Evolving a research tool meant for parents of NICU infants through Interaction Design

University of Turku
Department of Future Technologies

Master's Thesis
Computer Science
October 2019
Markus Willman

The originality of this thesis has been checked in accordance with the University of Turku quality assurance system using the Turnitin OriginalityCheck service.

UNIVERSITY OF TURKU
Department of Future Technologies / Faculty of Science and Engineering

MARKUS WILLMAN: Evolving a research tool meant for parents of NICU infants through Interaction Design

Master's Thesis, 74p.
Computer Science
October 2019

The purpose of the research in this master's thesis is to present and evaluate the evolution of a digital research tool, HAPPY, through the frameworks of interaction design and the three-cycle view of design science research while using participatory design (co-design) as the practical methodology for evolving the tool itself. More specifically, HAPPY is an Android smartphone application intended for use by parents with infants in the NICU during different kinds of research projects with interest in the subjective or qualitative experiences of participants.

The thesis approaches the subject using two stage literature review and the comparative evaluation of the produced artifact. The evaluation is done against a prior version of HAPPY, existing similar tools or applications and other prior research, in the theoretical or methodological areas that the now evolved tool touches on. The first part of the literature review process is concerned in finding ways, through interaction design, to evaluate the different versions of HAPPY using and interpreting three previously established models. The second part compares the artifact (the updated version), created using co-design as part of an iteration of the design cycle of the previous three cycles, with existing similar solutions and the previous body of knowledge.

In between the two parts of the literature review the co-design phase and the practical work is discussed, with focus on the subjective experiences, of the family that agreed to participate in this project, while using the new version of the tool through testing as well as the produced design and the rationale behind it.

Finally, since the work presented here is, in part, exploratory in nature, it is discussed whether and how much of the findings produced here could potentially be generalized. Specifically, to what extent this is needed, what kind of further research is considered necessary and how the tool could be used as it currently exists considering the nature of the work presented here.

Keywords: interaction design, participatory design, user centered design, experience sampling method, diary research, distributed systems

TURUN YLIOPISTO
Tulevaisuudenteknologioiden laitos / Luonnontieteiden ja tekniikan tiedekunta

MARKUS WILLMAN: Keskoslasten vanhemmille tarkoitetun tutkimustyökalun kehittäminen vuorovaikutusmuotoilun kautta

Pro gradu -tutkielma, 74s.
Tietojenkäsittelytieteet
lokakuu 2019

Tämän pro gradu -tutkielman tarkoituksena on esitellä ja arvioida digitaalisen tutkimusvälineen, HAPPY, kehitystä käyttäen vuorovaikutusmuotoilun ja suunnittelutieteen tutkimuksen, kolmen syklin näkemys, viitekehystä sekä käytännön metodologiana osallistavaa suunnittelua (co-design) itse työkalun kehittämiseen. Tarkemmin HAPPY on sovellus Android-älypuhelimille, joka on tarkoitettu eritoten keskoslasten vanhempien käyttöön erilaisten tutkimusprojektien yhteydessä, kun mielenkiinnon kohteena on osallistujien subjektiiviset tai kvalitatiiviset kokemukset.

Tutkielma lähestyy aihetta käyttäen kaksivaiheista kirjallisuuskatsausta ja toteutetun tuotoksen vertailevaa arviointia. Arviointi on tehty HAPPY:n aiemman version, olemassa olevien samankaltaisten työkalujen tai sovellusten ja aiemman muun tutkimuksen, joiden teoreettista tai metodologista alueita nyt kehittynyt työkalu koskettaa, välillä. Kirjallisuuskatsauksen ensimmäiseen osaan liittyy tapojen löytäminen, vuorovaikutusmuotoilun kautta, HAPPY:n eri versioiden arvioimiseksi käyttäen ja tulkiten kolmea jo olemassa olevaa mallia. Toinen osa vertaa päivitettyä versiota, joka toteutettiin käyttäen osallistavaa suunnittelua osana yhtä edellisen kolmen syklin mallin suunnittelusykliä, olemassa oleviin vastaaviin ratkaisuihin ja aikaisempaan teoriapohjaan.

Kirjallisuuskatsauksen kahden osan välissä käsitellään erityisesti osallistavan suunnittelun vaihetta ja tähän liittyvää käytännön työtä, keskittyen subjektiivisiin kokemuksiin perheeltä, joka suostui osallistumaan tähän projektiin, työkalun uutta versiota käytettäessä sekä tuotettuun muotoiluun (design) ja sen taustalla olevaan ajatteluun ja perusteisiin.

Lopuksi, koska tässä esitelty työ on luonteeltaan osittain tutkiskelevaa, käsitellään sitä, miten paljon ja mitä tässä tuotetuista tuloksista on mahdollista yleistää. Tarkemmin, missä määrin tätä tarvitaan, millainen mahdollinen jatkotutkimus koetaan tarpeelliseksi ja miten nyt toteutettua työkalua voitaisiin mahdollisesti käyttää sellaisenaan huomioiden tässä esitetyn työn luonne.

Avainsanat: vuorovaikutusmuotoilu, osallistava suunnittelu, käyttäjäkeskeinen suunnittelu, ESM-metodologia, päiväkirja tutkimus, hajautetut järjestelmät

CONTENTS

1. Introduction

This thesis is a report on an empirical, and in some respects exploratory, development project to further develop a data collection tool for qualitative research that had been used in studies focused on the experiences and perceptions around the care of infants in the NICU (Neonatal Intensive Care Unit) environment. The tool in question, a mobile application named HAPPY for the Android operating system for smartphones and mobile devices, is based on the idea of collecting short narratives about these experiences of the participants, recorded either as audio or text, throughout their day.

