we did open coding to review the text in detail. Finally, we cleaned the codes and grouped them based on major themes. These major themes are: participant's overall impression of the device, the feedback they gave us and what they would use the device for.

5.2.1 Overall impression

Besides the quantitative data gathered from the SUS scale, participants expressed their opinions about the usability of the device during the interview. The opinion about usability varied between easy to use, experiencing a learning curve and confusion. Additionally, participants expressed their personal impression of the device, whether they liked it or not. We further discuss this in this section.

Confusion. Participants found the device confusing without being able to figure out how it worked with confidence (Table 11 in Appendix D). They expressed their confusion about different aspects of the device: the time indications, normal visualization, and abstract visualization. We further discuss these in the section about the feedback participants gave (Section 5.2.2). There were no participants of condition A that mentioned confusion about the device. This indicates that ambiguous design can more likely create confusion in users compared to less ambiguous design.

Interestingly, the confusion of participants might not necessarily be a bad thing as participants were still positive about the experience, finding it either funny, nice, interesting or playful. However, there was one participant that thought it was weird (B10) and the other participants did not give their inherent opinion about the device.

Learning curve. Participants experienced a learning curve during the interaction with the device (Table 11 in Appendix D). The learning curve means that they had difficulty understanding the device at first, but accomplished understanding it at the end of the interaction. Aspects of the device participants experienced this most are: time indications, leaf-shaped knob, normal visualization, abstract visualization, and instructions. We further discuss these in the section about the feedback of participants (Section 5.2.2). Of the participants that experienced a learning curve most participants were positive about the interaction, finding it either funny, nice or interesting. Additionally, there were participants who did not express their opinions about the device.

Easy to use. Participants thought the device was clear or easy to use from the start of the interaction (Table 11 in Appendix D). They mentioned specific aspects of the device to be easy to use: normal visualization and abstract visualization. We further discuss these in the section about the feedback participants gave (Section 5.2.2). Participants in conditions A mentioned the device to be easy to use the most, compared to the other conditions. This indicates that ambiguous designs can be less easy to use, compared to less ambiguous designs. Additionally, these participants were overall positive about the device, finding it either funny, nice, or interesting. One participant thought it was weird and other participants did not express their opinion.

Playful interaction. Participants found it an interactive and playful experience. Some of these participants described it as a puzzle. Other participants found the experience interactive because it is a physical device.

5.2.2 Feedback

The previous section presents the opinion of participants about the usability of the device. In this section we further discuss the specific feedback made on the different aspects of the design. Participants suggested different improvements that could be made and some pitched in new ideas. In this section we present the feedback mentioned by the participants per aspect of the device and we discuss the different opinions.

Time indications. The time indications are the blue and orange dots on the device (i.e. Figure 24). These were perceived as confusing by some participants, while other participants had difficulty with figuring out how it worked precisely, but managed to do so in the end. The confusion with the time indications mostly came from participants in condition B, who received less instructions about the device. Participants suggested that adding timestamps to the time indication would make the design more clear.

“I noticed that I hesitated for a second like, ‘ok, but which blue knob is for what?’ However, then I argued that it is like a clock going back in time, so at every indication point that is more distanced from the orange one, that is how you go back in time. But I realize now, that is just an assumption because there are no numbers.”

A01

Leaf-shaped knob and instructions. One participant suggested that a slider from high to low would be clearer to use for displaying the history. Additionally, participants in condition B mentioned that extra explanation about how the device worked would have made them figure out how the design worked more easily. Specifically the purpose of the knob was unclear for some participants. Nonetheless, they thought the current instructions were sufficient enough.

“No, I thought it was actually good to kind of figure it out by yourself, but if there would have been timestamps on the dots, or clearer buttons, it would have been easier.”

B17

Normal visualization. Participants had comments on the moving numbers on the display, as this happens when multiple numbers are displayed. Participants also interpreted the scrolling of numbers and its pace as a way of alarming the user of the air quality. Participants suggested either adding a large display, or adding a display with higher definition. Nonetheless, one participant pointed out that is a fun feature of the design and the number could have been smaller, but they found it more interesting the way it was designed.

Abstract visualization. Participants were confused or experienced a learning curve with the abstract visualization. One participant mentioned to be confused about the background information on air quality that was shown in numbers, while the visualization was presented in a wave. This made them expect a number displayed on the device. Additionally, participants interpreted the waves of the visualization in different ways.

“It was semi confusing, absolutely. It is difficult to connect values to a continuous sinusoid [...] but I understand the essence of faster and slower moving sinusoids. So in that sense it is funny, because you can make the connection: when it moves faster there has to be more of something. Kind of like a geiger counter that goes very fast.”

D20

Size of the device. Participants pointed out that the device is too big to be carried around. They suggested either creating a smaller version or giving it the purpose to stay at one place.

“I think that if you would use this daily, it would be a bit of a thing. It’s like the first phone, it’s a thick unit. But I do think it would be easy to place it in public spaces. I know that they have air pollution sensors in Utrecht, but they generally put those in parks [...] but if you place it at location where there are a lot of people, or at fun location where people pass by, maybe in a forest, then people could be like ‘oh wow, the air quality is way better here than in the city center.’ This way you could maybe influence people’s behavior in a positive way.”

B7

Additionally, participants suggested integrating the data with their smartphone, instead of showing it on a separate device. Participant A13 pitched a more original idea, where the device would be a smaller version that can be worn as a necklace.

“It can probably become a small object that would be a bit more charming. You could even wear it as a necklace. I imagine people will look at me and be like ‘oh the air quality’ and that they will turn the wheel of the device.”

A13

Extra features. Participants pitched ideas for extra features to add to the device. One idea was to add an alarm that would go off when reaching a set limit. Another idea was to add a diary so users can connect the dots with the air quality and the places they have visited. It was also suggested to create a more modular system that makes it possible to add different sensors to the device, and include an SD-card to be able to track a longer time period.

5.2.3 Intention to use

During the interviews we asked the participant what they would use the device for if they had it. This gave us insight into whether they would actually use it, how they would use it and what they would use it for. Intentions of the participants were mixed. Intentions to use it for awareness and behavior change came to light, but also general uses were mentioned. Additionally, participants mentioned they would not use the device in their daily lives.

Awareness and behavior change. Participants mentioned that the device would help with awareness of air quality, of which some would also use it to make others aware of the air quality. Contrary, one participant would not create awareness in others as this is seen as an accusation. Other participants pointed out that the device would not be good for behavior change as it would be too much “*micromanaging*” or confronting.

“.. I don’t want to be confronted with how bad the air quality is all the time. Because, if you would use it too often at different locations, then that would make you dependent on it, or well not dependent, but you will keep questioning yourself ‘is the air clean enough here?’ and it might make you a bit scare of the air isn’t clean.”

A6

Curiosity and comparing. Participants mentioned they would mainly use the device out of curiosity. They wanted to find out all the different things it can measure and the results. Participants would also use it for comparing different things and locations, just to inform themselves about various situations.

