LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN
Department "Institut für Informatik"
Lehr- und Forschungseinheit Medieninformatik
Prof. Dr. Florian Alt

Masterarbeit

Design and Development of a Public Display Survey Platform

Lukas Ziegler
lukas@lukasziegler.com

Bearbeitungszeitraum:	6.10.2014 bis 30.04.2015
Betreuer:	Prof. Dr. Florian Alt, Jiamin Shi
Verantw. Hochschullehrer:	Prof. Dr. Florian Alt

Zusammenfassung

Kurzzusammenfassung der Arbeit, maximal 250 Wörter.

Abstract

Short abstract of the work, maximum of 250 words.

Rough outline

- Interactive screens in public are already deployed in urban environments, such as airports, shopping malls, train stations,
- This new form of XXX, not only beings a new way of multimedia presentation with it, but also enables a new variety of interactive applications and a new way of collecting quantitative and qualitative data.
- In this thesis we discuss ... (what is the central topic of the thesis?)
- We provide ... / We contribute ...

A shorter way, inspired by MyPosition [21]:

1. We **present** the design and development process of *PDSurvey*, an interactive survey platform for public displays. Additionally we propose a taxanomy of standardized questionnaires, usable for public display evaluation.
2. PDSurvey aims to facilitate the conduction of surveys on public displays...
3. We **evaluated** PDSurvey regarding two first research questions, <feedback channel, (question type), motivation for approaching>.
4. We **found** that most users prefer to complete survey directly on-site.

Inspiration

- Mueller 2010: Requirements and Design Space for Interactive Public Displays [15].
- MyPosition [21]

Aufgabenstellung

Development of a Public Display Survey Platform

Problem Statement Public displays are quickly proliferating in public space. At the same time, interactive applications are still scarce, since their development is costly and the effect on the user - and thus their benefit - is often not clear. Hence, interactive displays applications are usually developed, deployed, and carefully evaluated in research contexts. In most cases, evaluation focusses on particular aspects only, such as user performance, user experience, or social implications, due to the significant effort associated with planning, preparing and conducting public display evaluations.

Scope of the Thesis To tackle the aforementioned challenge, the objective of this thesis is to develop a survey tool that allows interactive public display installations to be comprehensively assessed. In a first step, an extensive literature review will be conducted with the aim to identify important aspects of public display deployments - both from a researcher as well as from a practitioners' perspective - as well as to develop an understanding of how these aspects could be addressed through surveys. Based on the literature review, a web-based survey platform will be implemented that can easily be used to evaluate and compare public displays through different channels. Such channels include both evaluation directly at the display or through a (mobile) website that allows participation also via a smartphone or tablet. The platform should allow public display owners to configure their own surveys based on their needs. Optionally, the survey tool itself will be evaluated with an interactive public display application.

Tasks (1) conduct a literature review to identify (research) questions that are of interest to researchers and practitioners
(2) produce a comprehensive set of questions that can be used to assess these questions by means of a survey
(3) develop a web-based public display survey platform consisting of (a) an administration interface that allows (groups of) questions to be selected for use within the tool and (b) a responsive UI that can be rendered on different devices (public display, smartphone, tablet, laptop)

Requirements Strong skills in web programming, independent scientific work and creative problem solving, experience in creating questionnaires is a plus.

Keywords Public displays, interaction, applications, survey, questionnaires, web

Ich erkläre hiermit, dass ich die vorliegende Arbeit selbstständig angefertigt, alle Zitate als solche kenntlich gemacht sowie alle benutzten Quellen und Hilfsmittel angegeben habe.

München, April 6, 2015

. . . . . .

Dedication / Acknowledgments

I would like to thank Jiamin Shi, Axel Hösl and Dr. Julie Wagner and the entire Mediainformatics chair, for their great spirit and support. It was an honor to be able to complete my Master studies and thesis at the chair. Additionally I would like to thank my parents how they have supported me, not just throughout the last five years, but especially during my studies. Much would not have been possible without their love and support.

But above all the thanks goes to Prof. Dr. Florian Alt, for his guidance, feedback, support and the time devoted to my Master thesis.

Contents

1 Introduction

1.1 Outline

How Ismail approaches it

1. Motivation
2. Problemstellung
3. Zielsetzung
4. Ueberblick

Notes gathered throughout the last weeks

1. Introduction to the topic: there seems to be a clear trend towards more public displays and an increasing ratio of ...
2. **Context and Goal of Thesis**: State the goal of the paper: why the study was undertaken / the paper was written. Do not repeat the abstract.
3. Sufficient **background information** to allow the reader to understand the context and significance of the question you are trying to address.
4. *not sure about this one...*: Proper acknowledgement of the *previous work* on which you are building. Sufficient references such that a reader could, by going to the library, achieve a sophisticated understanding of the context and significance of the question. » doesn't this belong in the next chapter?
5. **Motivate** your reader to read the rest of the thesis. You should draw the reader in and make them want to read the rest of the paper. An important/interesting scientific problem that your paper either solves or addresses.
6. Already state the **research question** here!?!? I would say so, because introducing it in chapter 4 is a little to late. How about stating the general research question here ("How to automatically evaluate public display setups via surveys carried out on the displays themselves."), and in chapter 4 getting more specific and talking about the 1) feedback channel, 2) question types, 3)motivation for approaching the display, and 4) the categorization of standardized questionnaires (literature review).
7. **Give Overview**: A verbal *road map* or verbal *table of contents* guiding the reader to what lies ahead.
8. **where??**: Explain the scope of your work, what will and will not be included.

Important things to keep in mind!

1. Make sure that it is obvious where the introductory material (the *old stuff*) ends and where your contribution (*new stuff*) starts?

1.2 First draft (notes)

Scope for the practical part of the thesis. What is the system supposed to look like. What is the current state of the practical work in research and in the industry. Which partners do we have, where do we want to deploy this system. What are the goals for this evaluation platform?

1. Describe the increase of interactive displays in public, get the shift to airports, shopping malls, universities. Growing research area.
2. What is the current situation like?

1.3 not sure where to mention it

1. What are public displays?
2. Survey vs Questionnaire

1.4 Research Question

1.5 Approach

Benefits of PDSurvey platform:

- being able to work with big data. collecting a large number of responses from a variety of displays in various settings, and assigning a specific context to every display connected to PDSurvey. Once enough data is collected, having the ability to evaluate and compare the displays between each other. Interesting questions for analysis would be, which role the context plays on how the users behave, when running identical software settings on the displays, but only varying the context (position, size of display, surrounding environment of the display, positioning it outdoors or indoors, influence of the weather, type of building it is positioned in) – see Chapter 4: Modeling

1.6 Overview

1.7 Our Contribution

1. categorization of questionnaires being useful for the evaluation through automated public survey display platforms
2. research platform

1.8 Other introductions for inspiration

1. paper 60 mueller2010mm - Requirements and Design Space for Interactive Public Displays: Good introduction from Mueller et al. [15]: “Digital immersion is moving into public space. [...] analogies to HCI (utility, usability, likeability) often hold true, but not always! These need to be evaluated individually.”
2. Another good introduction as an inspiration: “MyPosition: Sparking Civic Discourse by a Public Interactive Poll Visualization” » check out the introduction there! [21]
3. paper 18 cites Mark Weiser, on his predictions regarding Ubiquitous Computing and how it will be the third revolution after main frame (1st) and personal computer (2nd). for the quote see GDrive ‘Quotes and Notes from Literature’
4. copy further notes from https://docs.google.com/document/d/1f2MJHMh5Yvvh9d4hhIT0WmgdlBbZ9rX5b

2 Background and Related Work

Our research is based on an extensive literature review, with over 100 papers viewed. This led to the development of the public display survey platform (see chapter 3) and the categorization of standardized questionnaires (see section 2.2). A short summary of the most relevant paper is described in section 2.1. The goal of the literature review was to find out how other researchers evaluated public displays and to develop an understanding of how these aspects can be addressed through surveys. The aim was to identify important aspects of public display deployments, both from a researcher's as well form a practitioner's perspective.

2.1 Related Work

Public display evaluation has already been addressed in literature. Alt et al. [3] give an overview of study types, research paradigms, and evaluation methods used for evaluating public displays. Müller et al. present with MirrorTouch [16] an updated evaluation and additionally extract metrics used for quantitative field studies. According to their findings almost exclusively descriptive field studies are used in the area of public display evaluation. For a more in-depth introduction to public display evaluation the doctoral thesis of Alt [2] gives a good overview. For a general recap of how to best design, evaluate, and report experiments, the book by Field and Hole [8] was used. There, they lay a solid basis for evaluation. A good introduction for conducting surveys for practitioners provides the list of FAQs by Kirakowski [13].

Papers with a good approach for evaluation and for your own design:

Papers which inspired us and have a good approach to the evaluation of public displays:

Papers on which we did a meta-analysis, analyzing how they approached the evaluation via questionnaires. A list of exemplary papers with a good evaluation: Overcoming Assumptions by Huang et al. [10], Worlds of Information by Jacucci et al. [12], and Looking Glass by Müller et al. [17].

Based on the previous evaluation of related work and the time-consuming nature of field studies (even for small quantitative questionnaires) the demand for a simplification or automation of the evaluation process becomes clear. Due to the essential importance for the validation of public display installations and research in general, a toolset such as the PDSurvey platform can be a great relief.

One constraint of public display research represents the opportunistic nature of the setups and the discrepancy between lab studies and field studies [18]. Thus there is an additional demand for evaluating each public display setup individually and if possible directly in the field.

paper 23: Overcoming Assumptions and Uncovering Practices (Huang) paper 25: Worlds of Information (Jacucci): Best-practices / guidelines for designing good public display applications

A set of related survey platforms aiming for a similar goal, which differ from our approach:

LimeSurvey [1]

SosciSurvey [2]

eSurv `http://esurv.org/`

eSurvey Creator [3]

SurveyMonkey [4], might be the toughest opponent based on the functionality of the tool. But it is not visible that they offer running dedicated public displays in a dedicated mode for the evaluation platform. One They also offer a set of sample surveys, available for use in their platform

[1] `http://de.wikipedia.org/wiki/LimeSurvey` (last visited on November 26, 2014)

[2] `https://www.soscisurvey.de/` (last visited on November 26, 2014)

[3] `https://www.esurveycreator.com/` (last visited on April 6, 2015)

[4] `https://www.surveymonkey.com` (last visited on April 6, 2015)

[5], however they

An overview of additional web-based survey tools[6].

UX Suite by UsabilityTools [7], has a handy backend for configuring the surveys, although it is lacking the pre-configured standardized questionnaires.

SoGoSurvey [8]. One of the best solutions!

Free Online Surveys is missing some question types (e.g. the Likert scale) and freeonlinesurveys.com

SurveyPlanet () seems to be the best solution found so far! However the embed code is only based on iframes.

Qualtrics [9]

+++ clarify what the difference is between the already existing approaches to our approach. +++ REST API is missing, not designed for use on public displays. Thus no overlay as intended with the PDSurvey embed code is possible.