1.1 Background

The two preceding studies, a pilot study and a feasibility study, involving the first versions of the HAPPY application were both conducted in NICUs with the nurses working there using the application to record thoughts about moments of closeness and separation between the parents or nurses and the infant being cared for. The latter feasibility study also included entries provided by the parents of the infants themselves using the application. In both studies, the data was collected over a relatively short period, i.e. one day (24 hours) or the duration of a single work shift (Niela-Vilén, Feeley et al. 2017, Feeley, Genest et al. 2016).

Broadly, in terms similarities in methodology and sharing many similar benefits[1], the type of data collection the application is facilitating is perhaps closest to the experience/event sampling method (ESM), also known as ecological momentary assessment (EMA), used as a measurement tool e.g. in psychology and other social sciences interested in the subjective experiences of individuals (see e.g. Hektner, Schmidt et al. 2011). Where HAPPY significantly deviates from this method, as presented by Hektner et al., is that for both of the studies

[1] For example, minimal recall bias can be presumed in most cases, based on how participants were instructed to use the application in the two studies.

that used the application no equivalent of a signaling mechanism, e.g. a notification or message to prompt the participants to engage with the software, was used. Thus, the filling of the application with experiences was entirely up to the participants themselves. This naturally puts significantly more importance on participant motivation and the perceived value that the participant gets from creating entries in the application for themselves.

1.2 Motivations and research questions

The idea for the further development of the application began by asking the question of whether it could be something more than a simple data collection tool for its users or rather could and should it have more than one primary use case and user group. With regard to the original version of the application, even though the researchers are not the ones directly using the application, it still serves as a tool primarily for the researchers over the users using it. The original design of HAPPY was created to facilitate the specific needs of those researchers without involving the actual end users which for the purposes of this new development was now defined as the parents and families in the NICU.

That initial question was then refined to the following two primary questions: Could we provide more immediate or direct benefit, other than the research results, to the participants? What kind of formats in addition to text and audio could the application use and which of these would be interesting for and help the parents in expressing their thoughts and feelings about the time in the NICU? Other questions surrounding these two were of course also those of privacy and how or when is the application used, particularly when it may involve recording audio and later other formats, in places with public or shared areas and limited opportunities for privacy such as a hospital.

1.3 Methods

For its theoretical and methodological basis, this research project uses interaction design (IxD) and software engineering together with elements borrowed from design science research (Hevner 2007) where the creation of design artifacts, in this case the next versions

of the application, is a common method to evaluate design decisions before testing them more broadly in practice.

To help answer the previous research questions co-operative, or participatory, design was used during the development phase of the next version in which new features or changes were conceived and then introduced to be tested. The parents of an infant who had recently been released from an NICU agreed to participate in this design process based on their experiences during the time in the NICU. For the purposes of this project, from the point of view of the design process, the parents can be considered as the representatives of the end users.

For the participatory design phase, the methods used were free-form interviews or discussions together with a testing period for a version of the application based on the ideas conceived during the initial interview. Anna Axelin, a professor at the Department of Nursing Science from University of Turku, Finland, and a member of the SCENE research group, for which the original version of HAPPY was created, acted as a point of contact between the author and the family, and a second interviewer, during this phase.

1.4 Thesis structure

This thesis is divided into three bigger mostly self-contained parts consisting of multiple related chapters. The ordering of these parts is generally chronological with the practical work done on the application as part of this project.

In the first of these three parts, the basic concepts surrounding interaction design are first defined using a literature review in Chapter 2. This is then followed by an analysis, in Chapter 3, of the version of HAPPY that already existed at the start of this project and was used during the two studies previously mentioned in Section 1.1. The analysis in Chapter 3 will focus on interaction, but any information about the requirements of the related two studies that is relevant or likely had an impact on the original design of the application will also be discussed.

The second part focuses on the co-operative design phase with the family (Chapter 4) and their thoughts and observations about both the original and the updated versions of

HAPPY, latter of which was produced during this phase. This is then followed by similar but more in-depth analysis, in Chapter 5, to that done for the original version in Chapter 3. This analysis also serves to document the final design decisions made along with the design thinking and rationale behind those decisions.

In the third part the focus will be on comparing the newly produced version of HAPPY with existing related works, primarily using a second literature review in Chapter 6. This is followed by discussion on the updates made to the accompanying data collection portion (study server) of the application, in Chapter 7, and specifically how the study server influenced, and manifests in, the functionality of HAPPY and the execution of the previous design from the co-design phase.

Finally, in Chapter 8, the work done so far is related back to the concepts of interaction and co-design covered in all the previous chapters while discussing whether something can be generalized from the work presented here and what future work would be required to effectively do so.

Notably, certain typographical considerations have also been made in this thesis because the names chosen for some models by their authors used herein are also frequently-used words in the English language, such as the word ‘tool’, that must be used as both a reference to those models or concepts as well as with their original meaning. As such, when these words are used to reference one of such models, they have been emphasized with a *cursive font* to avoid any possibility of misinterpretation. Additionally, any names referring directly to entities in program code or logic, such as classes and keywords are written with `monospaced` font for clarity due to similar reasons.

2. Defining interaction design

The basis for the design work done as a part of this thesis is rooted in the field of interaction design (IxD). In this chapter, we will define the basic concepts of IxD and go over several helpful existing models and terms from relevant literature to facilitate the following analysis of the different versions of HAPPY.