Specific uses. There were participants that did not show interest in having the device for personal use. Part of that was the fact that they did not have interest in air quality. Some participants suggested that they would potentially use it if it had other sensors such as virus particles, noise pollution, CO₂, or something that detects how much water you drink. Participants also mentioned attributes and location they would use the device for (Table 12 in Appendix D). Attributes that participants would use the device for were: cigarette smoke, other types of smoke (i.e. incense, smokes from fire), dust in a room, pollution of cars. Participants also mentioned different locations to use the device in: at home, in the kitchen, in the bathroom, in public spaces, at a factory or at work, in public transport and at an airport.

6 Discussion

The goal of this research was to create an understanding of ambiguity as a design resource in HCI. We did this by performing a systematic literature review and conducting a follow-up study testing some of the findings from the review. We pointed out the gap between the abstract design tactics suggested in theory and the implementation of these design tactics in practice. To bridge this gap, we created an overview of concrete design tactics based on ambiguous designs in our corpus, opposed to the suggested design tactics currently found in literature. It is important to note that this list might not be complete, as it is only based on design from our corpus, and in time more design tactics can be added to the list when the creation of ambiguous design grows. To do this, we suggest clear documentation of the creation of ambiguity and also of potential unintended ambiguity that arose during the creation process.

In the follow-up study we created a design with three different ambiguous conditions that made use of different design tactics. Using this design, we tested whether the concrete design tactics for creating ambiguity affected the users' reflection, user engagement, system usability and perceived creepiness of the device. Although we did not find significant results, there are interesting indication which we further discuss our findings in this section.

Concrete design tactics. In the systematic literature review we found that design tactics and strategies have been suggested in different analyses on ambiguous designs (Boehner and Hancock, 2006; Gaver et al., 2003; Sengers and Gaver, 2006). However, in practice these suggested design tactics are often not applied directly. An explanation for this is that these suggested design tactics are presented on a high level of abstraction which makes them either difficult to interpret, or hard to implement directly. By analyzing the corpus, we were able to present concrete design tactics for creating ambiguous designs that are presented on a lower level of abstraction. This way, we aim to connect the abstract theories about ambiguity, with studies that focus more on the application of ambiguity in design. The concrete design tactics we present are not distinctive from the suggested design tactics and strategies and we do not try to diminish this previous work. In a way the concrete design tactics can be connected to the more abstract tactics. However, it is difficult to directly map the concrete tactics to the abstract tactics as these abstract tactics are overlapping and concrete tactics can appear in multiple abstract tactics. This is also the issue we saw in the creation ambiguous design in our corpus, which could be a reason why researchers did not refer or use the abstract design tactics. The concrete design tactics serve as an overview of the tactics that are currently used in practice for creating ambiguous designs.

Self reflection. The results of the systematic literature review show us the possibilities of using ambiguous systems to evoke reflection in users. Ambiguity is also included in the framework of design resources for supporting reflection in systems (Bentvelzen et al., 2021). Since these studies are based on qualitative data, we tested the influence of ambiguity on self reflecting quantitatively, using the TSRI scale. Although we did not find significant results, there was an indication of difference in standard deviations between the less ambiguous and ambiguous conditions. We can not draw conclusions on this, but it is something that can be further studied, which we discuss in our suggestions for future research (Section 6.3).

User engagement. In our systematic review we found that engagement in users was evoked by curiosity or in the form of social engagement, play and immersion. In our qualitative analysis participants from all conditions, so regardless of the ambiguity, mentioned to have an interactive and playful experience. One explanation for this was the fact that it is a physical device, which made it more interactive for the participants. We also tested the user engagement quantitatively, by

using the User Engagement Scale. We did not find significant results, but there was an indication of difference in means between the less ambiguous and ambiguous conditions. Interestingly, this was an unexpected difference, as the average user engagement for the less ambiguous condition seemed higher, than for the ambiguous conditions. This difference can be explained by a number of reasons. The first reason is that a number of factors can influence the user engagement and we did not further study these factors and thus were not able to take these into account. Second, it can indicate that ambiguous design might not be engaging, although this seems unlikely to us, as other researchers observed this in their studies (e.g., [Jorge et al., 2013](#)). Third, this can indicate a flaw in our design. For instance, it can be that condition A was already perceived as ambiguous, the design tactics made the design even more ambiguous, but maybe too ambiguous. We discuss this idea further in limitations (Section 6.2).

Balance and confusion. Finding the right balance of ambiguity is an important task in creating ambiguous design ([Sanchez et al., 2019](#)). If a device is made too ambiguous, it can create confusion and frustration in users, while a design that is made too clear can prevent users from reaching the desired outcomes, i.e. engagement. The qualitative findings of our study showed that the ambiguous conditions evoked confusion in some participants, while the unambiguous condition did not. This can indicate that our ambiguous conditions might have been made too ambiguous, which we further discuss in limitations (Section 6.2).

Furthermore, opinions about the usability of the design appear to be more diverse in the ambiguous conditions. Although this was not significantly found in the quantitative results, it indicates a difference to be further studied. Nonetheless, the qualitative results also showed an indication of more varied opinions on the usability of the design in ambiguous conditions. In the ambiguous conditions, participants found the device either confusing, easy to use or experienced a learning curve. While, in the unambiguous condition, no participants found the device confusing and thought it was easy to use, or experienced a learning curve. This result was also found in our corpus, where studies confirmed that ambiguous design can be confusing for one, but be eye-opening for another (e.g., [Cosley et al., 2008](#)). Additionally, we found that confusion in participants was not necessarily a negative thing, as a part of the participants still found the experience funny, nice, interesting or playful. There were also participants that found the device confusing, but did not give their opinion, which can also mean that they kept their negative opinions for themselves, we discuss these further in the limitations section (Section 6.2).

6.1 Design adaptations

The ambiguous design we created was not perfect. As we discussed above, it created confusion in participants, and although this might not necessarily be bad, we did not intend to confuse participants. In this section we discuss how we created other ambiguities in the design that we did not intend to create and how this implementation could be changed. Additionally, we discuss the different versions of the device (no instructions, no original context, and the combination of non-obvious connection with abstract visualization) which we implemented, and how this implementation could be different. Important to note is that there is no standardized way of measuring ambiguity and how it is balanced, which makes reviewing the ambiguity of a design a difficult task. Thus we merely give suggestions to how the design can be changed and the different design tactics can be tested.

Unintended ambiguities. Creating ambiguity in design tactics is a challenging task, because it is difficult to know in advance how users will perceive the ambiguity ([Fdili Alaoui, 2019](#)). We also found this in the interviews, where participants interpreted the normal visualization, shown in conditions A, B and C, in different ways. This was an ambiguity we did not mean to create, but

was perceived by participants. We could have anticipated this better by performing a pilot study. Although we reviewed the device, we did not perform a proper pilot study with a lot of participants. It is also the question how many participants we would have needed in order to find this ambiguity, as a lot of participants from our study did not see the scrolling numbers as an issue.