Unique about our approach is the fact that we will have the opportunity to conduct surveys across a broad number of devices and platforms, due to its modular and RESTful architecture. The benefit is that the whole platform retrieves all data via a RESTful API, allowing the greatest possible coverage of public displays and end consumer devices. Additionally we offer a range of standardized questionnaires, which specifically aim at usage on public displays, introduced in section 2.2.2.

2.2 Background

As a result of the literature review, besides getting a better understanding of how public displays were evaluated, a side effect was getting an overview of the questions asked to evaluate public displays and their applications. This turned out to be a quite valuable approach, since we haven't seen any compilation of questionnaires used for public display evaluation so far. Our goal was to find patterns and to build clusters of questionnaires being useful for the evaluation through automated public survey display platforms.

In the following we will first describe our methodology for gathering the information (section 2.2.1), followed by a categorization of standardized questionnaires (section 2.2.2), and wrapped up by a summary of our findings (section 2.2.3). The categorization of the standardized questionnaires can be found in Table 2.2.2.

2.2.1 Methodology

The procedure for the selection of papers to review, was as follows. As a starting point all papers form the appendix of Florian Alt's doctoral thesis [2] were read. Afterwards interesting related work and citations were followed based on the papers from the previous step. This was supplemented with targeted research on Google Scholar and the APM Digital Library. To round off the literature review the publications of two authors, who are very active in this field, were reviewed. A full list of all papers reviewed can be found in Appendix D.

The first step, the analysis of the appendix was fairly straight forward. All papers were read from start to finish (pages 335 to 343), in order to get a first overview of the current state of research. The second step, pursuing related work and citations of interest, was carried out in a more subjective manner. Whenever interesting papers or projects were mentioned, the cited paper

[5] https://www.surveymonkey.com/blog/en/sample-survey-questionnaire-templates/ (last visited on April 6, 2015)

[6] http://www.capterra.com/survey-software/ and http://www.idealware.org/articles/fgt_online_surveys.php (last visited on April 6, 2015)

[7] http://usabilitytools.com/ux-suite/ (last visited on April 6, 2015)

[8] http://www.sogosurvey.com/Features/List-of-All-Features.aspx (last visited on April 6, 2015)

[9] http://www.qualtrics.com/site-intercept/ (last visited on April 6, 2015)

was also skimmed through. For the third step, a more strategic approach was used. Based on the insights gained from the previous steps, Google Scholar and APM was checked for literature relevant to our research question. The Keywords, that were used amongst others for the research in these online libraries, were:

- Standardized Surveys for Usability
- Standardized Surveys for User Experience
- user satisfaction questionnaire
- public display evaluation
- standardized public display evaluation

The last step for collecting relevant papers consisted of profiling publications of two relevant authors in the area of public display research, namely Jörg Müller and Marcus Foth. The process started out by first finding a list of their publications. Since the literature review made by Florian Alt (see first step) already covered papers up to 2011, only ones published between 2012 and 2014 were viewed. On each opened papers from this time frame a keyword search was carried out, to see whether it contained an evaluation which might be relevant for us. These keywords were: *questionnaire*, *survey*, *question*, *interview*, *(field) study*, and *evaluation*. If none of these words could be found, the headlines and the abstract was skimmed through. All papers containing a reference to evaluating public displays were saved locally and analyzed in more detail. For Jörg Müller the best list of his publications were found on his personal website[10], and for Marcus Foth two websites were evaluated [11].

2.2.2 Standardized Questionnaires

All questionnaires found during the literature review phase were categorized into a schema, inspired by the research questions introduced by Alt [2] (chapter 2.8.2), serving as a guideline for our classification of standardized questionnaires. Since the categories *audience behavior* and *user performance* from Alt can not be evaluated through questionnaires, the research questions are not represented in the following. We extended the prior categorization with findings from the literature review phase, described in the previous section 2.2.1,

During the literature review phase a comprehensive list of widely used questionnaires was assessed. Other people's collections incorporated into our categorization can be found in the bibliography [6, 7, 9, 14, 22].

A full overview of all standardized questionnaires found in literature can be found in Table 2.2.2, grouped by the following categories: user experience, usability, user acceptance, user performance, display effectiveness, privacy, social impact, context, and demographics.

A list of other people's collections can be found here: Lewis and Sauro [14] list discuss 19 questionnaires at the HCI conference, Garcia [9] describes the SUMI, PSSUQ, and SUS questionnaire.

User Experience User experience describes the overall satisfaction and experience the user has with a display. The evaluation can be carried out through questionnaires.

Usability The category was added

[10] `http://joergmueller.info/publications.html` (last visited on November 17, 2014)

[11] `http://www.vrolik.de/publications/` (last visited on November 18, 2014) and `http://eprints.qut.edu.au/view/person/Foth,_Marcus.html` (last visited on November 18, 2014)

User Acceptance User acceptance analyzes user's motives and incentives for approaching the display. The evaluation can be carried out qualitatively (subjective feedback, focus groups) or quantitatively (questionnaires).

Display Effectiveness Display effectiveness evaluates the economic perspective of display efficiency.

Privacy Privacy takes a look at the users privacy concerns.

Social Impact Social impact considers everything related to social behavior, the influence on social interaction and communities, as well as social effects.

Context One category new category, is the collection of context data, relative to the public display. On most normal studies the context doesn't change during evaluation and thus is not as important. For the evaluation of public displays, especially when multiple displays are deployed in different locations running the same application, it will become if importance to also assess the static and dynamic context of each deployed display. External influences such as the weather, time of day, special events or semester break can have an influence on the number and type of people passing by a display. Additionally static context parameters, such as the display type, display size, position on wall and inside of the room might also influence how the display setup is perceived in public. Once recorded, these static and dynamic parameters can be evaluated with knowledge discovery algorithms for big data, a whole research field for itself. So far no previous works are known on this area so far, evaluating a large public display deployment based on their difference in context.

Demographics In most surveys background information about the participants is also of interest. This varies from general questions (gender, age, religion, education), more personal questions (relationship status, family, children, country of origin), skills (personal, professional, technical), personal beliefs, political affiliation or voluntary commitment.

Three background questionnaires, which we haven't used ourselves yet, but which go more in depth, are the Adult Literacy and Lifeskills Survey (ALL) [12], the PIAAC Conceptual Framework of the Background Questionnaire Main Survey [13] and a Police Background Questionnaire [14].

Miscellaneous All questions and questionnaires, which can not be assigned to any of the previous categories, belong to this category.

Cheverst et al. [5] evaluated whether there were any previous experience with Bluetooth, or recommendations for possible new features. For the evaluation of the Digifieds platform Alt et al. also evaluated: "We asked them about their mobile phone usage (e.g., how often they used it, if it had a touch screen, if they used it to surf the web, and if they had installed third party apps) and whether they had used the UbiDisplays before" [4].

1. POST OR REFERENCE THE TABLE WITH THE CATEGORIZATION HERE
2. My Categorization: `https://docs.google.com/document/d/1D925jJ7bmRc1EZdCTz32lmW2hniMiq7GzBWxX8rmhpE/edit` (Google Docs)

[12] `http://nces.ed.gov/surveys/all/` (last visited on April 1, 2015)

[13] `http://www.oecd.org/site/piaac/PIAAC(2011_11)MS_BQ_ConceptualFramework_1Dec2011.pdf` (last visited on April 1, 2015)

[14] `http://www.slmpd.org/images/hr_forms/commissioned/BackgroundQuestionnaire.pdf` (last visited on April 1, 2015)

2.2.3 Findings

Findings:

1. use both quantitative and qualitative methods for data collection (explain why this is important, teaser this as a requirement for the platform, how it could be implemented)
2. support mutliple sections, all displayed at once or (optionally) spread across multiple users
3. support various question types (e.g. 5-point and 7-point Likert scale, multiple choice, numeric responses, comments)

These findings bring us to the next chapter, the research platform to develop, capable of conducting all of these questionnaires.

3 Implementation

In this chapter we will deal with the infrastructure and technical realization of the public display survey platform. First off, we will start with the requirements for the survey platform (section 3.1). Subsequently the architecture resulting from the design decisions will be the main focus (section 3.2). To facilitate the training period for successors we will also take a brief look at the software model (section 3.3.5). For more specific information and for information regarding maintenance of the project, please refer to the Documentation found on the CD enclosed or on the GitHub repository (see Appendix B).

In figure 3.2 a brief overview of the *PDSurvey* platform and its components is given. The platform consists of three major parts: a backend for display providers (PDAdmin), a RESTful server (PDServer) and the user interface itself, being embedded on the end user devices (public displays, tablets, smartphones or other devices).

3.1 Requirements

The starting point for the PDSurvey platform and the Master's thesis itself was the official announcement[15], describing the scope of the thesis. This problem statement already included first requirements for the survey platform to develop, and was also a trigger for further literature research and talks with people from the industry.

1. development of a survey tool that allows interactive public display installations to be comprehensively assessed
2. a web-based survey platform will be implemented that can easily be used to evaluate and compare public displays through different channels
3. different channels to support: 1) evaluation directly at the display or 2) through a (mobile) website that allows participation also via a smartphone or tablet.
4. configuration options for public display owners

Additional requirements, dervived from the previous and complemented with additional ones, are listed below:

- easy **embedding of questionnaires** on websites of public display owners (provide API / embed code)
- support of **various devices**: public displays of all sizes, tablets, phablets, smartphones, desktop devices (responsive web design)
- create an **open research platform** (host project and documentation on GitHub, release it as open source, publish and invite fellow researchers)
- allow for both **quantitative and qualitative** metods of data collection

These derived requirements had an impact on the chosen architecture, which will be discussed in the following.

3.2 Design Decisions

After having assessed all requirements for the platform (see section 3.1), the next step was making design decisions for the software, programming language and frameworks to use, before starting with the practical implementation of the platform.

[15] `http://www.medien.ifi.lmu.de/lehre/arbeiten/detail.xhtml-php?pub=alt_pdsurvey` (last visited on March 24, 2015)

Two weeks were taken for assessing all of the possibilities, on the one hand to get informed again what is buzzing, on the other hand because every choice also has a significant impact on the later architecture. Changing a technology in between would not be a smart choice in such a short time-frame.

Programming language Due to the requirements and objective to support a large number of devices, operating systems, and form factors, a device-independent programming language was preferred. The choice fell on Javascript, not just due to the growing popularity[16], but also because it can be used on the largest number of devices. Furthermore another huge benefit, only having to use JavaScript on all tiers, from client to server to persistence layer.

This approach has become very popular in recent years, now often being referred to as the MEAN stack, consisting of MongoDB, Express.js, Angular.js, and Node.js. Some fundamental differences to the LAMP stack (Linux, Apache, MySQL, PHP) are its shift form server-side to client-side single-page applications (SPA), faster prototyping, shift from synchronous to asynchronous, fast page loading times, less time spent writing SQL (schemaless), and the shift to using RESTful services for the backend[17]. Additional, the feedback from the industry contacts also played a role.

Alternative languages considered were: PHP, Python, Ruby, Java and ASP.NET. The biggest drawback was the additional workload on having to maintain the object model on multiple platforms. With Javascript it is possible to work with the same object model and easily keeping it consistent across all platforms (backend, frontend, server).