According to Cooper et al. (2014) interaction design is "the practice of designing interactive digital products, environments, systems, and services" with specific focus on form and behavior. While their definition is concise, they choose to take a more process-oriented view towards IxD in large parts of their work. In contrast Murray (2012) in her work takes a more humanistic and cultural view on interaction and IxD. Her way of defining different models for interaction and interactivity is also generally more useful for analysis, which we will focus on here, as opposed to Cooper et. al whose chosen process presents methods applicable mainly in the context of the software development industry rather than a research project where analysis and critical evaluation is also an equally important part of the results in addition to the produced artifact or software itself.

First, we should answer the fundamental question of what interaction as a term means at a high level. For example, Merriam-Webster dictionary[2] defines interaction as "mutual or reciprocal action or influence", in other words any action followed by reaction impartial to who or what initiates the interaction or how significant or meaningful it is. Of course, in the interest of relevance, most analysis will focus on the more significant interactions or any perceived lack of interaction or feedback, which usually means the absence of some expected reaction when talking about software.

[2] https://www.merriam-webster.com/dictionary/interaction [Jun 24, 2018]

2.1 Ways of describing digital interaction

In this section, the three most relevant models of interaction or interactivity out of the four presented by Murray are briefly covered and elaborated with examples. Starting with the related *machine* and *tool* models and followed by a brief look at the more distinct *companion* model that explores the idea of digital systems being presented as social entities to be interacted with.

The reason for excluding the fourth model presented by Murray is because by its very nature that style of interaction is almost exclusive to systems and situations where the interactor expects or wants to be challenged or engaged in some way. This can go as far as setting goals or expectations for the interactor beforehand, which is very different from the other three models, and is often the case in video games, gamified systems or in gameful design for example. Hence the name of this fourth model the *game* model (Murray 2012 p. 379-408). However, this model is not directly applicable in the context of this project because we already have an existing extrinsic motivator in that the software is used to participate in research projects. We are also not looking to increase quantity at the expense of quality, so adding additional extrinsic motivators would at least initially appear to be counterproductive.

2.1.1 Tool or a Machine, the opposing yet related models

In the *machine* model of interactivity (Murray 2012 p. 321-344) the emphasis is on the output of computer-controlled procedures or processes. When a product or software, which are digital artifacts, has dominant characteristics of this model, the interactor is presumed to be most interested in what the process that is being run produces as output for a given input rather than how it happens. Moreover, they may also be interested in controlling the process at a high level particularly if it takes a non-trivial amount of time.

The ability to control *machine* interactions often means controlling when they start or end sometimes augmented with some high-level information about the underlying process as it is underway. Essentially with the *machine* model the view to the interactor is either a con-

trol panel, a report or any combination of the two. In systems which exhibit dominant *machine* characteristics this reporting can be real-time, a part of the output or the output itself. Although many common *machine* interactions, such as those that appear often in software dealing with a collection of data, can also be entirely controlled by their inputs. Examples of this are user actions such as searching, sorting and visualizing data in different ways. This is the usual result of the common desire for the output, i.e. the interaction results or feedback, of to appear nearly instantly or as fast as possible. This explains why the control is front-loaded even if it means more complex inputs or repeated actions in some cases, e.g. if the interactor fails to find what they are looking for.

Looking at the previous descriptions of the *machine* model it could almost be thought of as the default model of interaction for software, and specifically more traditional applications. This is because often if one looks hard enough and deconstructs interactions enough eventually smaller aspects or features that use this model of interaction can be identified at the lower levels. This is simply a byproduct of the way in which we traditionally interact with most computer-based systems. According to Murray (2012 p. 322): “Tools augment human effort; machines replace human effort.”. The previous way of thinking is directly in line with a common design goal, in many pieces of software, to reduce such manual effort through automation of tasks. Which means that often the simplest forms of interaction, the axioms or interaction primitives provided by the platform itself, are biased towards the *machine* model for historical reasons.

Then, in contrast, when thinking about what exactly a *tool* is, we can already see that according to Murray (2012 p. 291-319) the *tool* model of interactivity has the focus on empowering the user; or extending the expressive power of the hand as she puts it. Importantly, there is usually much less automation as a result and any automation that exists should be secondary while also ideally being optional. This leaves the user with a high degree of control. An example of a software that is closer to a *tool* could be a word processor while its *machine* counterpart for a similar task would be a voice recorder application. The corresponding analog equivalents here are a pen, with a notebook or journal, and a tape recorder, respectively.

However, looking at the example of a word-processing application, depending on the features it provides, even applications like this which could be considered *tools* based on the primary ways and the context in which they are often used can still have multiple prominent *machine*-like interactions. Examples of features like this for a word-processing application would be automated grammar checks, especially the kind that largely bypass user confirmation, or reference management features etc. In practice most applications are located somewhere in between these two models incorporating elements from both. These *tool-like machine*s which Murray (2012) calls "expressive machines" still in large parts represent the *machine* model but are more *tool* like because they empower the interactor in some way to express themselves. The specific examples of such expressive *machines* given are photo and video editing software (Murray 2012 p. 337).

This relationship between *tool* and *machine* interaction is thus easy to see as two somewhat opposing models which can and do often co-exist (see Figure 1). We can easily see this from the previous examples and how the concept of an expressive machine is needed to fill in the gap. *Tool* is often simpler, with a well-defined purpose and function, and the interactor remains involved or hands on while a *machine* is often complex and the interactor in some ways less involved or more distant.