Furthermore, as previously mentioned, participants were engaged because of the physical appearance of the device. The majority of ambiguous designs in our corpus were also physical artifacts that people could interact with. In the systematic review, we recognize data physicalization as a concrete design tactic. Although we did not physicalize the data shown to users, i.e. air quality, the device itself was physical. This can indicate that creating a physical design opposed to digital design, ambiguity can be stimulated. This can be further tested by creating a physical design and a digital design with the same features, to see whether physicalization of design adds to the user engagement.

Absence of instructions. We gave participants less instructions about the device compared to the other conditions. Nonetheless, they still received instructions, which we did to make sure the device did not become too ambiguous. However, in the end this still resulted in confusion with participants, specifically regarding the use of the time indications, and which way to turn to go back in history. In this case, the ambiguity was found in the usability of the device. In future work we could test different types of absence of instructions. This way we can find out when the absence of instructions work and what not. For instance, we could leave out the background information about air quality instead of instruction about the device. A possible side effect that would happen with this, is that it also creates a non-obvious connection, as it might not be clear for participants what data is presented.

Absence of original context. The device had no original context, meaning it could be placed in any setting. We tried testing the absence of the original context by adding extra context to the designs in the other conditions. This was mainly done out of convenience, as it was difficult for us to be bound to one location in view of the pandemic. We also tested the device at different locations, but did not see differences between these settings. In another study it would be interesting to test how the interpretation of the design is influenced in different settings as was done in the study by [Rasmussen et al. \(2013\)](#). In the qualitative results, one participant also pointed out that it can be put in different public spaces which makes people compare the different locations. Additionally, it would be interesting whether different locations influence the ambiguity of the device.

Non-obvious connection and abstract visualization. The non-obvious connection and abstract visualization was created by visualizing a wave that represented the air quality. The implementation of these design tactics worked out well. One aspect that can be changed in a follow-up study, is the background information of air quality. The list that indicates the goodness of air quality in numbers, might not have been useful in condition D. We decided to keep this for this research, to make sure that we only changed the visualization in the design. However, as a participant pointed out in the qualitative results, this design choice led them to confusion. Another aspect to study in the future is the relationship between non-obvious connection and abstract visualization. We chose to combine these in this study, as these seem to be closely related. It would be interesting to see how the ambiguity is influenced when these are tested separately.

6.2 Limitations

There are several limitations to this research that have possibly impacted the results of this study. In this section we discuss these limitations and give advice on how this can be improved in a

follow-up study.

Self reflection. In the qualitative reviews we found that some participants did not feel connected with air quality and would not care to use the device in their daily lives. This is an issue that can influence the reflection of the participants. We presented the air quality as general data on which every participant can reflect. However, if participants do not care for this data, reflecting on it might be a difficult task for them. Thus, we suggest that in a follow-up study, participants could be selected based on their interest in the data that is to be reflected on.

Sample size. The sample size used in this experiment is limited, which makes generalization for the population not possible. There are different reasons for the limited sample size. First, this experiment was conducted during the pandemic, which made recruiting difficult, especially since the experiment had to be done physically. Second, we chose an between-subjects design with four conditions, which made that one participant could only test one condition. As previously mentioned, we chose between-subjects design because the literature review showed that it is difficult to create new interpretations of design once these are made (Sanches et al., 2019). We could have chosen to test one condition less, however, because we suggested new concrete design tactics, it seemed appropriate to test more than one design tactics. Nonetheless, in terms of our sample size, this was not in our advantage. In future work we suggest to create a study with less conditions, so it is easier to have a larger sample size per condition.

Culture. Participants were recruited via convenience sampling, which made recruitment easier for us, especially during the pandemic. Because of this, a large part of the participants were students from information- or computer sciences studies, and friends and family members of the researcher. Here we can clearly see different cultures of participants that are familiar with technologies and study different aspects of it, and participants that might be familiar with technologies, but do not study this. Additionally, the majority of participants were young adults (ages 17-32), while two participants were 63. This difference in age can also have an influence on the way technology is seen, and thus how the interaction with our ambiguous design is perceived. Although we tried to distribute the participants evenly in different groups, these differences in culture are still of influence. We suggest in future research to keep this in mind and try recruiting participants on a more specified means of sampling.

Social desirability bias. In our study we also need to take into account the possibility of social desirability bias, especially for the qualitative part. The participants that participated in the study were all acquainted with the researcher that conducted the experiment and knew that the device was created by her, although this was not explicitly stated in the experiment. Because of this, participants might have kept their negative opinions to themselves. In a new experiment this bias can be reduced by recruiting unknown participants and stating that the device was made by someone else.

6.3 Future Research

We suggest several future research directions based on our work. First, even with the use of our concrete design tactics, creating ambiguous designs is a complex task. A standardized way of testing ambiguity can help solve this issue. Currently, this does not exist and it is the question whether this is possible, as ambiguous design can evoke different things in users in different ways. Based on this, a format for reviewing ambiguity in a quantitative way does not seem to be a good choice and a more abstract approach seems more suitable for this job. When a standardized way of testing is achieved, it will be easier for follow-up studies to find the right balance in ambiguous design.

Second, our research indicates that ambiguity can possibly influence the user reflection, user engagement, perceived usability and perceived creepiness in a way that the opinions of users are more diverse. Moreover, the interaction between perceived confusion in the device and the overall impression of users can be further studied. It would be interesting to see whether this interaction exists, and to which fields it applies. Preferably, these aspects are again tested in a quantitative way, as quantitative studies in research on ambiguity are scarce. A challenge will be to take into account all the aspects that can influence the more personal variables to be tested, i.e. user reflection.

7 Conclusion

The aim of our research was to create an understanding of ambiguity as a design resource in HCI. We did this by performing a systematic literature review on a corpus of fifty five papers. In this review, we created an overview of the current understanding of ambiguity, the effects of ambiguous designs and the suggested design tactics. The suggested design tactics showed a high level of abstraction, which could make it difficult for other researchers to directly implement them in their design. To undertake this, we presented ten concrete design tactics that were based on ambiguous designs from our corpus: absence of purpose, absence of obvious connection, absence of information, absence of object, connect external situation, abstract representation, data physicalization, ambiguous data, and uncomfortable design.

Based on the concrete design tactics, we created an ambiguous prototype ourselves. We tested the different effects of our ambiguous design on participants. The effects tested were: users' reflection, user engagement, system usability and perceived creepiness. While the quantitative results were not significant, the qualitative results indicate that ambiguous design evokes a more varied opinion. Participants found the device either confusing, experienced a learning curve or thought it was easy to use. Interestingly, the majority of participants were positive about the device, even when they perceived it as confusing. This is something that should be further studied. Furthermore, in this research we discussed the importance of balance of ambiguity, but did not study this further. We suggest studying the correct balance of ambiguity in future research, and developing a standardized way to analyze ambiguity of a design.