Based on these requirements and the feedback received from industry experts, a choice towards Node.js and to fully go along with MEAN stack was self-evident.

Frontend The next question to be answered was which technologies to use in the frontend, leading to the question whether to follow the single-page application approach or not. As of 2014 the JavaScript model-view frameworks most frequently used for creating single-page apps are Angular.js, Ember.js and Backbone.js. Purely based on numbers Angular.js is the clear favorite, it has by far the largest user base on GitHub, Stackoverflow, and Youtube[18]. When comparing the number of third-party modules, Angular.js also takes the lead with 800 ngmodules vs. 236 Backbone.js backplugs vs. 21 emberaddons. All these factors together indicate a short training time and give hope for making good progress for beginners. These were amongst other things the reasons why we chose Angular.js for this project, since the desired requirement is also for other students to further enhance the PDSurvey platform in the future.

- two-way data-binding, syncing, templating, being able to create your own html-tags with custom directives.

To speed up frontend development we chose Bootstrap[19] as our CSS framework of choice. Reasons for choosing Bootstrap were the large community, extensive documentation with helpful examples, large number of free tutorials and templates, its integration with Angular.js (AngulatStrap[20] and AngularUI), and its short training time. Alternatives considered were Foundation Framework by Zurb, however at the time of writing there was no prefabricated integration for Foundation and Angular.js. A good overview[21] and a comparison[22] of currently popular frontend

[16] `http://www.sitepoint.com/javascript-internet-things/` (last visited on November 27, 2014)

[17] `http://www.ibm.com/developerworks/library/wa-mean1/index.html` (last visited on March 26, 2015)

[18] `https://www.airpair.com/js/javascript-framework-comparison` (last visited on January 11, 2015)

[19] `http://getbootstrap.com/` (last visited on December 1, 2014)

[20] `http://mgcrea.github.io/angular-strap/`

[21] `http://www.sitepoint.com/5-most-popular-frontend-frameworks-compared/` (last visited on December 2, 2014)

[22] `http://www.sitepoint.com/grid-system-comparison-bootstrap-vs-foundation/` (accessed March 24, 2015)

frameworks was also considered.

Backend For the backend it was most important to have a solid and scalable solution with good performance, since our system might need to scale in the future, having a multiplicity of clients attached to the survey platform, all submitting responses and querying for different questionnaires. Additionally an administrator interface and easily being able to exchange data with all clients was of importance. For this reason a backend built solely on the principles of a RESTful API was preferred, being able to use the same data on no matter which platform.

Based on the decision to use JavaScript for all tiers, it was clear to also use Node.js as the underlying platform for building web applications. Reasons speaking for Node.js are its scalability and modular approach, only requiring the parts needed for your project. Another benefit is the easy implementation of authentication or internationalization, due to the concept of middlewares [23].

(Discussion of pro/con Node.js: http://www.heise.de/developer/artikel/2x-Nein-4x-Ja-Szenarien-fuer-Node-js-2111050.html)

The question which web framework to use for Node.js was not as straight forward to answer. A list of current options being Express.js, Conncet, Koa and Restify.

+ say which one I chose why

To facilitate the object modeling process Mongoose[24] was chosen. Mongoose is a object modeling package for Node.js, allowing to model the application data based on schemas. This simplifies making incremental changes to ... and keeping the model synchronized across all layers. Mongoose also takes care of performing CRUD applications.

Due to the decision to build a single-page application it became vital to separate the data from presentation layer. Using a RESTful service is the current de facto standard. An alternative would be to use SOAP for message exchange, however leading to an increase of data overhead, more logic on server.side, and the loss of statelessness.

Should a component not support HTML or JavaScript execution, then the required surveys can still be communicated directly with the REST API through rudimentary HTTP function calls, being another benefit for using a RESTful API.

Database Another fundamental aspect presented the question where to store the data persistently. Criteria for choosing the right database management system (DBMS) for this project were again the size of community, suitability for prototyping, and ease of integration with Node.js/Angular.js.

The first question presented whether to choose a SQL or a NoSQL DBMS. Because NoSQL is better for rapid prototyping, its Schema can be mixed inside of one collection and develop over time. Combined with the benefit of having a better scalability, a schemaless data representation, faster response time and a decreased development time [20], these are all reasons speaking for using a NoSQL DBMS for our scenario.

Out of the NoSQL databases MondoDB[25] represents the most popular DBMS, especially since it integrates seamlessly into the MEAN stack. With the help of Mongoose, a object modeling package for Node.js, performing CRUD applications on MongoDB and maintaining a solid object model (schema) gets even easier.

Benefits of MongoDB are being non-relational (and schemaless), plus its ability to directly store JavaScript object inside the database, being the biggest advantage. One disadvantage is that MongoDB does not support joins or transactions. For our use case, however, this is no major drawback. The benefits outweigh the disadvantages.

[23] `http://www.heise.de/developer/artikel/REST-Webservices-mit-Node-js-Teil-1-Connect-als-Fundament-1802258.html?view=print` (last visited on November 24, 2014)

[24] `http://mongoosejs.com/` (accessed November 14, 2014)

[25] `http://www.mongodb.org/`

Alternatives looked at were CouchDB and Redis[26]. Redis being useful for fast changing data, which is not the case for our platform. CouchDB would be an alternative worth looking at, having a better replication and conflict resolution. This additional securitry is however not needed. The speed benefits of MongoDB, also being able to make dynamic queries, are preferred.

Hosting For the hosting of the platform a free and easy scalable solution was of importance. Our first choice was Heroku[27], due to its simplicity of setup, its native support of Node.js and the seamless integration with Mongolab[28], hosting our MongoDB.

Alternative were Google App Engine, IBM BlueMix, Amazon Web Services (Amazon EC2) or hosting everything on local or virtualized machines at our university. However for our scenario all of the above options had their drawbacks in comparison to Heroku. Google App Engine (as of December 2014) still had no native support for Node.js and custom runtimes had to be used to get Node.js support up and running. IBM BlueMix just got overhauled, offered full out-of-the-box Node.js support, however they only the first 30 days were free and the pricing model wasn't as attractive. Amazon Web Services offering a Infrastructure as a Service (IaaS), would have required too much administration of the server, which would have slowed down the main objective of the project, the development of the survey platform. The same goes for the last option, hosting a MEAN-stack environment on our own servers at LMU Munich. All of the above are well-known solutions in the industry, however due to simplicity and ease of use we chose Heroku.

MEAN In the end, as already predicted in the beginning, it turned out to be the classic MEAN stack.

The clear benefit, being able to use JavaScript from client to server to persistence level.

+ a fantastic short recap of the the last ten years of web development with an introduction to the MEAN stack, also explaining what the differences are between traditional web development (using the LAMP stack), to using the MEAN stack: `http://www.ibm.com/developerworks/library/wa-mean1/index.html` (last visited on March 26, 2015)

[26] `http://kkovacs.eu/cassandra-vs-mongodb-vs-couchdb-vs-redis` (accessed March 26, 2015)
[27] https://www.heroku.com/
[28] https://mongolab.com/

3.3 Modeling

The model for the *PDSurvey* platform is maintained with the help of the Node package Mongoose. Node.js maps the route parameters and routes all requests to the corresponding Mongoose model. Angular.js builds its model upon the REST API and maps it via dynamic two-way-binding to it's scope. Thus all changes to the model originate from Mongoose.

3.3.1 Development Process / Modeling

The development process of the PDSurvey platform was inspired and influenced by the following approaches:

- working agile and user centered... TODO: look up how to best explain it.
- User Centered Design: (paper nr 31): "constitutes an iterative process of system design, deployment and evaluation" (quote from paper 31). Work iteratively, continuous deployment and evaluation.
- Concept of extreme programming[29]: First user stories were written and assessed in a small group. The next step was to transfer these stories to user models, describing in detail which functionality the stakeholders of PDSurvey are supposed to have. Later a first software architecture and software model was built, getting more specific. Dependencies between models were defined and this model was continuously refined and improved throughout the dvelopment phase. The last step of the modeling represented screen designs, getting a clear view of what the interface might later look like.
- Used the extreme programming[30] approach: user stories, release planning, release schedule, small releases, iterating
- criteria for good user stories: http://tigertechtalk.wordpress.com/2012/10/17/wie-schreibe-ich-eine-gute-user-story-und-was-ist-das-uberhaupt/

3.3.2 User roles

As of now only two roles are implemented, the admin-role and the guest-role.

In the long term it would be desirable to have the following user roles: Admin, Operator, Evaluator, DisplayApplication.

3.3.3 Software model

In total there are the following classes.

1. list all classes
2. 1 UML diagram is enough, according to Florian

3.3.4 Dependencies

Of special interest are the following four models: Survey, Display, Campaign and Responses.

[29] `http://www.extremeprogramming.org/rules.html` (last visited on November 14, 2014)
[30] `http://www.extremeprogramming.org/rules.html` (last visited on November 13, 2014)

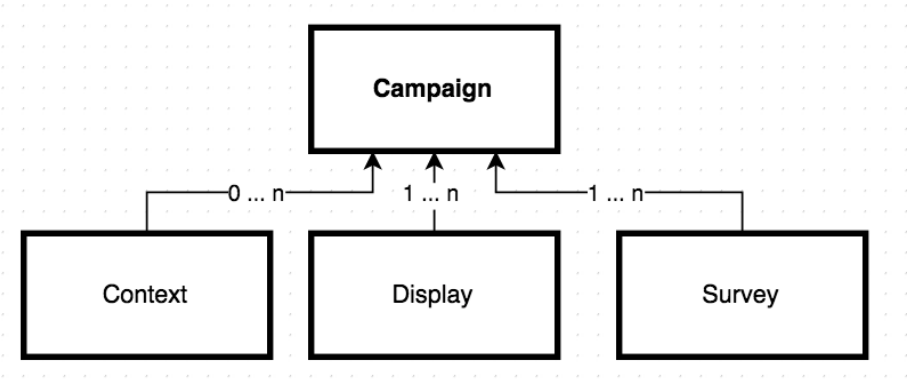


Figure 3.1: Campaign model dependencies

Surveys: Surveys resembles the foundation of PDSurvey, with the aim of reuse and standardization. A survey consists of multiple sections, being built of of multiple questions. Each question is of a corresponding question type and every survey belongs to a category. This allows the filtering for relevant surveys. To be able to create private surveys, not being shared across the entire platform, every survey is assigned to an individual user.

Displays: In the display collection all displays connected to the PDSurvey platform are contained. To allow for an evaluation across multiple display models and based on the context of the displays, the display model and a static and/or dynamic context is assigned to every display.

Campaigns: Campaigns resemble the most integral part, since they glue all of the pieces together and allow the distribution of surveys to public display networks. A campaign consists of displays and surveys and creates the mapping of the questionnaires to public displays. Additionally to each of those mapping an individual context can be assigned, enabling the later comparison of results in between the public displays.

Responses: All responses made to each survey are logged in the Response collection. The queries are carried out individually per user, per display and per campaign. This model will be the base for further extensions, amongst others the automatic evaluation of the survey responses and the comparison inbetween an entire display network, to be able to find out which properties of a display might be related to certain effects.