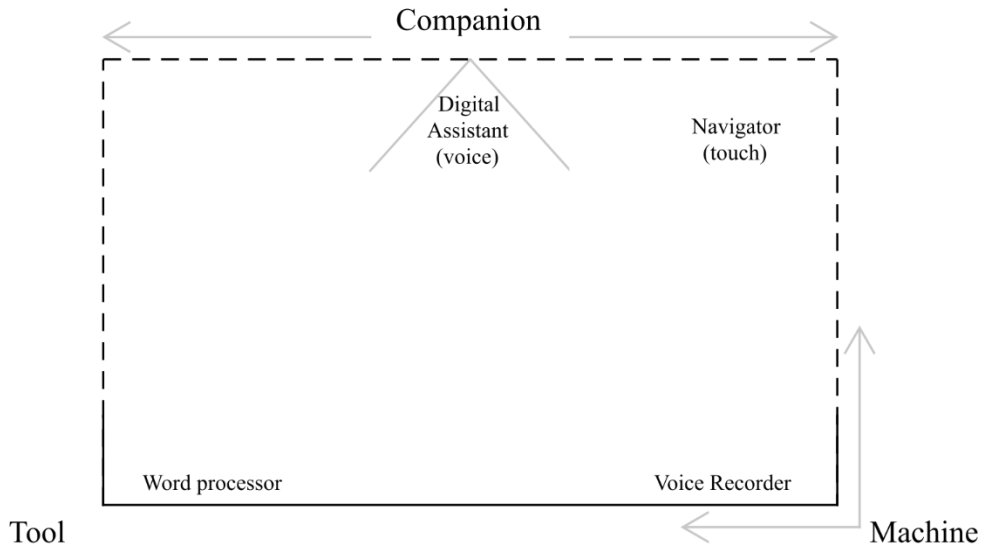

Figure 1 One possible mental representation of the three related models of digital interaction with each model placed relative to the others as a rectangle.

Alternatively, when looking at these two models from the point of view of versatility, a versatile *tool* is such because of how it is used but a *machine* can also be versatile because of its different functions. For instance, using analog examples, a pen would be a versatile tool because of how we make use of its rather simple ability to leave a mark and a food processor, on the other hand, is clearly a versatile machine because it usually has multiple settings or modes for different tasks built into the machine itself rather than because of how we engage with or use it.

Next to look more closely at the specific examples given for an expressive machine by Murray, as it is the ever more relevant grey area. The term of an expressive machine remains intuitive if the emphasis is on a *machine* transforming something that already exists, such a video. After all regular videos are not usually created frame by frame from nothing but shot with a video camera before then possibly being edited digitally. However, suppose that the emphasis is instead on creating something truly new and let us take an image editor that is used by the interactor only to do freehanded drawing as an example. If the interactor is working with a blank canvas and only a set of *tools* such as the pen, brush, ruler or eraser, their way of interacting with the software at a high level is still very much that of a *tool*. On the other hand, if the same image editing software is only used to remove red eyes, frame and crop or fix exposure on vacation photos, then the way in which the interactor interacts with and sees the image editor as is indeed a *machine*, simply further processing the existing product of another *machine*, in this case a regular camera.

Of course, if deconstructed further and looked at more closely, even the familiar pen tool in most image editing software will again default to more *machine*-like interactions at a low level. This happens especially when the program tries to expose control to the interactor while trying to simulate the pressure and angle of a physical pen or a brush. These can typically correspond to e.g. the width or thickness and darkness or opacity in the drawn line. The *machine*-like interactions clearly show in the interface design of the advanced properties of these tools specifically. Without the use of an external peripheral such as a drawing tablet and an electronic pen to more closely retain the original style of interaction from its analog counterpart, the pen tool will at that point have styles of interaction from both the *tool* and

the *machine* models. The significance of these interactions depends on whether we choose to emphasize the specific actions taken or the interactor's intent and context more in the analysis.

The takeaway from the previous example and the idea of an expressive machine, is that the same digital artifact can be both a *tool* and a *machine* depending on how it is perceived and used by the interactor. This can also be influenced by the context and external factors such as the used peripherals or even the hardware of the device it is being used on (e.g. do we point, drag, tap or click using a touch screen or with a mouse).

2.1.2 Companion model

Compared to the previous *tool* and *machine* models, which emphasized functions that extend the expressive power and efficiency respectively, the third *companion* model emphasizes presence and behavior. As a model, it is distinctly social compared to the other two. (Murray 2012 p. 345-377) If interaction is communication between the interactor and the interaction system, in a system placing heavy emphasis on aspects of the *companion* model, this interaction can sometimes even take the form of a literal dialogue.

The obvious examples of systems and services making use of this social model of a *companion* are those that recommend or suggest the interactor do something based on user input. Examples of these are car navigators or digital assistants which aim to clearly solve the user's query by acting on it or proposing solutions with usually significant autonomy and elaborate responses or continuous interaction. These kinds of systems are often present, effectively waiting to react to the first indication of user engagement. However, unless voice recognition is used for input then it is primarily the reactions, i.e. feedback, from the system that follows this model. Without voice recognition the way in which the interactor engages with the system, a navigator in this case, is still exclusively *machine*-like in terms of user inputs when done with a traditional touch screen or a keypad for example. Although this also means that an application does not need to be like Siri or Cortana to be thought of as having *companion* interactions, however, this identity is easily reinforced the more humanized the interaction is, in both directions.

Elements of all three of the models discussed here can indeed be included even if the primary way of interacting with the system is either a *tool* or *machine* (Figure 1 p. 8). Let us consider our previous example of a word-processing application from earlier. For simplicity's sake, we can think of it as a *tool*, or a set of tools for composing a document. However, some of its features, such as the auto-correction and proofreading, could be done in a way that could be considered more *companion*-like rather than *machine*-like. For instance, by doing and presenting these processes as a series of interactive suggestions, rather than automated corrections or actions. This can be taken even further if the application learnt off the user's habits or the presentation was closer to a dialogue between the user and the program. Of course, the interaction itself would remain simplistic as the options given would still likely be either to accept, from a possible list of choices, or reject the application's suggestion. However, taking it one step further and supposing the application had features such as proposing topics for writing or how to continue writing a document through a writer's block then the application would be more actively participating in the writing process and it would thus have a more constant presence as a *companion*.