References

- Back, J., Segura, E. M., and Waern, A. (2017). Designing for transformative play. ACM Transactions on Computer-Human Interaction (TOCHI), 24(3):1–28.
- Bae, S. S., Kwon, O.-H., Chandrasegaran, S., and Ma, K.-L. (2020). Spinneret: Aiding creative ideation through non-obvious concept associations. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, pages 1–13.
- Baumer, E. P., Khovanskaya, V., Matthews, M., Reynolds, L., Schwanda Sosik, V., and Gay, G. (2014). Reviewing reflection: on the use of reflection in interactive system design. In Proceedings of the 2014 conference on Designing interactive systems, pages 93–102.
- Benford, S., Greenhalgh, C., Giannachi, G., Walker, B., Marshall, J., and Rodden, T. (2012). Uncomfortable interactions. In Proceedings of the sigchi conference on human factors in computing systems, pages 2005–2014.
- Bentvelzen, M., Niess, J., Woźniak, M. P., and Woźniak, P. W. (2021). The development and validation of the technology-supported reflection inventory. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, pages 1–8.
- Boehner, K. and Hancock, J. T. (2006). Advancing ambiguity. In Proceedings of the SIGCHI conference on Human Factors in computing systems, pages 103–106.
- Briggs, P., Blythe, M., Vines, J., Lindsay, S., Dunphy, P., Nicholson, J., Green, D., Kitson, J., Monk, A., and Olivier, P. (2012). Invisible design: exploring insights and ideas through ambiguous film scenarios. In Proceedings of the Designing Interactive Systems Conference, pages 534–543.
- Brooke, J. (1986). System usability scale (sus): a quick-and-dirty method of system evaluation user information. Reading, UK: Digital Equipment Co Ltd, 43:1–7.
- Brynjarsdottir, H., Håkansson, M., Pierce, J., Baumer, E., DiSalvo, C., and Sengers, P. (2012). Sustainably unpersuaded: how persuasion narrows our vision of sustainability. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pages 947–956.
- Cosley, D., Lewenstein, J., Herman, A., Holloway, J., Baxter, J., Nomura, S., Boehner, K., and Gay, G. (2008). Artlinks: fostering social awareness and reflection in museums. In Proceedings of the SIGCHI conference on human factors in computing systems, pages 403–412.
- Creighton, E. (2010). Jogo: an explorative design for free play. In Proceedings of the 9th International Conference on Interaction Design and Children, pages 178–181.
- Dagan, E., Márquez Segura, E., Altarriba Bertran, F., Flores, M., and Isbister, K. (2019). Designing 'true colors' a social wearable that affords vulnerability. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, pages 1–14.
- Desjardins, A., Biggs, H. R., Key, C., and Viny, J. E. (2020). Iot data in the home: Observing entanglements and drawing new encounters. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, pages 1–13.
- Devendorf, L., Lo, J., Howell, N., Lee, J. L., Gong, N.-W., Karagozler, M. E., Fukuhara, S., Poupyrev, I., Paulos, E., and Ryokai, K. (2016). "i don't want to wear a screen" probing perceptions of and possibilities for dynamic displays on clothing. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, pages 6028–6039.

- Doherty, K. and Doherty, G. (2019). Engagement in HCI: Conception, theory and measurement. ACM Computing Surveys, 51(5).
- Fajardo, N. and Moere, A. V. (2008). Externaleyes: evaluating the visual abstraction of human emotion on a public wearable display device. In Proceedings of the 20th Australasian Conference on Computer-Human Interaction: Designing for Habitus and Habitat, pages 247–250.
- Fdili Alaoui, S. (2019). Making an interactive dance piece: Tensions in integrating technology in art. In Proceedings of the 2019 on Designing Interactive Systems Conference, pages 1195–1208.
- Gaver, B., Dunne, T., and Pacenti, E. (1999). Design: cultural probes. interactions, 6(1):21–29.
- Gaver, W., Bowers, J., Boucher, A., Law, A., Pennington, S., and Villar, N. (2006). The history tablecloth: illuminating domestic activity. In Proceedings of the 6th conference on Designing Interactive systems, pages 199–208.
- Gaver, W., Bowers, J., Kerridge, T., Boucher, A., and Jarvis, N. (2009). Anatomy of a failure: how we knew when our design went wrong, and what we learned from it. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pages 2213–2222.
- Gaver, W. W., Beaver, J., and Benford, S. (2003). Ambiguity as a resource for design. In Proceedings of the SIGCHI conference on Human factors in computing systems, pages 233–240.
- Grant, A. M., Franklin, J., and Langford, P. (2002). The self-reflection and insight scale: A new measure of private self-consciousness. Social Behavior and Personality: an international journal, 30(8):821–835.
- Håkansson, M. and Gaye, L. (2008). Bringing context to the foreground: designing for creative engagement in a novel still camera application. In Proceedings of the 7th ACM conference on Designing interactive systems, pages 164–173.
- Hassenzahl, M., Heidecker, S., Eckoldt, K., Diefenbach, S., and Hillmann, U. (2012). All you need is love: Current strategies of mediating intimate relationships through technology. ACM Transactions on Computer-Human Interaction (TOCHI), 19(4):1–19.
- Höök, K. (2006). Designing familiar open surfaces. In Proceedings of the 4th Nordic conference on Human-computer interaction: changing roles, pages 242–251.
- Jeng, T., Ma, Y.-P., and Shen, Y.-T. (2011). iawn: designing smart artifacts for sustainable awareness. In International Conference on Universal Access in Human-Computer Interaction, pages 193–202. Springer.
- Jorge, C., Hanna, J., Nisi, V., Nunes, N., Caldeira, M., and Innella, G. (2013). Fostering ambiguity: decontextualizing and repurposing a familiar public display. In Proceedings of the Biannual Conference of the Italian Chapter of SIGCHI, pages 1–10.
- Karyda, M., Ryöppy, M., Buur, J., and Lucero, A. (2020). Imagining data-objects for reflective self-tracking. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, pages 1–12.
- Laaksolahti, J., Tholander, J., Lundén, M., Solsona Belenguer, J., Karlsson, A., and Jaensson, T. (2010). The lega: a device for leaving and finding tactile traces. In Proceedings of the fifth international conference on Tangible, embedded, and embodied interaction, pages 193–196.