Context: One of the benefits of creating such a survey platform is being able to collect and evaluate large amounts of data, without having to pay more people for conducting and evaluating the survey. The idea would be to collect a large number of responses from a variety of displays in various settings, and assigning a specific context to every display connected to PDSurvey. Once enough data is collected, having the ability to evaluate and compare the displays between each other. Interesting questions for analysis would be, which role the context plays on how the users behave, when running identical software settings on the displays, but only varying the context (position, size of display, surrounding environment of the display, positioning it outdoors or indoors, influence of the weather, type of building it is positioned in).

++ neue Namensgebung, um in der domain specific language zu bleiben –> application provider / display provider / space provider (anstatt Operator). Wir werden aber nur mit dem Application Provider (anstatt Operator) im System arbeiten

3.3.5 REST interface

Defining the REST API.

Notes from before I started writing:

1. Think about using a Extreme Programming approach http://www.extremeprogramming.org/rules.html

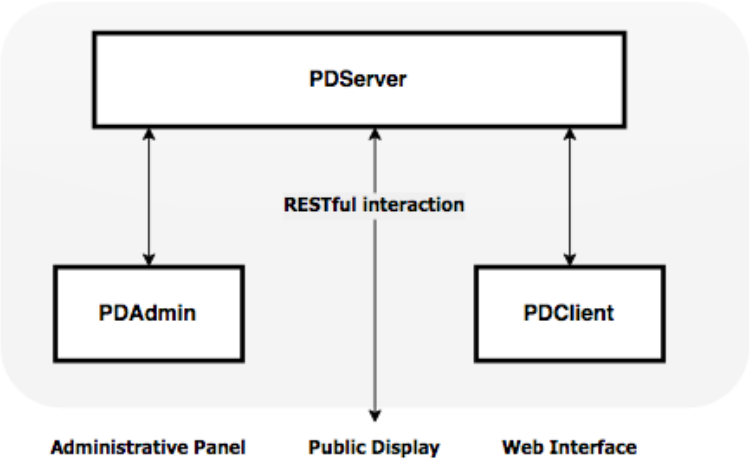


Figure 3.2: Overview of the PDSurvey platform: (a) PDServer containing the Node.js server, (b) PDBackend the Angular.js frontend for administrators and (c) PDClient the interface included on public displays and optimized for mobile devices.

3.4 PDSurvey Platform

+ describe what it is about + describe the main goals / ambitions

The architecture for the public disply survey (PDSurvey) platform can be split into three main sections: PDAdmin, PDServer, and PDClient (see Figure 3.2). *PDAdmin* contains the administrative interface, allowing display providers to configure questionnaires for their public displays. *PDServer* accommodates the REST service, the persistence layer, and the majority of the application logic. *PDClient* encloses

1. begin with describing what the platform does, which views it has (show screenshots).
2. describe the client

3.4.1 PDAdmin

For administrative purposes we created an admin interface, enabling display providers to create, manage and distribute surveys to public displays.

Display providers have the ability to create their own questionnaires or to select from a list of standardized questionnaires (introduced in section 2.2).

advancing > modular aufgebaut

All data and actions get transmitted to the server via HTTP request/response calls.

3.4.2 PDServer

PDServer makes a relatively simple impression. It consists of a Node.js server, which to the outside only acts as a REST server. Processing REST calls, performing CRUD operations and responding with JSON objects. Besides this REST functionality a rudimentary authentication mechanism is already implemented on the server and the capability for further logic, determining which client should ask which question next. This functionality might become of interest when trying to spread standardized questionnaires of longer length across multiple users or multiple displays. It would be intended for the server to keep track which questions have already been answered and to tell each instance of PDClient which question to ask next, in order to achieve a balanced question profile.

The specification of PDServer's REST API can be found in the documentation (see Appendix B).

3.4.3 PDClient

Our client tool was kept as simple and minimalistic as possible. It is running on a separated code base than PDBackend, the only communication between the two is via REST, exchanging JSON objects. Reasons for this were on the one hand reduction of the application size, on the other hand different requirements for constructing a PDAdmin interface for a limited number of users, compared with PDClient, being embedded at large scale.

The goal is to reduce logic and complexity on client-side. Currently PDClient is implemented as follows: The client either receives all questions for the questionnaire (caching the questions), or it loads just the next question on-demand. For this purpose the REST interface provides the /nextQuestion API is provided. For production use one sub-page per campaign gets created, individualized via the campaign ID.

PDClient has three main components (see Figure **??**). The principal part being the *Survey page*. All questions are loaded at once on initial startup, then one question gets displayed at a time. Settings for the survey can be modified in the PDBackend (e.g. number of questions to display and duration of the survey). Once the user makes a choice, it is directly logged on the server. In case that a participant aborts answering the survey, the questions answered so far are still recorded. The *About page* was added, since employees from university gave feedback to us regarding the public display installation, prior to the beginning of development for the PDSurvey platform. They said they were skeptical and had doubts regarding the research project, when having no information whatsoever regarding which information is logged. To motivate people to participate, a *Welcome page* was added. It turned out that a significant larger number of people were willing to participate in a survey, after knowing that it doesn't take long, the research is university-related and that it will be used for a Master's thesis. These arguments were amongst others stated in semi-structured interviews, carried out as part of the field study (see chapter 4).

3.4.4 EmbedCode

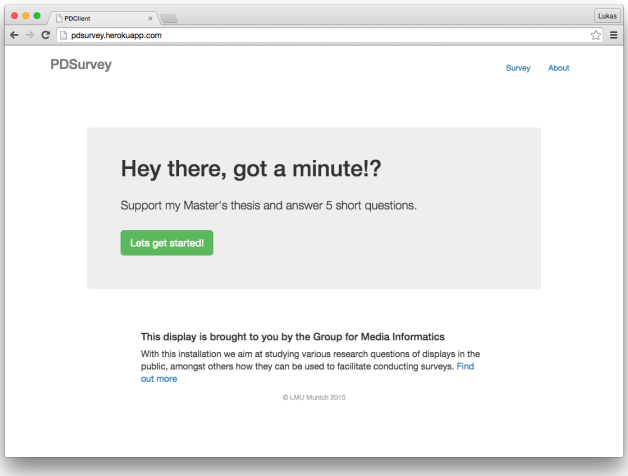
The embed code, JavaScript Code Injection, turned out to be a pure proof-of-concept, since it was not needed for the field study at the university. The problem was that the application on which PDSurvey should be integrated did not support any HTTP calls, thus we had to fall back on another solution. This embed code was intended to be used by display operators, wanting to include optional questionnaires hovering over their normal application.

An example use case is exemplified here ... **TODO TODO TODO**

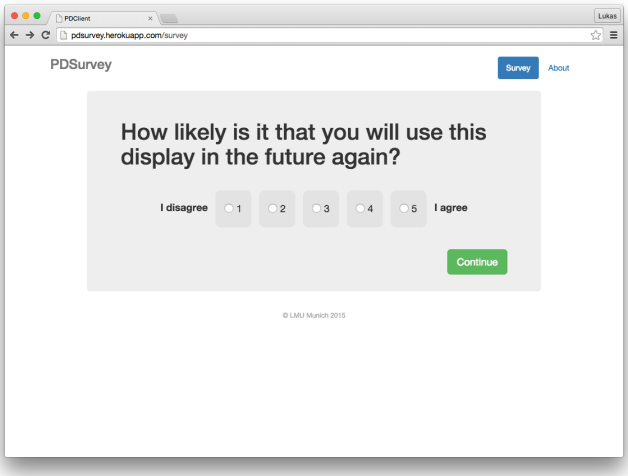
The implementation is quite simple. A JavaScript code snippet will be given to the display provider, which has to be added before the closing body-tag. This minified line of JavaScript code adds a HTML <script>-tag to the DOM of the HTML page, injecting a remote JavaScript file from the PDSurvey platform. This personalized scripts first loads jQuery and/or Angular.js asynchronously, and thereafter creates another instance of the PDClient on client side, inside of the primary website DOM. All questions for the questionnaire get loaded via REST API from the server and the responses get sent back to the server for logging.

One important aspect is to prefix all classes and files with a unique namespace, to prevent any sort of collisions with the main program, where the code gets injected into. For this prototype all CSS Bootstrap classes were prefixed with $pd-$.

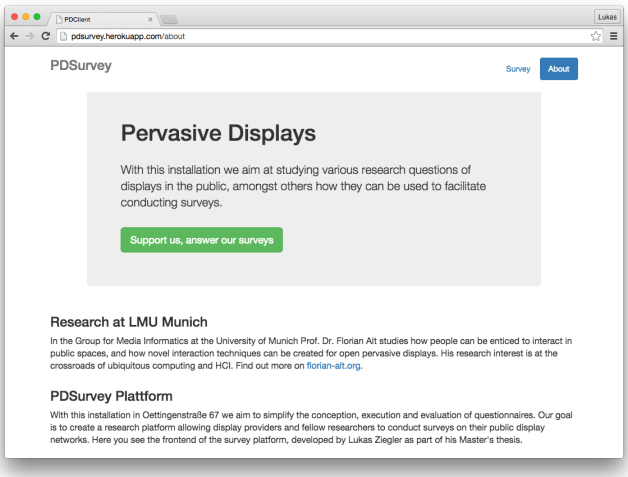
Useful links for the development of this prototype were Google Analytics concept of track-



(a) Welcome Page



(b) Survey page



(c) About page

Figure 3.3: Overview of PDClient

ing[31], Wikipedia describing the Web Bug [32], and a Stackoverflow discussion regarding embed codes [33].

For inspiration, of what the finished embed code functionality could have looked like, have a look at the Qualtrics blog article "Website Feedback Surveys"[34] and at the demo of Qualtrics Site Intercept.

3.4.5 Feedback Channels

As of now PDSurvey offers a ready built survey tool for all devices being capable of running a browser and displaying the PDClient website. Thus a large number of feedback channels is conceivable. For our scenario these were a tablet, smartphone, and laptop/desktop. Integrating PDClient on a public display itself is also no problem, as long as the public display application is a web application itself (embed code), or it supports embedding a browser window on top of the application.

In case that the application does not integrate well with a web page being displayed on top of the actual application, then a custom integration needs to be built making use of the REST calls. All REST calls needed to receive the questionnaire and send responses to the PDSurvey platform (in JSON format) can be found in the documentation (see Appendix B).

3.4.6 (Future Work)

PROBABLY PUT THIS IN THE VERY END OF THE PLATFORM; BEFORE CONCLUSION

Authentication: using passport, HTTPS was not offered in the beginning. As of now it is not needed, since we only had one client. All REST Update and Delete functionality was restricted to the URI of the PDServer, which is also the desired approach in the production setting.

Evaluation of results: dynamic queries, information visualization

Context: refine the Context model and build the evaluation of the responses based on the Context model, assigned to each and every display and/or campaign.