When it comes to the *companion* model though, in order to include elements that go beyond simple tone and presentation, the application would need to have some capacity to make decisions about and observe the interactor. This is a needed for any suggestions given to be more likely to be helpful but also introduces the problem of reliability through the possibility of incorrect solutions. At this point it is often tempting to attempt to simply give the most likely correct solution, and when it works the results are quite impressive indeed. However, what if it does not work, in these cases who would be accountable for such errors and would they also carry the consequences? Considering the previous dilemma, when working with applications that appear smart like these, regardless of how smart they actually are[3],

[3] See also: the *(Generalized) Eliza Effect* (Ekbia 2008 p. 311-312), provided that it is a result of intentional design in an appropriate context, the effect itself is not inherently good or bad but as an unintended, or unexpected, side effect of design choices it can be typically viewed as negative aspect of any interaction system regardless of complexity.

it is important to think about the individual responsibilities of the interactor and the *companion* which is the application.

For dividing these responsibilities, Cooper et. al. (2014 p. 179) identified a helpful design principle which they come to when talking about the division of labor and digital etiquette between software and user: “The computer does the work, and the person does the thinking.” So, unless the problem that the *companion* application is aiming at solving is one that can be correctly and reliably solved by a computer with presumed minimal repercussions in case of total or partial failure, or at least the ability for easy correction, an all-in approach to *companion* interaction would be a poor experience.

The focus of a *companion* application should be on helping the interactor to find and evaluate the solutions they need to a problem that they have identified or to help them notice a change, rather than necessarily directly identifying problems or solutions for the interactor and attempting to act on them. Additionally, any direct goals set should originate from the interactor or the environment not from the system.

3. The original HAPPY application

In this chapter, we will take a more detailed look at HAPPY as it existed at the start of the project. The purpose here is not to bring up criticism towards the original work but to create a viable point of comparison for analyzing the work that followed and to ground the early decisions made during the project.

It is also important to do this because there is no doubt that the original version of the application has influenced the further work done both deliberately and unintentionally. Therefore, to be able to critically evaluate the work done as part of this project the same must be done for the work that preceded it and acted as basis for the work that followed.

As mentioned in the introduction the HAPPY application or the "**H**andy **A**pplication to **P**romote **P**reterm infant happ**Y** life" is a research tool for collecting user narratives for qualitative research. What this meant in practice at the time was that the application behaved like a simple voice recorder, even though text input was also technically possible, with a server backend to which the application would then send all input as files categorized by numeric ids for the installation and input category, such as closeness or separation in the case of the prior studies, and time of the entry. Additionally, it was possible to mark a time as a bookmark, an empty uncategorized entry, which the user could then later return to and make an actual content entry for it easily.

3.1 Interaction and interactivity of the original version

Looking at this version of the application from the point of view of the of its interaction using the previous models of interactivity and our examples of the *tool* and *machine* interaction (see Section 2.1.1) it is then easy to categorize this kind of software as a *machine*. However, what makes this initial version of HAPPY interesting is when we consider the question of who the users and interactors in this system are. Because this version of the application was still purely a data collection tool the primary user of the complete system was the researcher interested in collecting data even if they were not interacting with or using the client application themselves. When thinking of the entire system like this the user of HAPPY itself

at this point was delegated to simple source of input data. This idea of one user and two interactors was further emphasized by how the feedback or output from the application was divided and handled making the client application appear as quite selfish towards its user.

This selfishness was because of the feedback to the actual user of the client application being minimal, as seen in Figure 2, any actual produced data being visible and directed only towards the researcher. The user facing feedback being limited to colored markers on the ring of a clock dial that could be navigated either by using the category specific arrow buttons or touch.

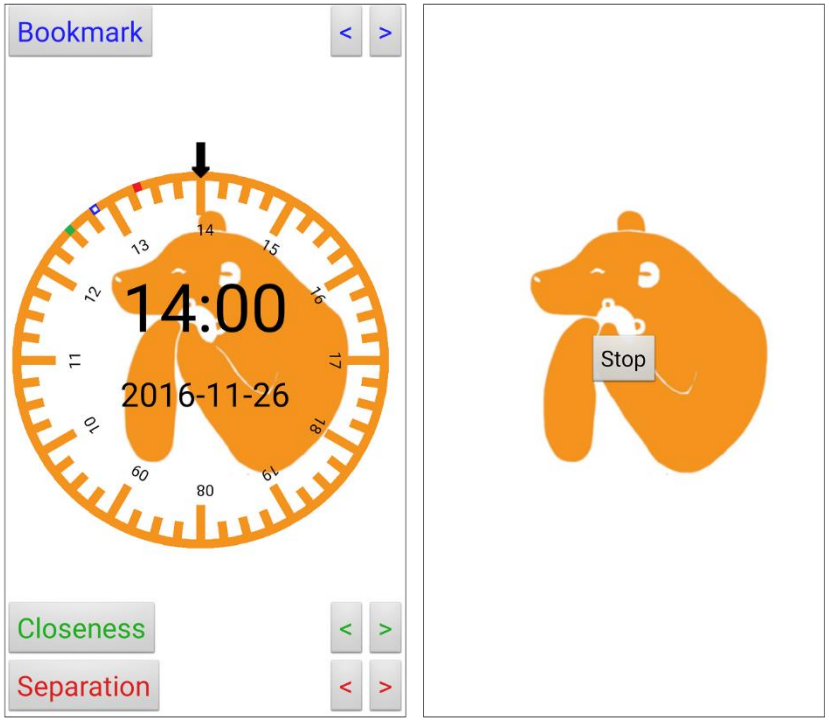

Figure 2 The user interface for the original version of HAPPY, using inputs like those used in the referenced studies (recreated). The main view on the left and the recording view on the right.