- Leahu, L., Schwenk, S., and Sengers, P. (2008a). Subjective objectivity: negotiating emotional meaning. In Proceedings of the 7th ACM conference on Designing interactive systems, pages 425–434.
- Leahu, L., Thom-Santelli, J., Pederson, C., and Sengers, P. (2008b). Taming the situationist beast. In Proceedings of the 7th ACM Conference on Designing interactive Systems, pages 203–211.
- Lee, M.-H., Cha, S., and Nam, T.-J. (2015). Patina engraver: Visualizing activity logs as patina in fashionable trackers. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, pages 1173–1182.
- Lewis, J. R. and Sauro, J. (2009). The factor structure of the system usability scale. In International conference on human centered design, pages 94–103. Springer.
- Li, Z., Patibanda, R., Brandmueller, F., Wang, W., Berean, K., Greuter, S., and Mueller, F. (2018). The guts game: towards designing ingestible games. In Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play, pages 271–283.
- Li, Z., Wang, Y., Sheahan, J., Jiang, B., Greuter, S., and Mueller, F. F. (2020). Insideout: Towards an understanding of designing playful experiences with imaging capsules. In Proceedings of the 2020 ACM Designing Interactive Systems Conference, pages 601–613.
- Lindley, S. E., Banks, R., Harper, R., Jain, A., Regan, T., Sellen, A., and Taylor, A. S. (2009). Resilience in the face of innovation: Household trials with bubbleboard. International Journal of Human-Computer Studies, 67(2):154–164.
- Lockton, D., Zea-Wolfson, T., Chou, J., Song, Y., Ryan, E., and Walsh, C. (2020). Sleep ecologies: Tools for snoozy autoethnography. In Proceedings of the 2020 ACM Designing Interactive Systems Conference, pages 1579–1591.
- Mendes, M., Ângelo, P., Nisi, V., and Correia, N. (2012). Digital art, hci and environmental awareness evaluating play with fire. In Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design, pages 408–417.
- Merrill, N. and Cheshire, C. (2017). Trust your heart: Assessing cooperation and trust with biosignals in computer-mediated interactions. In Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, pages 2–12.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., Group, P., et al. (2009). Preferred reporting items for systematic reviews and meta-analyses: the prisma statement. PLoS medicine, 6(7):e1000097.
- Mols, I., Van Den Hoven, E., and Eggen, B. (2016). Technologies for everyday life reflection: Illustrating a design space. In Proceedings of the TEI'16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction, pages 53–61.
- Mueller, F., Vetere, F., Gibbs, M., Edge, D., Agamanolis, S., Sheridan, J., and Heer, J. (2012). Balancing exertion experiences. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pages 1853–1862.
- Nagargoje, A., Maybach, K., and Sokoler, T. (2012). Social yoga mats: designing for exercising/socializing synergy. In Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction, pages 87–90.
- Nold, C. (2004). Bio mapping / emotion mapping. <http://biomapping.net/>. Retrieved February 22, 2022.

- Noortman, R., Schulte, B. F., Marshall, P., Bakker, S., and Cox, A. L. (2019). Hawkeye-deploying a design fiction probe. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, pages 1–14.
- Normark, M. and Tholander, J. (2014). Performativity in sustainable interaction: the case of seasonal grocery shopping in ecofriends. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pages 271–280.
- Núñez-Pacheco, C. and Loke, L. (2014). Aesthetic resources for technology-mediated bodily self-reflection: the case of eloquent robes. In Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: The Future of Design, pages 1–10.
- O'Brien, H. L., Cairns, P., and Hall, M. (2018). A practical approach to measuring user engagement with the refined user engagement scale (ues) and new ues short form. International Journal of Human-Computer Studies, 112:28–39.
- Papworth, N. (2010). ispoons: an audio focused game design. In Proceedings of the 5th Audio Mostly Conference: A Conference on Interaction with Sound, pages 1–8.
- Patel, R., Schnädelbach, H., and Koleva, B. (2018). Come and play: interactive theatre for early years. In Proceedings of the Twelfth International Conference on Tangible, Embedded, and Embodied Interaction, pages 376–385.
- Pichlmair, M., Mech, L., and Sicart, M. (2017). Designing for immediate play. In Proceedings of the 12th International Conference on the Foundations of Digital Games, pages 1–8.
- Raijmakers, B., Gaver, W. W., and Bishay, J. (2006). Design documentaries: inspiring design research through documentary film. In Proceedings of the 6th conference on Designing Interactive systems, pages 229–238.
- Rasmussen, M. K., Grönvall, E., Kinch, S., and Petersen, M. G. (2013). "it's alive, it's magic, it's in love with you" opportunities, challenges and open questions for actuated interfaces. In Proceedings of the 25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration, pages 63–72.
- Rogers, Y., Hazlewood, W. R., Marshall, P., Dalton, N., and Hertrich, S. (2010). Ambient influence: Can twinkly lights lure and abstract representations trigger behavioral change? In Proceedings of the 12th ACM international conference on Ubiquitous computing, pages 261–270.
- Rogers, Y., Sharp, H., and Preece, J. (2011). Interaction design: beyond human-computer interaction. John Wiley & Sons.
- Ryding, K. (2020). The silent conversation: Designing for introspection and social play in art museums. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, pages 1–10.
- Sanches, P., Höök, K., Sas, C., and Ståhl, A. (2019). Ambiguity as a resource to inform protopractices: The case of skin conductance. ACM Transactions on Computer-Human Interaction (TOCHI), 26(4):1–32.
- Seevinck, J. and Edmonds, E. (2009). Open in art, nature and emergence. In Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group: Design: Open 24/7, pages 365–368.

- Seitinger, S., Taub, D. M., and Taylor, A. S. (2010). Light bodies: exploring interactions with responsive lights. In Proceedings of the fourth international conference on Tangible, embedded, and embodied interaction, pages 113–120.
- Sengers, P., Boehner, K., Mateas, M., and Gay, G. (2008). The disenchantment of affect. Personal and Ubiquitous Computing, 12(5):347–358.
- Sengers, P. and Gaver, B. (2006). Staying open to interpretation: engaging multiple meanings in design and evaluation. In Proceedings of the 6th conference on Designing Interactive systems, pages 99–108.
- Seok, J.-m., Woo, J.-b., and Lim, Y.-k. (2014). Non-finito products: a new design space of user creativity for personal user experience. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pages 693–702.
- Sokoler, T. and Svensson, M. S. (2007). Embracing ambiguity in the design of non-stigmatizing digital technology for social interaction among senior citizens. Behaviour & Information Technology, 26(4):297–307.
- Ståhl, A. and Höök, K. (2008). Reflecting on the design process of the affective diary. In Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges, pages 559–564.
- Terzimehić, N., Häuslschmid, R., Hussmann, H., and Schraefel, M. (2019). A review & analysis of mindfulness research in hci. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems-CHI, volume 19.
- van Huut, T. (2018). Zo woekert de natuur op utrecht centraal via schermen naar binnen. <https://www.brainwash.nl/bijdrage/zo-woekert-de-natuur-op-utrecht-centraal-via-schermen-naar-binnen>. Retrieved February 11, 2022.
- Vidyarthi, J., Riecke, B. E., and Gromala, D. (2012). Sonic cradle: designing for an immersive experience of meditation by connecting respiration to music. In Proceedings of the designing interactive systems conference, pages 408–417.
- Wensveen, S., Overbeeke, K., Djajadiningrat, T., and Kyffin, S. (2004). Freedom of fun, freedom of interaction. Interactions, 11(5):59–61.
- Woźniak, P. W., Karolus, J., Lang, F., Eckerth, C., Schöning, J., Rogers, Y., and Niess, J. (2021). Creepy technology: What is it and how do you measure it? In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, pages 1–13.
- Wright, P. and McCarthy, J. (2008). Empathy and experience in hci. In Proceedings of the 2008 SIGCHI conference on human factors in computing systems, pages 637–646.
- Wright, P., Wallace, J., and McCarthy, J. (2008). Aesthetics and experience-centered design. ACM Transactions on Computer-Human Interaction (TOCHI), 15(4):1–21.

A Corpus overview

In this section we present two tables that give an overview of the papers in our corpus. Table 9 presents an overview of the papers in which ambiguity was discussed, but not created. Table 10 presents an overview of the papers in which ambiguous design was created.