Unit testing: not done yet

[31] `https://developers.google.com/analytics/resources/concepts/gaConceptsTrackingOverview` (last visited on November 26, 2014)

[32] `http://en.wikipedia.org/wiki/Web_bug` (last visited on November 26, 2014)

[33] `http://stackoverflow.com/questions/3534524/how-does-the-embedded-google-analytics-javascript-work` (last visited on November 26, 2014)

[34] `http://www.qualtrics.com/site-intercept/website-feedback/` (last visited on April 6, 2015)

4 Field Study

The field study took place during the first two weeks of March, from 3/3/2015 to 3/15/2015 in Oettingenstrasse 67, a faculty building of Ludwig-Maximilians-Universität München. Data was collected from the displays on 14 consecutive days and 28 semi-structured interviews were carried out on five working days during the same two weeks. A total of 117 interactions were registered with the public display installation and 57 survey responses were recorded.

The goal of this study was to validate our research questions, and to see how users respond to questionnaires being conducted on displays in the public. We chose to conduct a descriptive study, with a focus on ecological validity, since our research prototype is still in an early development stage.

4.1 Research Questions

One of the main reasons why we performed this field study, was to get a better understanding of our assumptions and to see how users react to questionnaires on displays in public. Besides, it was of importance to conduct a study "in the wild", because there often is a discrepancy between lab studies and field studies. This phenomena has been discussed by Ojala and Kostakos in 2011: "The first important conclusion we have arrived at is that there exists a huge difference between results obtained in a lab and in the wild using the exact same configuration" [18].

An assumption we made for the development of our first research prototype of the PDSurvey platform was that we can simplify the conduction and deployment of surveys to large public display networks. Since this is a rather large claim, we broke down this hypothesis to more fine-grained statements.

We already had an application running on a public display in a faculty building which attracted lots of regular and new users. Our interest was how we could best integrate questionnaires on and after the application itself (a balloon shooter game, see section 4.2.3).

The following research questions are the basis for the field study:

1. Which channels are best suited for completing surveys in public?
2. Why did the users approach the display? What motivates them to fill in surveys in public?
3. How did the user notice and perceive the survey on the display?
4. Which question types are best suited for questionnaires carried out in public?

These four research questions were represented in the PDSurvey and semi-structured interviews. Due to the time restriction of this Master thesis could be assessed which came up during the literature review phase. Questions which would go beyond the scope of this thesis, and might serve as follow-up questions for further research, are gathered here:

- In which situations is the user most willing to answer surveys on public displays? What influence does the context and surrounding environment have on the survey responses?
- How many questions are acceptable and tolerated? Does this variable differ between the feedback channels, location of the display setup, and its surrounding environment? Which other factors play a role here?
- What is the ideal placement for surveys to pop up on a public display? Where should an overlay be positioned, when embedding it into a foreign application? How large/small, how obtrusive should it be?

- Getting a better understanding of the influence of the environment, e.g. how personal questions can get in public, and how much privacy the display should offer (the smaller the display, the more private the context seems).
- How can we best break down a standardized survey with 10+ questions across multiple users, each getting their own subset of questions?

In addition to these questions we were also interested in user stories, in the feedback real-world users gave us in regards to answering surveys on screens in the public. For this reason we also conducted semi-structured interviews in parallel to the quantitative evaluation of the PDSurvey platform. In order to get as authentic and personal feedback as possible, we stuck only roughly to the designated questions of the semi-structured interview (see Appendix E).

4.2 Study Setup

A descriptive research type was as the study type, aiming to describe and observe how users react to the new display setup. One single prototype is deployed, without varying any variables. The goal was to get first feedback on how people perceive filling in questionnaires in public, before getting into more fine-grained research (see Future Work, chapter 5). Both quantitative and qualitative data was gathered as part of the field study. Quantitative data was obtained through the PDSurvey system and qualitative data was collected through semi-structured interviews.

4.2.1 Design

Our primary goal was to find out which channel users preferred to respond to surveys in public. Each user had the choice to respond to the questionnaire on a TV display, on a tablet to the right of the TV screen, via their own smartphone or via email. We displayed the same five questions on all four feedback channels (see Table **??**). We chose to limit the number of questions asked via all four channels to 5, to avoid a low participation and response rate.

Wording	Question Type
1. How often have you used this display before?	Numeric
2. How likely is it that you will use this display in the future again?	5-point Likert scale
3. Which devices do you possess or use regularly?	Multiple choice, 5 options
4. In which area do you study / work?	Text field
5. What was your motivation for approaching and using this display?	Text field

Table 4.1: Questions asked on all four feedback channels

In order to also get first insights into how well certain question types are suited for surveys in public, where a short completion time is crucial, we varied between the following question types and kept them in the same order: numeric questions, Likert scale, multiple choice (based on check boxes) and two text fields for responses of undefined length.

To increase the motivation for participation, additional intrinsic motivation was given to increase the response and acceptance rate of the public surveys, as proven by Richard Ryan in his self-determination theory[35] [19]. We stated that the questionnaire consists of five questions, that it will only take one minute to complete and the results are for a Master's thesis at the university. This information was displayed as a splash screen on the tablet (see Figure 4.1).

Since we conducted a descriptive study, we only observed how users used our study setup. The parameter of interest was the feedback channel chosen to respond to the survey. Due to the fact that we didn't have any conditions, no independent variables are present.

[35] http://www.selfdeterminationtheory.org/

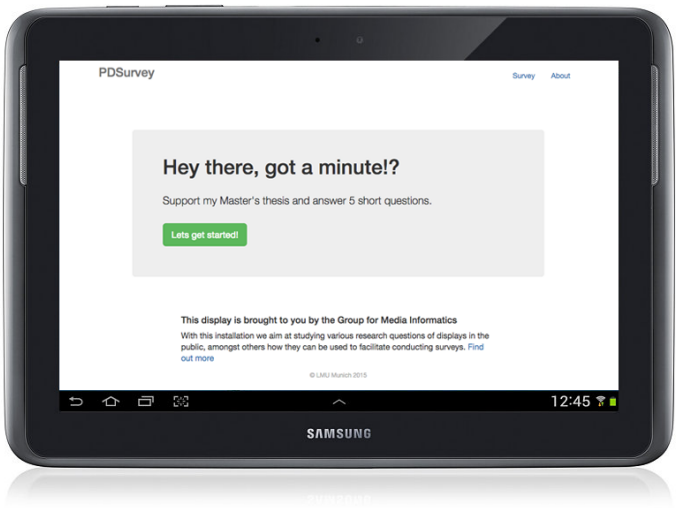


Figure 4.1: PDClient: Motivating users to participate in a short questionnaire

To find out more about the users motive for approaching the display setup, we also carried out semi-structured interviews in parallel to the field study of the PDSurvey platform. The goal of the interviews was to get qualitative feedback from all age groups and backgrounds. Getting a better understanding of how people respond to questionnaires in public, helps us develop the PDSurvey platform more target-orientated.

4.2.2 Participants

In total 54 questionnaires were submitted and 28 semi-structured interviews were conducted during the two week study period. As for the study size, we took the findings from Alt et al. [3] as a rough guide, according to which most field studys have an average of 26.9 interviews and 38.4 questionnaire responses.

Questionnaires were completed mainly by students, but with various backgrounds. Since our research focus was to find which feedback channel is best suited for conducting automated questionnaires in public, we only included one question in the quantitative PDSurvey questionnaire to collect demographic data.

Based on the fourth question ("In which area do you study / work?") we can draw a conclusion about the study field of the survey participants. Out of the 54 questionnaire responses 42 submissions could be assigned to a study field. The remaining 12 submissions consisted of responses such as *bavaria, bib, home, munich, muc* or were left empty. The study fields most frequently represented were Computer Science (23,8%), followed by Political Science (14.3%), Japanese Studies (11.9%), and Anthropology (11.9%). Other study fields mentioned were Cultural Science, Business, Physics, Sociology, Ethnology, Communication Science, Sports, and Science & Technology.

For the semi-structured interview we collected more detailed information about demographics of the participants. Out of the 28 participants, 20 were male and 18 were female. The age of all people interviewed ranged between 20 and 69, with an average age of 31,5 years. Due to the wide variety of faculties and a library being located in the same building, various technical backgrounds were present. What they all had in common was their affiliation to LMU Munich, either because of being a student themselves, working at the university or being otherwise related to the university. In total 23 students, three employees, and two retirees were interviewed. The study fields which were most frequently represented are Computer Science (16.7%), Japanese

Studies (16.7%), Ethnology (12.5%), and Political Science (12.5%). Other areas mentioned were Sociology, Communication Science, Law, Physics and Engineering. The full list of demographics, including exact ratios, can be seen in Table 4.2.

	people participated in survey		**people interviewed**
10	Informatics	4	Informatics
6	Political Science	4	Japanology
5	Anthropology	3	Ethnology
4	Cultural Science	3	Political Science
4	Business	1	Communication Science
2	Physics	1	Sociology
2	Sociology	1	Law
1	Ethnology	1	Physics
1	Communication Science	1	Engineering
1	Sports	3	workers (PhD, public officer, SysAdmin)
1	Science and Technology	2	in pension

Table 4.2: Demography for the survey data (left) and the semi-structured interview (right).

The selection of the participants for the completion of the surveys was not influenced by us. All survey responses were made in their own interest, no reward was given for participating in this "in the wild"-study. The selection of the participants for the semi-structured interview was influenced by how users reacted to the display setup. Our primary goal was to observe and interview active users of the public display setup, in order to get a better understanding of how they perceived the setup and to get insights into which feedback channel they chose why. In order to also understand why people did not approach, or if they have any concerns, people passing by were also interviewed.

Before starting the semi-structured interviews, all people participating were asked whether they have previously noticed the display setup and/or the option to fill in a survey.

This allowed us to differentiate three groups: participants who approached the display by themselves (and were observed doing so), people passing by the display (noticing the display, however not approaching it) and the last group of people simply passing by (not having noticed the display). The distribution between the groups was as follows: 11 active participants, 17 passerby (14 noticed the displays before, 3 haven't).

Out of all people passing by no one has noticed the option to fill in a survey. Out of the active participants, 4 out of 11 have noticed the option to respond to a survey on different channels.

To increase the amount of feedback, we approached people from all three groups. The number of survey responses was not artificially increased by asking passersby was.

> TODO: think about which information is RELEVANT and which is not! for me the devices they possess might be of relevance. people not possessing a smartphone or tablet, might be less willing to use these devices.

... and 25 out of 28 noticed the public display setup before. For three people interviewed it was the first time that they have noticed the displays.

4.2.3 Apparatus

The permanent setup consisted of a XXX-inch TV screen, connected to a laptop running Windows 7, and a Samsung Galaxy Tab 10.1 tablet positioned to the right of the TV screen on a console. Our object of investigation was the TV screen with touch support, running an interactive game called *Balloon Shooter*, developed by Jiamin Shi. After users completed the game, they were asked via a notice to fill in a questionnaire on one of the four provided feedback channels (see Figure 4.4).