Thinking of the interaction the user had with the application as a one-way email, where any record of the email being sent is removed from the system except for the information of an email having been created, would be an apt comparison. The reason for this kind of one-way design is evident when looking at the method descriptions of both studies conducted with this version of the application. Because the initial studies, with the nurses, used loaned devices, from the researchers, as opposed to participants' personal devices. These devices

could then potentially pass from one user to the next throughout the day without the researcher being involved, so obviously the data could not be retained on the device even just through-out the duration of the study. (Niela-Vilén, Feeley et al. 2017, Feeley, Genest et al. 2016)

Despite of the selfishness, of the application, the implication is not that the participant or the user did not get any value out of participating in these studies, however, any value for them would have been related to the study itself and likely not facilitated by the application, aside from the added convenience, thus being similar even if the same study was conducted using more traditional data collection methods. In fact, as far as fulfilling its intended purpose in the context of those studies and their specific requirements the application seems to have performed well (Niela-Vilén, Feeley et al. 2017).

3.2 Additional design goals identified from the original version

Even before the start of the co-design phase several clear design goals could be identified just by looking at this version of the application. First and foremost, addressing the imbalance between the two distinct users, the researcher and the participant, as interactors without yet aiming to change the way the application is interacted with.

In practice, this would mean giving the participant access to a copy of their own data. This did raise some interesting questions concerning who owns this kind of data given for research purposes and who is responsible for its handling, particularly if multiple copies are kept and managed separately. To make this goal realistic the only decision to be made here was that the researcher's responsibility is, as before, the confidential handling of any data entrusted to them while final responsibility of the participant's own copy would lie with the participant themselves and their handling of the application and the device it is used on. The responsibility of the software in this regard would only extend as far as with any other application. The researcher's responsibility would then be to vet the application they intend to use and the software vendor's responsibility to follow good practices when handling the local storage of application data. Notably in some cases the researcher and the effective software vendor are the same person or group as is the case in this project and the preceding studies.

During this early phase, which in terms of practical work concentrated on learning the original application not just in terms of the apparent design[4] but also in terms of implementation, another design goal would end up being improving or replacing the bookmark feature. This need further reinforced the research question about support for additional formats.

The rationale behind this additional goal being that, in its implementation the bookmark feature was only a timestamp or empty entry as described earlier and, due to its characteristics as such, while it was a reasonable assumption at this point that it may work well during short data collection periods it would be likely to lose effectiveness, if the participant could only return to fill the bookmark the next day or even later for example. Additionally, even though it is likely just an unforeseen side-effect of the implementation if a significant one, a bookmark that a user would never return to would not be reflected to the data collected in any way nor would it be easy to see, if user created an entry under such a bookmark or directly after the experience or how much time had passed which is particularly relevant when considering any potential recall bias.

From the research side, the desire for automated transcription of collected audio recordings also came up at this point, however, since it does not concern the existing feature set of this version of the application and has more to do with background analytics than interaction directly, although it would act as an enabler more diverse interaction in the future, it will be returned to during later chapters. In particular why it did not end up manifesting as a concrete goal in the end and any related considerations throughout the work that followed.

[4] The original design process of the application was scarcely documented, so this was also an exercise in trying to see the design behind the implementation and codebase that was provided.

4. Co-operative design phase

This chapter focuses on the co-operative, or participatory, design phase of this project. It consists of two interviews, with a period of testing with an updated version of the application in between, with the family that agreed to participate in this project with us.

The updated version of HAPPY conceived here was developed following the first interview with the family. They then participated in a brief (1 week) period of testing with it before the second interview. The focus in the follow-up interview was on collecting open feedback from this testing and about the participants' views and experiences with the changed application. This second interview also contained some forward-looking discussion and thoughts on how the application could be developed further.

Both interviews were held at the family's home with both parents present in a four-way discussion format in Finnish and recorded as audio recordings. These recordings were then transcribed verbatim and anonymized by removing or changing any identifiable information. A. Axelin, from the Department of Nursing Science at University of Turku, was present during all meetings with the family, during the interviews as a second interviewer, and an active participant in all discussions with the family.

Sections 4.1 and 4.3 contain translated overviews for both interviews based on the final interview transcripts in a summary format. To better preserve readability some re-organizing of the order in which points came up has been done to group related themes together. Between the interview summaries, Section 4.2, is a very high-level summary of the development of the test version to provide some additional context for the follow-up interview. For a more thorough look at the updated version created here, as well as the detailed design thinking behind it, see Chapter 6 for the analysis and Chapter 7 for discussion on the data collection aspect, developed later, and its influence on the realization of the design created here.

4.1 Initial interview (Context and Design)

The first interview (March 2017) with the family can be roughly divided in two parts. First, the interview started warm up questions about the context, that is the family's experiences of preterm birth and the NICU and how and to whom had they talked about these experiences about, led by A. Axelin. This was followed with a brief look at the version of HAPPY that had been used previously, as covered in Chapter 3, without going into any technical details or many of the specific design goals from Section 3.2 although the need to add ability to view created entries was mentioned during the interview. The high-level comparison used for the software was that of a diary and the idea of describing the original version as selfish also began here.