Papers	Presentation of ambiguity	Application field
Back et al. (2017)	Discussion of ambiguity in a discussion on something else	Play
Bae et al. (2020)	Ambiguous design was created	Education / work
Benford et al. (2012)	Ambiguous design was created	Art / Culture
Boehner and Hancock (2006)	Discussion on ambiguity	Communication
Briggs et al. (2012)	Ambiguous design was created	Design Research
Cosley et al. (2008)	Ambiguous design was created	Art / Culture
Creighton (2010)	Ambiguous design was created	Play
Dagan et al. (2019)	Ambiguous design was created	Play
Desjardins et al. (2020)	Ambiguous design was created	Home
Devendorf et al. (2016)	Ambiguous design was created	Daily life
Fajardo and Moere (2008)	Ambiguous design was created	Health / Well-being
Fdili Alaoui (2019)	Ambiguous design was created	Art / Culture
Gaver et al. (2003)	Discussion on ambiguity	Design Research
Gaver et al. (2006)	Ambiguous design was created	Home
Gaver et al. (2009)	Ambiguous design was created	Home
Håkansson and Gaye (2008)	Ambiguous design was created	Daily life
Hassenzahl et al. (2012)	Discussion of ambiguity in a discussion on something else	Communication
Höök (2006)	Ambiguous design was created	Daily life
Jeng et al. (2011)	Ambiguous design was created	Home
Jorge et al. (2013)	Ambiguous design was created	Art / Culture
Karyda et al. (2020)	Ambiguous design was created	Home
Laakolahti et al. (2010)	Ambiguous design was created	Art / Culture
Leahu et al. (2008a)	Ambiguous design was created	Daily life
Leahu et al. (2008b)	Ambiguous design was created	Daily life
Lee et al. (2015)	Ambiguous design was created	Health / Well-being
Li et al. (2018)	Ambiguous design was created	Play
Li et al. (2020)	Ambiguous design was created	Play
Lindley et al. (2009)	Ambiguous design was created	Home
Lockton et al. (2020)	Ambiguous design was created	Daily life
Mendes et al. (2012)	Ambiguous design was created	Art / Culture
Merrill and Cheshire (2017)	ambiguous design was created	Communication
Mols et al. (2016)	Discussion of ambiguity in a discussion on something else	Design Research
Mueller et al. (2012)	Ambiguous design was created	Health / Well-being
Nagargoje et al. (2012)	Ambiguous design was created	Health / Well-being
Noortman et al. (2019)	Ambiguous design was created	Health / Well-being
Normark and Tholander (2014)	Ambiguous design was created	Daily life
Núñez-Pacheco and Loke (2014)	Ambiguous design was created	Health / Well-being
Papworth (2010)	Ambiguous design was created	Play
Patel et al. (2018)	Ambiguous design was created	Play
Pichlmair et al. (2017)	Discussion of ambiguity in a discussion on something else	Play
Raijmakers et al. (2006)	Ambiguous design was created	Design Research
Rasmussen et al. (2013)	Ambiguous design was created	Daily life
Rogers et al. (2010)	Ambiguous design was created	Daily life
Ryding (2020)	Ambiguous design was created	Art / Culture
Sanchez et al. (2019)	Ambiguous design was created	Health / Well-being
Seevinck and Edmonds (2009)	Ambiguous design was created	Art / Culture
Seitinger et al. (2010)	Ambiguous design was created	Art / Culture
Sengers and Gaver (2006)	Discussion on ambiguity	Design Research
Sengers et al. (2008)	Ambiguous design was created	Daily life
Seok et al. (2014)	Discussion of ambiguity in a discussion on something else	Design Research
Sokoler and Svensson (2007)	Ambiguous design was created	Home
Vidarthi et al. (2012)	Ambiguous design was created	Art / Culture
Wensveen et al. (2004)	Ambiguous design was created	Home
Wright et al. (2008)	Ambiguous design was created	Daily life
Wright and McCarthy (2008)	Discussion of ambiguity in a discussion on something else	Design Research

Table 9: All papers in the corpus, how they discuss ambiguity and in which it is applied.

Papers	Design	Effects	Design tactics
Bae et al. (2020)	Spinneret	reflection, creativity	absence of obvious connection
Benford et al. (2012)	Breathless	curiosity, immersion	uncomfortable interactions
Briggs et al. (2012)	Invisible Design (method)	creativity	absence of object
Cosley et al. (2008)	ArtLinks	social engagement, play, reflection	absence of information
Creighton (2010)	Jogo	play	absence of information
Dagan et al. (2019)	True Colors	social engagement, immersion	abstract representation
Desjardins et al. (2020)	Data Dunes	play, understanding	abstract representation, data physicalization
Devendorf et al. (2016)	Ebb	social engagement, understanding	abstract representation, data physicalization
Fajardo and Moore (2008)	External Eyes	reflection	abstract representation
Fdili Alaoui (2019)	SKIN	immersion, awareness	absence of obvious connection
Gaver et al. (2006)	History Tablecloth	play, understanding	connect external situations
Gaver et al. (2009)	Home Health Monitor	understanding	absence of information
Håkansson and Gaye (2008)	Context Camera	play, understanding, creativity	connect external situations
Höök (2006)	Affective Diary	understanding	abstract representation
Jeng et al. (2011)	iAwn	awareness, creativity	absence of purpose
Jorge et al. (2013)	MStoryG	curiosity	absence of original context
Karyda et al. (2020)	Data-objects (method)	reflection	data physicalization
Laaksoaho et al. (2010)	Lega	understanding	absence of purpose
Leahu et al. (2008a)	Arousal Maps	creativity	absence of purpose
Leahu et al. (2008b)	Situationist Agent	reflection	absence of original context
Lee et al. (2015)	Patina Engraver	reflection	abstract representation
Li et al. (2018)	Guts Game	understanding	abstract representation
Li et al. (2020)	InsideOut	social engagement, play, understanding, reflection	connect external situations, abstract representation
Lindley et al. (2009)	The bubbleBoard	play	n/a
Lockton et al. (2020)	Making Time	creativity	absence of purpose
Mendes et al. (2012)	Play With Fire	play, awareness	n/a
Merrill and Cheshire (2017)	Heartrate monitor	reflection	abstract representation, ambiguous data
Mueller et al. (2012)	Jogging Over a Distance	curiosity	connect external situations
Nargargoje et al. (2012)	Social Yogamat	social engagement	absence of purpose
Noortman et al. (2019)	Hawkeye	social engagement, reflection	absence of information
Normark and Tholander (2014)	EcoFriends	reflection	absence of information
Núñez-Pacheco and Loke (2014)	Eloquent Robes	reflection	abstract representation, data physicalization, ambiguous data
Papworth (2010)	iSpooks	immersion	absence of information
Patel et al. (2018)	Suspended Discs	play, creativity	absence of purpose
Raijmakers et al. (2006)	Design documentary	creativity	n/a
Rasmussen et al. (2013)	Actuated Interfaces	understanding	absence of original context
Rogers et al. (2010)	Clouds	curiosity, understanding, reflection	abstract representation
Ryding (2020)	Never Let Me Go	curiosity	absence of purpose
Sanches et al. (2019)	Affective Health	curiosity, understanding	absence of purpose, abstract representation
Seevinck and Edmonds (2009)	+now	immersion, understanding, creativity	data physicalization
Seitinger et al. (2010)	Light Bodies	curiosity	absence of purpose
Sengers et al. (2008)	Affector	play, awareness	abstract representation
Sokoler and Svensson (2007)	The Presence Remote	social engagement	absence of purpose
Vidyarthi et al. (2012)	Sonic Cradle	immersion	absence of obvious connection
Wensveen et al. (2004)	Alarm Slider	curiosity, play	absence of obvious connection
Wright et al. (2008)	Blossom	play, reflection	connect external situations

Table 10: Papers in which ambiguous design was created, with the name of the design, its effects and design tactics.