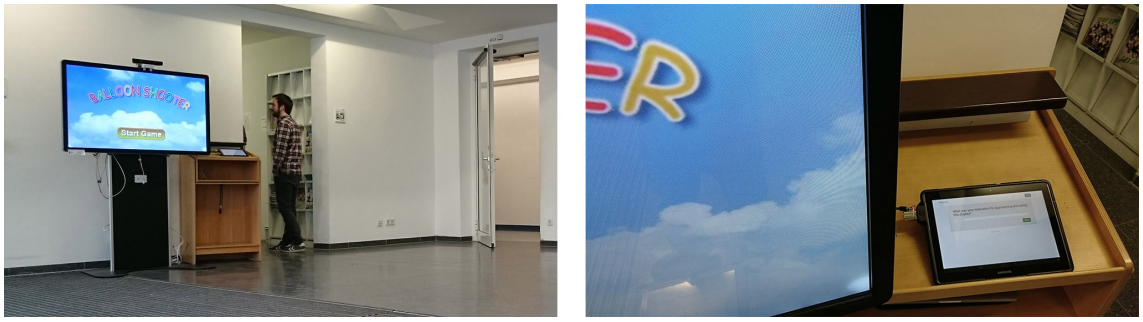


Figure 4.2: Overview of the study setup in the entrance hall of the faculty building.

The courtesy for the Balloon Shooter game and the survey implementation on the TV screen goes to Jiamin Shi.

Each user had the opportunity to respond to the questionnaire either directly on TV screen (1), on a tablet to the right of the big TV (2), via their own smartphone (3) or via email (4). The first option was embedded natively into the Balloon Shooter game, offering a consistent UI and the most direct feedback channel. Choosing the tablet as an option, the users were prompted to move to the right and to answer five questions on the tablet. The Samsung Galaxy Tab 10.1 was displaying the responsive frontend of PDClient, being enclosed in an Android Kiosk App, namely KioWare Lite[36]. Choosing the third option prompted the user to either scan a QR code with their smartphone or to open a URL in their mobile browser. The last option consisted of an input field embedded into the Balloon Shooter game on the TV screen, asking the user to enter their email address. The address was logged to a txt-file, which was scanned every 5 minutes by a Windows task scheduler. An email reminder was sent to the user with the request to complete the survey. For sending the email from the TV screen a Python script was written to send the email via the universities SMTP server[37]. Screenshots of all four options can be found in the Appendix on page 35.

For the permanent setup the following data was logged on all four feedback channels: The timestamp of the users choice, which feedback channel the user chose to respond to the survey, and whether they skipped the call to participate in the survey or if they stopped playing the game (determined via timeout).

For conducting the semi-structured interviews two questionnaires (one for participants, one for passerby), a voice-recorder (smartphone) were used in addition to the permanent setup.

For the evaluation in the field study itself a self-made questionnaire was used, since the focus was on finding which channels and question types are best suited in general for being used on public display. This was the reason why we did not use any of the standardized questionnaires mentioned in section 2.2. Screenshots of the questionnaire run on the PDClient can be found on the enclosed CD.

The main application installed on the public display was a game called *Balloon Shooter* developed and run by Jiamin Shi, a PhD student at the Group for Media Informatics at LMU Munich. It was first installed on January 7th 2015 and has been running in different versions since then. Public audience was already used to it for roughly two months and adapted to it well.

4.2.4 Location

All parts of the field study were carried out in Oettingenstrasse 67, the faculty building for Computer Science. In the same building there are also research institutes for Ethnology, Political

[36] http://www.kioware.com/android.aspx
[37] https://github.com/lukasziegler/python-send-mail

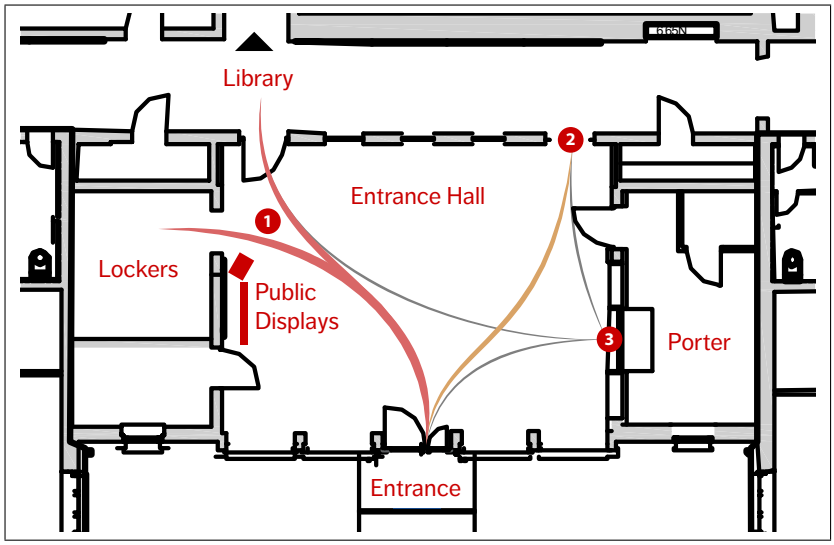


Figure 4.3: Floor map of the entrance hall, where the field study was carried out. User paths, and the surrounding environment including facilities such as the library can be seen.

Science, Japanese Studies, and Physics. The study was carried out in the entrance hall of the university building. Figure 4.3 gives an overview of the entrance hall and of the paths most people take while crossing the room. The excerpt is based on the universities floor plan[38] and was inspired by Sandra Zollner [1]. There she also published that at the time of her study "approximately 59% of all passers-by used path 1", to get something from the lockers or to leave through the door to the library. 28% of the people were taking path 2 and 13% were taking path 3.

In our field study it was also evident that the majority of the visitors took path 1 were usually fairly target-orientated or in a hurry. Otherwise, on days with bad weather people had their break in the entrance hall or waited for someone. On days with good weather people usually took their breaks outside and only passed through the entrance hall, coming from the library, picking up something from the locker room and going outside.

4.2.5 Procedure

All participants of the semi-structured interview were asked a similar set of questions (see Appendix E). Based on the group they belonged to, either questionnaire 1 (for participants of the display setup) or questionnaire 2 (for passersby) was chosen. In order to speed up the interviewing process and to get away from a plain question-response schema, the questions on the printed out questionnaire only survey as a rough guideline.

For people having trouble understanding the concept of the public display installation, the situation was described as follows. "Imagine you are in a shopping mall or at an airport using one of those large displays to find some information. After having found what you were looking for, you get asked to answer a short questionnaire. How would you react to it?" A full transcription of all questions and responses can also be found on the enclosed CD.

The participants for the PDSurvey questionnaire were not additionally motivated. All they saw was the options panel after completing the Balloon Shooter game (see Figure 4.4) or the welcome screen of the tablet (see Figure 4.1) while passing through the entrance hall. A complete copy of what the users were able to interact with, can be seen on the attached CD (see Appendix A).

[38] `http://www.uni-muenchen.de/funktionen/gebaeudeplaene/7070_d_00.pdf` (last visited on March 22, 2015)

Figure 4.4: Options panel to choose a feedback channel

4.3 Results

We received a total of 57 filled in surveys, submitted via all four of the provided feedback channels, and carried out 28 semi-structured interviews. No treatments were applied to the dataset, descriptive statistics will follow bellow. The presentation of the evaluation is divided into three parts. First we will have a look at which feedback channel is most popular, followed by the quantitative results of the PDSurvey questionnaire, and rounded off with the results from the semi-structured interview.

4.3.1 Feedback Channels

NOT SURE HOW TO BEST POSITION THIS PART

The majority of the surveys were submitted on the tablet (87.72%). Four responses were recorded directly on the TV (7.02%), two via smartphone (3.51%), and one via email (1.75%). Since this listing only contains the number of responses, it should not be taken as a base for the comparison of the feedback channels' popularity. Due to the tablets sole purpose to be used to fill in surveys in our setup, and the additional intrinsic motivation given on this channel (see section 4.2.1), this ratio has to be treated with caution. For a comparison of the **feedback channels** the log data from the option panel and the responses from the semi-structured interviews are better suited (see Table 4.3).

After having completed one session of the game, users had the option to choose a feedback channel. Based on this log data of the TV screen a better comparison of the feedback channels can be made, since all direct responses made on the tablet are excluded from this summary. The most popular feedback channel was the tablet (46.15%), followed by the TV screen (30.77%), smartphone (15.38%), and email (7.69%). In order to have another source of input, the same question was asked at the end of every semi-structured interview. Based on this quantitative data from the interviews the response via tablet (42.86%) was most popular again, followed by the TV screen (32.14%). Interestingly, for the interview data the option to respond via email (17.86%) is more popular than smartphone (7.14%).

4.3.2 Survey Responses

How often have you used this display before? In our questionnaire executed on the public displays we asked five questions. Out of the 57 responses people have on average used the display

From Survey Data		From Interviews
30.8%	on public display	32.1%
46.9%	on tablet	42.9%
15.4%	on smartphone	7.1%
7.7%	at home / via email	17.9

Table 4.3: Preferred feedback channel for answering surveys.

6.9 times before. For 25 people (43.9%) of the users it was the first time using the display setup, 11 people (19.3%) have used it once before, 18 people (31.6%) between two and ten times, and the remaining 3 people (5.3%) more than ten times.

How likely is it that you will use this display in the future again? For the second question, based on a 5-point Likert scale, the response was fairly uniformly distributed (average=3.04, SD=1.46). The whole scale from 1 (not likely at all) to 5 (very likely) was represented. No clear trend could be seen. When only considering the responses collected from the large TV screen, a better perception can be noticed. There the responses to this question had an average of 4.5 (SD=0.866), showing a trend towards a positive perception of the large display setup. However, due to the low number of responses for the TV display, this conclusion can not be regraded as significant.

Which devices do you possess or use regularly? Taking a look at the devices users possess might give us first insights into why users chose which feedback channel. Overall, the majority of the people participating in the survey already owned a smartphone (79.3%). The second most popular response was laptop (73.6%), followed by tablet (41.5%), and desktop computer (26.4%). Still, 18.9% of the users indicated that they possess a feature phone and use it regularly. On average each participant possessed 2.4 devices. When looking at which combinations of devices were most frequent, twelve people responded that they own a smartphone + tablet + laptop. Twelve other people indicated that they possess a smartphone + laptop, and six people own a smartphone + laptop + desktop.

In which area do you study / work? The fourth question was used to get a little insight into the background of the survey users. Only the occupation of each participant can be derived from the questionnaire. As far as was indicated all people responding to the questionnaire installed on the public display setup were students. The majority of people interacting with the TV screen study Coputer Science (23.8%), followed by Political Science (14.3%), Japanese Studies (11.9%), Anthropology (11.9%), Cultural Science (9.5%), and Business (9.5%). Table 4.2 shows a full list of which study field or work field the participants specified.

What was your motivation for approaching and using this display? The main reasons why people have approached the display were “curiosity” (12x), “fun” (10x), “boredom” (8x), “interest” (2x), and “during breaks” (2x). Other reasons mentioned were “it is there, so why not?”, “it is there and colourful”, or “I’ve never seen it before in this spot, wanted to know what it is about”.