During both this and the follow-up interview, two things had to be taken note of to be able to consider the contents and presented ideas in context. Firstly, of the parents the mother had a smartphone, although not an Android device, while the father had a more traditional basic cell phone as their day to day devices. Secondly, while the mother had experience in writing personal diaries or journals, the father did not. On the other hand, for the father, his professional background as a doctor also had an impact on the things he chose to focus on and find inspiration from.

The themes to highlight from the first half of the interview, apart from the rich description of the events that lead to NICU from both parents, are eventually the feeling of safety and trust in the NICU, after the initial shock, as well as there being some things that you can only talk about with people who have gone through the same thing or are in the same situation. Their description of the communication about the situation was overall centered around each other or the close family, but this is of course dependent on the individual as was noted.

The second part was more centered on the possibilities of the application itself, ways it could be used and actual design ideas. One question that came up here early on from the mother, and was later revisited, was whether this software should be something for the individual or something shared, for example a social software environment for the whole NICU, even the notion of NICU internal "Facebook" came up later.

For the father, his way of looking at the application and its future possibilities was to think about the idea of recording his thoughts also for someone other than just himself, in that sense bringing up the idea of a recipient, such as psychotherapist. He also noted that he would probably be more likely to focus on recording more about negative things such as anxiety or fear, noting that he would not see the need to record something if everything is fine because those things are already easy to talk about.

While the father clearly focused more on externally driven and structured way of using the application, the mother also brought up the idea of a concrete recipient at one point when thinking about the scenario of having something on her mind but a member of the staff not being immediately available, saying that in those instances it could be useful to have something to note things down on.

Here two kinds of scenarios could be clearly identified: the application as a personal tool for reflection, closer to the original idea of the diary, or a tool for recording thoughts, feelings or questions for some recipient whether that was to be a therapist, nurse, doctor, family member or even potentially more than one person.

To move the discussion more towards the idea of expanding the design, and to try and distance from the singular idea of recording just audio, the scenario and example that was given was a situation in which the user might not know how to start writing or talking and then whether it could be combined with something abstract such as starting with a color for instance. This quickly led into talk about using different symbols such as icons or emoji, and then towards the end the idea of also using photos or even video came up.

This then eventually, through discussion, refined into the idea of having three distinct elements after pointing out that it would not be good if the act of making these entries was too simplified to the point of it becoming trivial. These elements would be for example emoji, color or icon, followed by grading e.g. on a scale from 1 to 10, then combining that with text or audio. Which, as was brought up later, could then be turned into a graph or a timeline or

otherwise visualized in different ways. The father also pointed out that when offering different formats like this alongside each other, then even the change in user habits can be valuable indicator to a professional in and of itself.

Towards the end the discussion turned back the idea of sharing information, for example, between family members. It came up that it could be interesting, or as the mother pointed out, even natural in some way because this is a [phone] application. One idea that was specifically proposed at one point to the parents here, by A. Axelin, was giving parents access to each other's entries, or using the application to more literally communicate with each other, such as writing messages.

During the discussion that followed it was noted that the idea of sharing and whether it is applicable would depend on the individual and the kinds of entries being saved into the application. The entries would quite likely focus on different things depending on who has access to them. The mother described the idea of sharing in this way as potentially charming but also maybe a bit strange for her personally. The idea of a timeline or a day view with both parents' entries visually presented also came up around this point.

When the discussion turned to aesthetics the parents mentioned that the original HAPPY application, using the icon of a bear, felt from their impression approachable. The father specifically brought up making it simple, with few clear steps, referring to the design idea using emoticons from earlier.

As the discussion then briefly turned back to the idea of using the application in the NICU specifically the idea of using the application to help the parents to make better use of some of the routines present there was also brought up here, revisiting the idea of making entries for the express purpose of bringing them or questions from them up later with a staff member, such as during the doctor's rounds or in a weekly meeting.

To end this interview, it was agreed that a next version would be developed based on this discussion and the ideas that it highlighted. This would then be followed by the parents trying out this updated version through a short test period.

4.2 Development of the test version

This section presents a brief high-level overview of the development of the test version with focus on the things brought up during the initial interview. The purpose here is simply to give additional context and help relate the work done between the interviews to the interviews themselves giving this chapter a more cohesive and chronological structure.

At the start of the work towards the next version, the idea of automated transcription was still being considered, however, as a result of certain technical challenges and limited options when it came to supporting speech to text for the Finnish language, at the time, together with the fact that the idea of text to speech or anything requiring it on the backend did not come up in the first interview the natural decision was to focus on addressing the other goals presented in Chapter 3 together with implementing the main design idea of using an emoji or other symbol along with grading as a type of an entry. The idea of automated transcription will be revisited in a later chapter when looking at future possibilities.

The goal with the test version was to present a version close to a minimum viable product (MVP) of the new HAPPY that included support for text, audio, photo (picture) and "emotion" input types. The last of which is referring to the form of input conceived during the initial interview with the parents. The supported types would also need to be able to be combined freely under entries and past entries should be easily viewable and able to be amended with more input types later. Basically, to allow mixing and matching different input types under the same top-level entry. Notably the existing network dependency, that is the data collection or study server, was not updated until the test period was over and evaluated. The reasoning for this being in case the evaluation of the parents would show that this updated version was not a workable concept along with the fact that the test would have to be conducted using loaned devices with no guaranteed network access. Essentially this meant that the version would have to include a local mode usable with the device effectively in airplane mode. The details of the later updates to the study server will be covered as part of Chapter 7. The inclusion of video as input format was omitted because of the increase in storage and memory requirements that video files would create, in addition to being less

feasible to be collected over a, potentially metered, network connection if used in research context.