B Additional resources

A link and QR code are made that refer the reader to Google Drive map with additional resources (Figure 30). The additional resources added in this map are: 1) 3D models of the prototype, and 2) example videos of the prototype. All credits for creation of the 3D models go to Jelle Algra.



Figure 30: QR code and link to the additional resources: https://drive.google.com/drive/folders/1I-mGmESVQ8mOZoC1GZppCTDOR_1Aeabk?usp=sharing

C Questionnaire

This section includes the questionnaire that was presented to the participants. Because different conditions received different instructions, we present condition A (the control condition below) and indicate the changes made per condition.

C.1 Informed consent

Thank you for participating in this study for my master's thesis. My name is Pia Herbes and I am studying Human Computer Interaction at Utrecht University. If you have any questions or concerns about the study, you can address them at any time and contact me at: s.f.herbes@students.uu.nl.

During this session you will be asked to interact with a device, perform two small tasks during the interaction and fill in a survey after the interaction with the device. At the end of the session you will be interviewed shortly. The session will take around 10 minutes of your time.

The data collected from the survey is and will remain anonymous. During the short interview you will be asked to be recorded, you have the rights to decline the recording or stop the recording at any time. All data will be kept safe in the Qualtrics environment and Google Drive. The data will be used for the purpose of this study only.

As a participant to this study you have the right to stop participating at any time. Once you stop participating in this session, it is not possible to finish it another time. If you feel uncomfortable during the session, you can address this at any time.

[] I have read the consent form, I recognize my rights within this study and I agree to participate under these terms.

C.2 Participant number

Participant number that is given to you: ____

C.3 Given instructions and tasks

C.3.1 Introduction (all conditions)

You have been given a device called airQualityTime. AirQualityTime shows the PM2.5 value that is generally used for measuring air pollution. PM2.5 are very small particles usually found in smoke and long exposure to high PM2.5 values can be damaging. The table below shows the categorization of the air quality levels per PM2.5 values measured over an hour.

Air quality category	PM2.5 averaged over 1 hr.
Good	< 25
Fair	25-50
Poor	50-100
Very poor	100-300
Extremely poor	> 300

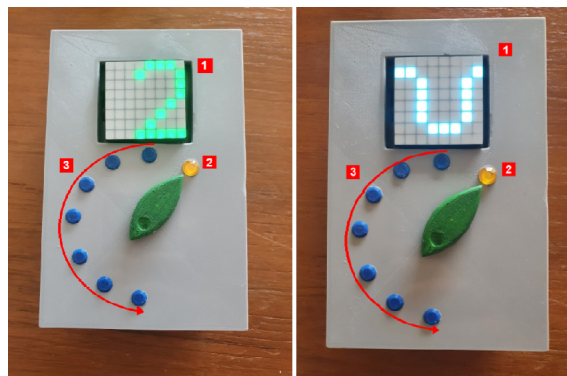
C.3.2 Instructions (all conditions)

AirQualityTime has two main features:

- It can display the current air quality
- It can display the average air quality per hour that was saved over a period of seven hours.

C.3.3 Additional instructions (only for conditions A, C, D)

The device has a pixel display that shows the measured PM2.5 value (see 1). Using the knob, which is shaped like a leaf, you can switch to different display views. Turning the knob to the shining orange dot, the current air quality is displayed (see 2). Turning the knob to one of the blue dots, the air quality history is displayed (see 3). Each blue dot to the left is one hour backwards in history.



C.3.4 History (all conditions)

Imagine the following scenario: You have been using airQualityTime and saved the data of different places during the day from 6:00 AM to (and including) 12:00 PM. The device still shows this data in the air quality history.

C.3.5 History (only conditions A, B, D)

Imagine you were going on a small trip this morning to explore the city. 6:00h - you were in the kitchen at home and ate breakfast there 7:00h - you were waiting for the bus at the bus stop going to the city centre 8:00h - you arrived at the city centre and walked around for a while 9:00h - you were inside a museum 10:00h - you were at the cafeteria of that museum 11:00h - you were walking outside in the city centre towards the bus stop 12:00h - you were going back home and were waiting at the bus stop

C.3.6 Tasks

Please, take the device and explore it. You can use the matches to play with the current air quality (try not to burn anything or yourself!). When ready, perform the tasks and answer the questions. You are allowed to scroll back and reread the information mentioned above if needed. Take your time, there is no time limit to the tasks.

- What is the current air quality? (you may give an estimation)
- When was the air quality at its best?

C.4 Questions

C.4.1 Technology Supported Reflection Inventory

Please answer the statements below. (These statements are answered on a 7-point Likert-scale). Possible answers: 1. Strongly Disagree, 2. Disagree, 3. Somewhat Disagree, 4. Neutral, 5. Somewhat Agree, 6. Agree, 7. Strongly Agree.

- Q1.1 Using the system has led to a wake-up call to make changes in my life
- Q1.2 As a result of using the system, I have changed how I approach things
- Q1.3 Using the system gives me ideas on how to overcome challenges
- Q1.4 I enjoy exploring my data with the system
- Q1.5 The system makes it easy to get an overview of my personal data
- Q1.6 The system makes it easy to review my long-term personal data
- Q1.7 I reflect on my data in the system with others
- Q1.8 The system helps me to discuss my data with others
- Q1.9 The system makes me think about how my personal data relates with that of others

C.4.2 Self-Reflection and Insight Scale

Please answer the statements below. (These statements are answered on a 6-point Likert-scale). Possible answers: 1. Strongly Disagree, 2. Disagree, 3. Somewhat Disagree, 4. Somewhat Agree, 5. Agree, 6. Strongly Agree.

- Q2.1 I don't often think about my thoughts
- Q2.2 I rarely spend time in self-reflection
- Q2.3 I frequently examine my feelings
- Q2.4 I don't really think about why I behave in the way that I do
- Q2.5 I frequently take time to reflect on my thoughts
- Q2.6 I often think about the way I feel about things
- Q2.7 I am not really interested in analyzing my behaviour
- Q2.8 It is important for me to evaluate the things that I do
- Q2.9 I am very interested in examining what I think about
- Q2.10 It is important to me to try to understand what my feelings mean
- Q2.11 I have a definite need to understand the way that my mind works
- Q2.12 It is important to me to be able to understand how my thoughts arise

C.4.3 User Engagement Scale - Short Version

Please answer the statements below. (These statements are answered on a 5-point Likert-scale). Possible answers: 1. Strongly Disagree, 2. Disagree, 3. Neutral, 4. Agree, 5. Strongly Agree.