4.3.3 Interview Responses

As mentioned earlier, we also conducted semi-structured interviews. The evaluation of the semi-structured interviews was based on Grounded Theory, for a systematic evaluation of the interview transcripts. A total of 28 semi-structured interviews were conducted, of which 72.4% of the participants were male and 28.6% were female. The average age was 31.5 years, with an age distribution

ranging from 20 years up to 69 years (median=25.5, SD=13.2). Eleven of the 28 interviews were conducted with actual participants of the public display study setup (39.3%), the remaining 17 interviews (60.7%) consisted of people passing by the display.

To avoid any interferences between the two groups, each passerby was asked before starting the interview whether he has noticed the public display installation, and whether he has already interacted with the installation. Out of all passerby no one has previously been interacting with the game or survey platform. 82.4% (14 of 17) of the passerby have already noticed the public display installation before, however, none of the passerby has previously participated in the game. The remaining 17.6% have neither approached the display nor noticed it previous to the interview.

Looking at the scientific background of all 28 participants, 79.2% are students, the remaining 20.8% either already worked full-time or were in pension. The majority of students studied Computer Science (16.7%), Japanese Studies (16.7%), Ethnology (12.5%) or Political Science (12.5%).

From what has been mentioned, the main reason for approaching the public displays was in 6 out of 8 cases "curiosity" (6x). Two other reasons were "for fun" (1x) and "waiting for someone" (1x). Reasons for not approaching the display were "no time" (2x) and "it is in the entry zone of the university, it feels strange when one plays with it" (1x).

From the open coding phase the following patterns can be seen:

ASK JULIE!

- reason for approaching: see above - number of questions found acceptable: 5 - 10 - reasons for choice of feedback channel: - requirements for a survey, to attract users:

- most interesting feedback: reason PRO / CON using a certain channel

The semi-structured interview was most useful to get a better insight into why certain users chose which feedback channel. Reasons mentioned influencing their choice were ""

see Table 4.4.

4.3.4 Additional Observations

The response time for responding to the five questions was on average 1:02 minutes, ranging from 0:36 to 3:06 minutes.

How many questionnaires were fully completed, how many were aborted? Which questions were left empty? Does this infer anything for the quantitative vs. qualitative surveys?

Questionnaires on public displays are best suited for quantitative surveys. Users want a short interaction time, not having to think much about their answers and for roughly XXXXX percent of the participants it holds true, that they do not like being observed while making responses in public. From this observation, the implication for the **question types** can be derived: question types ideally with a single-click interaction are preferred (e.g. Likert scale, multiple choice with all options given, yes/no-questions). Then followed by numeric, dropdown and multiple choice questions with one option for open-end responses. For these question types the user has to think a little bit more, he has to assess more precisely in order to make his response. One example stated by a participant, in regards to the numeric question 'How often have you used this display before?', was that "It would be great if you had the possibility to choose from a predefined range, because typing is not always optimal. I would prefer if areas would be given instead of oneself having to think about the exact number." Last, being no big surprise, are text fields combined with open-ended questions. As a take away for text fields: wherever possible rephrase the question so that you can respond to it as short as possible.

4.4 Discussion

It is interesting to see that the tablet is the most popular **feedback channel** in all scenarios, although responding via the TV screen would be more a more direct approach and not require

moving to another device. Nevertheless all offered feedback channels were present in the evaluation and during the semi-structured interviews for each channel a good reason was given. What can be said that the crowd usually distinguishes into three groups. The first (and slightly larger) group preferring the option of *direct response*. They are not as considerate about answering questions in public and their privacy aspect. For them it is more important to complete the survey as quick as possible and not having to think about it later, as long as nothing too private or personal gets asked. One person said "If something too private would be asked, I would simply abort and go away from the display". The second group is more *privacy* concerned, often older of age, or actually wanting to take the time to think about all of their responses in depth in order to give high-quality responses. This group prefers to take the questionnaire away from the public setting into their home. The third group chose the feedback channel purely based on their *habit* and what they are accustomed to. Two ladies in their mid-twenties responded immediately "on my smartphone, because I am most used to it".

These observations go along well with the five adaptation factors stated by Huang et al. [11]: task specificity and deep integration, tool flexibility and generality, visibility and exposure to others' interaction, low barriers to use, dedicated core group of users.

Another assumption we had was confirmed by our observations and the semi-structured interviews: the smaller the display, the safer and more private the users feel. An exception to this finding could be old people. Once people become short-sighted or more insecure and unconfident with using new devices, they prefer having a large input surface. But for the majority of younger people this held true.

Additionally we made the observation that users responded ...

We are aware of certain limitations of our descriptive study. Our limitations are consistent with the findings found by Ojala et al. [18]. The effects of curiosity, impact of novelty, and influence of weather had an influence on our field study. Due to the novelty effect caused by the tablet, and the intrinsic motivation we added through the splash screen on the tablet (see section 4.2.1, selt-determination theory), the participation rate on the tablet was increased. For our primary research question, which feedback channels is best suited, the impact of novelty, curiosity and of the always-visible tablet, should not have an impact. We based the evaluation of the feedback channel not on the overall number of responses, which was therefor distorted, but on the option panel and on the interview responses.

Despite these effects, it was striking to see a response rate of 42.4%, when comparing the 50 responses made on the tablet with the total number of 118 interactions made with the public display setup. When we exclude all participants who directly accessed the tablet and did not see the option panel to use one of the four feedback channels, the response rate on the tablet was still 5.1%.

Otherwise, it should be mentioned that both the TV screen and the tablet were always on and that all questions were optional. One suggestion for improvement is to only turn on the screen of the tablet when it is selected on the TV screen as the desired feedback channel.

All in all, it can be said that people prefer to respond to questionnaires in public directly, as long as the questions don't get too private. Nonetheless the more feedback channels one offers, the better it is, since the variety of user backgrounds also bring different preferences and attitudes. When designing public display setups for getting more sensible user input, the display size should also be taken into consideration. So far we have made the observation, that users feel more secure on smaller screens.

For the development of our public display survey platform the study showed that we are on a good path. - people are willing to respond in public - put a higher priority on how users are motivated to participate, currently not embedded yet for scalable solutions, currently it was manually embedded - a higher focus on fast, one-click responses

TV Screen	
Pro	**Contra**
	Display is too large (4x)
Most direct, immediate feedback (4x)	Everyone could watch me (2x)
I am already standing here (2x)	Feels too public (3x)
Seems easiest	That is mean, when the screen is so large
Requires less personal information	The keyboard on the display would have been too large and confusing
All on one device	Display is uncomfortable for reading long questions
I can use it without putting my glasses on	Don't feel comfortable standing in focus in such a large room
Seems to be the fastest option	The system is too innovative, that is why I wouldn't trust it yet
	Because of social desirability influencing my responses
Tablet	
Pro	**Con**
The display is smaller and better laid out (5x)	
Better sensitivity / user experience (2x)	
It feels more private (2x)	
Because it is its sole purpose	
You are not in the way of others	
I am more used to it	Redundancy (2x), why do I need a tablet when I can respond on the T
Most interactive option	Personal aversion (1x), he had bad experiences with tablets
Less people watching	
Because I expect a better input	
Requires less personal information	
More comfortable standing here	
Smartphone	
Pro	**Contra**
	Too much effort (4x)
	Too indirect (3x)
	Requires more personal information (3x)
	I am not sure how complex and time-consuming it would be to set it u
I use it most often	I don't know if I would know how to do it
It belongs to me	Too small display for comfortably answering surveys and long questi
	Too cumbersome
	I would assume that I would have to install some sort of software
	Privacy aspect
Email	
Pro	**Contra**
	I would forget about it (5x)
I can do it at home (4x)	I don't like to submit my email address (4x)
I have more time to complete the survey (3x)	I don't like to postpone it (3x)
Better warranty of privacy	It would take too long to complete (2)
I could deliver qualitatively better results	Too much effort (2x)
I wasn't sure which kind of questions to expect	Requires more personal information
	Too Indirect

Table 4.4: Reasons stated for or again each feedack channel

5 Future Work

Survey Platform

1. Logging, adding more data sources (for tracking of "User performance", see chapter 2.2)
2. Support the logging of video feeds. One possible approach would be to save the raw data in a dropbox account, to submit the file names via REST calls, and to access the files from PDSurvey via the Dropbox API
3. Dynamically evaluating *reliability* and *validity* in the platform, ideally on-the-fly.
4. also for pervasive displays?

Evaluation

1. Number of questions tollerated on each evaluation channel (public display, tablet, smartphone, laptop/desktop)
2. + also see chapter 5, subsection "Research Questions"
3. to also impose variables: context, content, location / setting and to vary those in experimental studies
4. splitting
5. finding the similarities between different standardized questionnaires. As Jacucci et al. [12] mentioned, there are often significant similarities between standardized questionnaires. It might be possible to break down each questionnaire to its principal parts, to bundle all conformities, in order to reduce the total amount of questions and to be able to split all questions across multiple users on the same display.

Other

1. How should a survey be constructed to take best advantage of the PDSurvye platform? How many quantitative and how many qualitative questions?

6 Conclusion

Outlook, be inspired by `http://www.wlu.ca/forms/1676/Conclusion.pdf`

1. which feedback channel:
2. which question types:
3. quantitative vs qualitative:

Appendix

A Content of enclosed CD

1. /docs/
2. /pdsurvey/
3. /pdemail/
4. /pdclient-static/
5. Google Docs
6. transcribed interviews!

B Documentation of PDSurvey Platform

A user, developer and maintenance documentation for the PDSurvey platform can be found in the GitHub repository [39] and on the enclosed CD.

C Papers Evaluating Public Displays

List of relevant papers, which include an evaluation of public displays.
TABLE: 1st column (paper), 2nd column evaluation (quantitative, qualitative, no evaluation)
TODO

D Questionnaires for Field Study

E Questionnaires for Field Study

Embed the following PDFs

1. interview-participants.pdf
2. interview-passerby.pdf
3. semi-structured-interview.pdf

F Screenshots of Platform

All screenshots including the copyright of the Balloon Shooter game belong to Jiamin Shi. // with courtesy to JIAMIN SHI.
1st option: tv screen
2nd option: tablet
3rd option: smartphone
4th option: email

[39] https://github.com/lukasziegler/masterarbeit/tree/master/docs

Figure F.1: All four four options for completing a survey, the order being randomized on every instance.