- Q3.1 I lost myself in this experience
- Q3.2 The time I spent using airQualityTime just slipped away
- Q3.3 I was absorbed in this experience
- Q3.4 I felt frustrated while using airQualityTime
- Q3.5 I found airQualityTime confusing to use
- Q3.6 Using airQualityTime was taxing
- Q3.7 This airQualityTime was attractive
- Q3.8 This airQualityTime was aesthetically appealing
- Q3.9 This airQualityTime appealed to my senses
- Q3.10 Using airQualityTime was worthwhile
- Q3.11 My experience was rewarding
- Q3.12 I felt interested in this experience

C.4.4 System Usability Scale

Please answer the statements below. (These statements are answered on a 5-point Likert-scale). Possible answers: 1. Strongly Disagree, 2. Disagree, 3. Neutral, 4. Agree, 5. Strongly Agree.

- Q4.1 I think that I would like to use this system frequently
- Q4.2 I found the system unnecessarily complex
- Q4.3 I thought the system was easy to use
- Q4.4 I think that I would need the support of a technical person to be able to use this system
- Q4.5 I found the various functions in this system were well integrated.
- Q4.6 I thought there was too much inconsistency in this system
- Q4.7 I would imagine that most people would learn to use this system very quickly
- Q4.8 I found the system very cumbersome to use
- Q4.9 I felt very confident using the system
- Q4.10 I needed to learn a lot of things before I could get going with this system

C.4.5 Perceived Creepiness of Technology Scale

Please answer the statements below. (These statements are answered on a 7-point Likert-scale). Possible answers: 1. Strongly Disagree, 2. Disagree, 3. Somewhat Disagree, 4. Neutral, 5. Somewhat Agree, 6. Agree, 7. Strongly Agree.

Q5.1 I think that the designer of this system had immoral intentions

Q5.2 The design of this system is unethical

Q5.3 Using this system in public areas will make other people laugh at me

Q5.4 I would feel uneasy wearing this system in public

Q5.5 The system looks bizarre to me

Q5.6 This system looks as expected

Q5.7 I don't know what the purpose of the system is

Q5.8 This system has a clear purpose

C.5 Demographics

Q6.1 What is your age?

Q6.2 What is your gender? (Male/Female/Non-binary/Prefer not to say)

Q6.3 What is the highest degree or level of school you have completed? (no formal education / high school / post-secondary vocational education (MBO) / University of Applied Sciences Bachelor (HBO) / University Bachelor (WO) / Doctorate/PHD / Other, please specify.

C.6 Updates on results

Q7 If you would like to receive the results of this study, please fill in your email address. Otherwise, skip this question.

D Qualitative data

In this section we present the qualitative data of the results. Table 11 presents the overall impression of the participants. Table 12 presents an overview of the participants' intention on how to use the device.

P.	Impression		P.	Impression	
	Usability	Overall		Usability	Overall
A01	Learning curve	Funny	C01	No opinion	No opinion
A02	Easy to use	Nice	C02	Learning curve	Nice
A03	Easy to use	Interesting	C03	Easy to use	Nice
A04	No opinion	Funny	C04	No opinion	No opinion
A05	No opinion	Funny	C05	Learning curve	Nice
A06	Learning curve	Nice	C06	Easy to use	No opinion
A07	Learning curve	Funny	C07	No opinion	Interesting
A08	No opinion	Nice	C08	No opinion	Funny and nice
A09	No opinion	Interesting	C09	Confusing	No opinion
A10	No opinion	Nice	C10	Easy to use	Weird
A11	No opinion	Funny	C11	Easy to use	Funny
A12	Easy to use	Funny and nice	C12	Confusing	Nice
A13	Easy to use	No opinion	C13	Learning curve	Funny
A14	No opinion	Funny	C14	Learning curve	Nice
A15	Easy to use	No opinion	C15	No opinion	Interesting
A16	Learning curve	Funny and nice	C16	Learning curve	No opinion
A17	Easy to use	No opinion	C17	No opinion	Nice
A18	No opinion	Funny	C18	No opinion	Nice
A19	No opinion	Nice	C19	Learning curve	Nice
A20	No opinion	Interesting	C20	Easy to use	No opinion
A21	No opinion	Funny	D01	No opinion	Nice
B01	Confusing	No opinion	D02	Confusing	No opinion
B02	Learning curve	Funny and nice	D03	Confusing	Funny
B03	Learning curve	Interesting	D04	Learning curve	Nice
B04	No opinion	Nice	D05	Learning curve	Interesting
B05	Learning curve	Interesting	D06	Learning curve	Interesting
B06	Easy to use	Funny	D07	Learning curve	Funny
B07	No opinion	Funny	D08	No opinion	Interesting
B08	No opinion	Funny and interesting	D09	Confusing	No opinion
B09	Learning curve	Funny	D10	Easy to use	Interesting
B10	Confusing	Weird	D11	Learning curve	Funny and interesting
B11	Learning curve	Funny	D12	Easy to use	Interesting
B12	Learning curve	No opinion	D13	Learning curve	Funny and nice
B13	No opinion	Funny and nice	D14	Learning curve	Interesting
B14	Learning curve	Interesting	D15	Learning curve	No opinion
B15	No opinion	No opinion	D16	No opinion	Nice
B16	Learning curve	No opinion	D17	No opinion	Interesting
B17	Learning curve	Nice	D18	Easy to use	Funny
B18	Confusing	Interesting	D19	No opinion	Interesting
B19	Confusing	Funny and nice	D20	Learning curve	Funny
B20	Confusing	Funny and nice	D21	Confusing	No opinion

Table 11: The participants' overall impression of the device, per participant number.

Category	Code	Total (N)	A (N)	B (N)	C (N)	D (N)
<i>General intention</i>	Not use it	23	8	5	6	5
	Curiosity	9	3	3	2	1
	Comparison	9	3	3	3	0
	just informative	3	0	3	0	3
<i>Specific intention</i>	Opening a window	9	1	3	5	2
	Cigarette smoke	6	1	4	1	0
	Other smoke	3	2	1	0	0
	Dust	3	3	0	0	0
	Medical	3	2	0	0	1
	Plants	2	1	0	1	0
	Crowds	2	2	0	0	0
<i>Location</i>	Home	24	7	5	6	6
	Kitchen	9	3	2	2	2
	Public spaces	9	2	3	1	2
	Factory	6	3	0	3	0
	Work	4	1	1	0	1
	Bathroom	3	2	0	1	0
	Highway	3	1	0	1	1
	Public transport	3	1	1	1	0
	Car	2	0	0	0	2
Airport	1	1	0	0	0	

Table 12: The participants' intention to use the device, categorized by general intention, specific intention and location. Columns A, B, C, D (N) represent the amount of participants that interacted with condition A, B, C, or D.