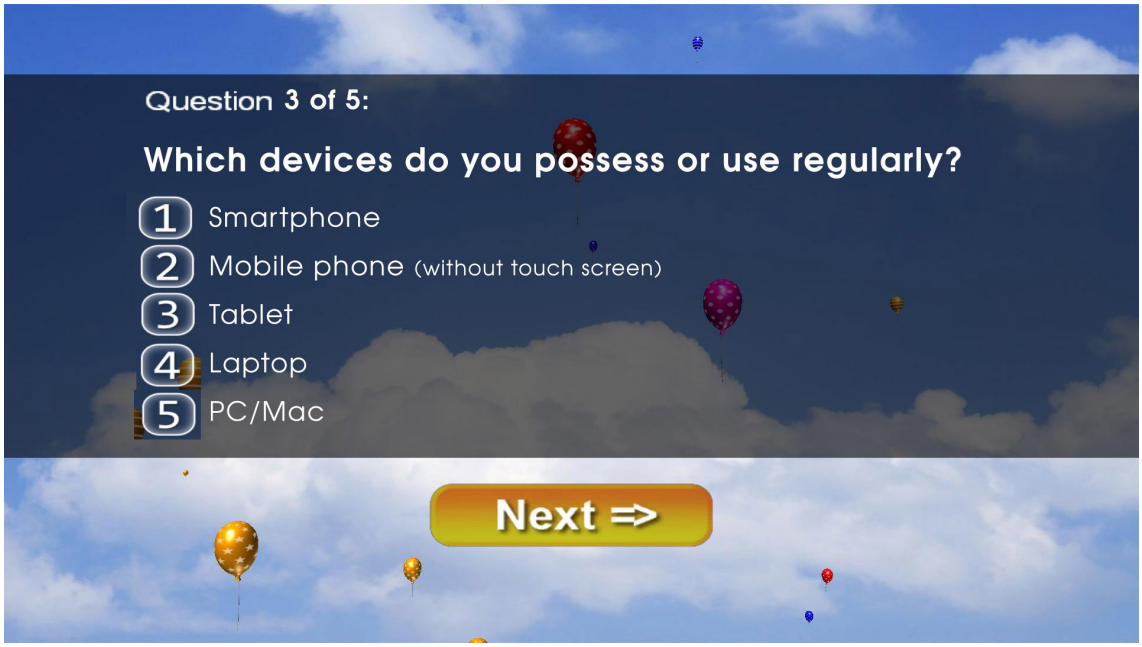


Figure F.2: Option 1, directly answering on the TV screen. Here you see a sample question getting asked on the interactive display.

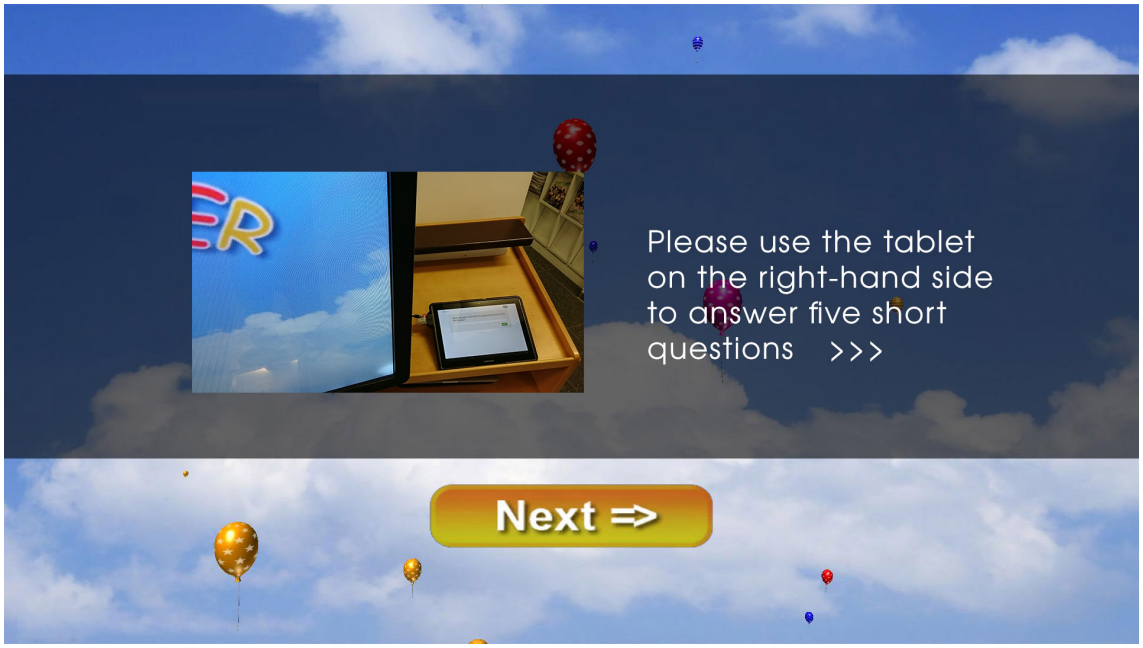


Figure F.3: Option 2, the screen the user sees when choosing to complete the survey on the tablet.

Figure F.4: Option 3, participating with your own smartphone, either by scanning the QR code or by typing the URL in the mobile browser.

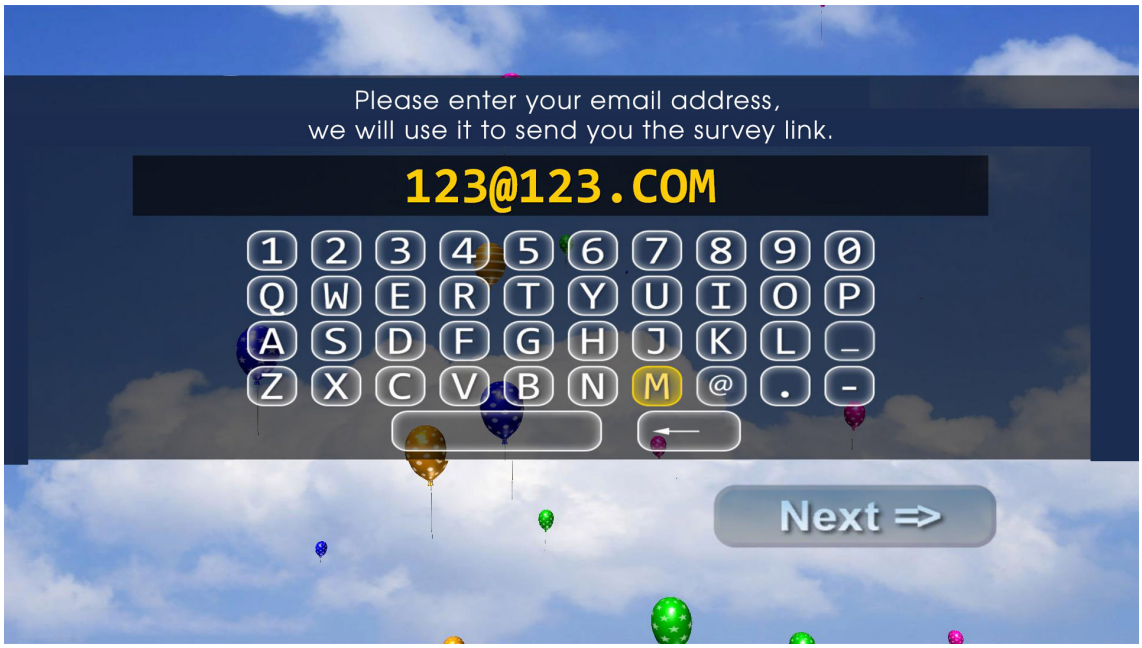


Figure F.5: Option 4, submitting ones email address and getting the survey link to participate in response.

Bibliography

References

[1] Bachelor's thesis.

[2] Florian Alt. LULU Press.

[3] Florian Alt, Stefan Schneegass, Albrecht Schmidt, Jörg Müller, and Nemanja Memarovic. How to evaluate public displays. In *2012 International Symposium on Pervasive Displays (PerDis'12)*, page 17, 2012.

[4] FlorianT Alt, Thomas Kubitza, Dominik Bial, Firas Zaidan, Markus Ortel, Björn Zurmaar, Tim Lewen, Alireza Sahami Shirazi, and Albrecht Schmidt. Digifieds: Insights into Deploying Digital Public Notice Areas in the Wild. *Adjunct Proceedings of the 9th International Conference on Pervasive Computing*, pages 165–174, 2011.

[5] K. Cheverst, Alan Dix, D. Fitton, C. Kray, M. Rouncefield, Corina Sas, J. Sheridan, and G. Saslis-Lagoudakis. Exploring Bluetooth based Mobile Phone Interaction with the Hermes Photo Display. pages 47–54, 2005.

[6] HTW Chur. Fragebögen. http://www.cheval-lab.ch/was-ist-usability/usabilitymethoden/frageboegen, 2014. Accessed April 6, 2015.

[7] Université de Genève. Usability and user experience surveys. http://edutechwiki.unige.ch/en/Usability_and_user_experience_surveys, 2014. Accessed April 6, 2015.

[8] A.P. Field and G. Hole. *How to Design and Report Experiments*. Sage publications Limited, 2003.

[9] Adrian Garcia. UX Research | Standardized Usability Questionnaires. http://chaione.com/ux-research-standardizing-usability-questionnaires/, 2013. Accessed April 6, 2015.

[10] Elaine M Huang, Anna Koster, and Jan Borchers. Overcoming assumptions and uncovering practices: When does the public really look at public displays? In *Pervasive Computing*, pages 228–243. Springer, 2008.

[11] Elaine M Huang, Daniel M Russell, and Alison E Sue. IM Here: Public Instant Messaging on Large, Shared Displays for Workgroup Interactions. 6(1):279–286, 2004.

[12] Giulio Jacucci, Ann Morrison, Gabriela Richard, Jari Kleimola, Peter Peltonen, and Toni Laitinen. Worlds of Information: Designing for Engagement at a Public Multi-touch Display. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2267–2276. ACM Press, 2010.

[13] Jurek Kirakowski. Questionnaires in usability engineering, 2000.

[14] James Lewis and Jeff Sauro. Standardized Usability Questionnaires. http://2013.hci.international/index.php?module=pagesmith&uop=view_page&id=44, 2013. Accessed April 6, 2015.

[15] Jörg Müller, Florian Alt, Daniel Michelis, and Albrecht Schmidt. Requirements and design space for interactive public displays. In *Proceedings of the international conference on Multimedia*, number Figure 1, pages 1285 – 1294. ACM Press, 2010.

[16] Jörg Müller, Gilles Bailly, Thor Bossuyt, and Niklas Hillgren. MirrorTouch: Combining Touch and Mid-air Gestures for Public Displays. In *Mobile HCI '14*. ACM Press, 2014.

[17] Jörg Müller, Robert Walter, Gilles Bailly, Michael Nischt, and Florian Alt. Looking glass: a field study on noticing interactivity of a shop window. *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI '12*, page 297, 2012.

[18] Timo Ojala and Vassilis Kostakos. It's a Jungle Out There: Fantasy and Reality of Evaluating Public Displays in the wild. In *CHI 2011*, pages 1–4. ACM Press, 2011.

[19] Richard M Ryan and Edward L Deci. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist*, 55(1):68, 2000.

[20] G. Vaish. *Getting Started with Nosql*. Packt Publishing, 2013.

[21] Nina Valkanova, Robert Walter, Andrew Vande Moere, and Jörg Müller. MyPosition: Sparking Civic Discourse by a Public Interactive Poll Visualization. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*, pages 1323–1332. ACM Press, 2014.

[22] Ina Wechsung and Anja B Naumann. Evaluation methods for multimodal systems: A comparison of standardized usability questionnaires. In *Perception in Multimodal Dialogue Systems*, pages 276–284. Springer, 2008.

Web References

References

[1] J. Nielsen: Alertbox: Current Issues in Web Usability http://useit.com/alertbox/, accessed April 24, 2005.