After the development work was nearly concluded a short additional meeting was scheduled with the participating family to hand over the test devices. This meeting included a brief demonstration of the updated version together with verbal instructions on its use. Written instructions were deliberately not provided. This way the parents could still easily pick up the use of the software, since any kind of first time use in application instruction did not exist, without the temptation to refer to written instructions later on how to "correctly" use it. The test period was one week in June 2017 and the parents were instructed to mainly make entries of anything related to their children.

4.3 Follow-up interview (Evaluation)

The second interview (June 2017) began with the parents' free form account of their experience of using HAPPY during the preceding week. As per their own expectations the mother with previous experience of keeping a diary was more active compared to the father who had no such experience. Moreover, the level of the father's participation was also clearly influenced by other commitments during the test period which has to be taken into account along with him being less familiar with smartphones in general. Irrespective of that the contribution of both parents during the test was valuable, specifically due to their different backgrounds and level of previous experience, along with the forward-looking discussion that also happened as part of this interview. In this section, unlike for the first interview that was presented more chronologically, the different themes covered during the interview have been gathered into their own independent sections to improve readability.

4.3.1 General opinions about using the test version

To highlight some things from the mother's account specifically she though that the application was easy enough to use, approachable and mentioned liking the emotion input, or "emoticons" as she referred to them, perhaps the most. Although she did also note that because entries could not be titled as such, only appearing as "audio recording", perhaps that

did contribute to her preferring the emotion input type also. Since by using them the resulting icon would be clearly visible when viewing the log of previous entries. In general, her concern was that viewing and tracking what previous entries were about was difficult if one was to return to older entries to view or add to previous entries with how the log or history viewer was implemented in this version. She also noted that having it on a separate device and not the phone that she was regularly using was cumbersome, as was to be expected, since this was the case out of necessity rather than it being ideal.

Here it should be noted that the possibility of adding support for titles had come up, from the author, during the meeting where the use of this version of HAPPY was demonstrated and the test devices given to the parents. Even though it was something that had been considered during development, as it was a feature that did not exist in the test version it could be argued that mentioning it should not have been part of that meeting.

Just as before, it was possible to see two clearly different ways to use the application. One way was to focus on using it for entirely personal entries and the other with some clear recipient or at least with the expectation to go over the entries done systematically at a later time. The father was clearly more inspired about the latter while the mother seemed more interested in the former. This clearly correlated with their respective experiences in doing this kind self-chronicling when considering the mother's previous experience in keeping a journal and the father's experiences from both his work as a doctor and what he had told us previously about his personal background. However, a central theme irrespective of how the application was or could be used, was clearly how to better interact with previous entries to get more out of using the application than just the inherent benefits that come from in some way recording one's own thoughts at the moment it is being done.

When asked what kind of things they had been making entries of during the week both parents said that the entries were about children mainly, as was suggested at the start of the test, but the father did mention that he made one entry related to work out of frustration. But in general, for both parents, the entries were, according to them created based on a particular emotion or situation where emotion was present. The father said that for what he had time to use the application for he mostly used it in retrospective way, with the motif being "How've

things been going…" the mother likened this to using it more like dictation, which makes sense considering his profession. Although, the mother also mentioned that the fact that it can be used in multiple ways like this is a good thing.

The parents' comments about HAPPY, here specifically in the NICU context and whether it could be useful there, were positive and both seemed optimistic but noted that it would again depend on the person, their situation, and in what context specifically HAPPY would be used in the NICU. The mother mentioned specifically that, compared to the week-long test here, at least in the NICU there could be more time to use this kind of application, provided the person is interested in such a thing in the first place.

4.3.2 Usability and technical issues

Apart from the application being installed on a separate device being an inconvenience the mother also noted that she experienced some technical issues with using the touch navigation implemented in the history viewer to switch between days, specifically with it feeling unresponsive. She also noted that she had some hard crashes related viewing entries towards the end of the week. Although part of the unresponsiveness could be attributed to the touch screen of the test device and at the time it was difficult to identify which it was. She mentioned typing using the device's on-screen keyboard being horrible. She also mentioned that while editing most entries would not likely be needed, being able to fix typing mistakes in text-based entries would have been nice.

From the father's side, the discussion almost immediately turned to the idea of having a common element between the entries, such as the emoticon being required, which was revisited again later in the interview. There was also comment on the division of the history shown on screen at once. Whether it should be per day as it was implemented, per week or all entries at once. He also mentioned that he found it difficult to get back to the history at times. The possibility to visualize the history more like a graph was also brought up by him, echoing the idea of going through the contents of the entries in some way with a professional that also came up during the first interview.

As the discussion later returned to usability, along with the issues of accessing history view, which was reiterated, another issue they had been having seems to have been related to the clock widget or dial that is part of the main view of the application visually identical to the original version discussed in Chapter 3 (see Figure 2 p. 14, left screen). Specifically, how its implementation at the time managed the currently selected or displayed time made it difficult to reliably access the history view by the button being context sensitive to the selected time as well as how it was sometimes tedious to create entries in the present moment if a time in the past was previously selected or the time did not otherwise update when the application was in the background for example.

The father's comments on the clock widget were that it might help usability if it remained more in the background, and thus optional, without the need for it to be manipulated so often. This led to the conclusion that creating all entries by default for the present moment would be best assumption for the application to make from both parents. The mother specifically mentioned that she did not think that it was important to even necessarily to be able to create entries in a specific point in the past. This she justified by noting that if she was going to create an entry about, for example, the early morning later in the day that it being related to the morning would likely be apparent from the contents of the entry. Thus, the real creation time might as such be enough or even more valuable than if she specifically created it under a point of time in the past using the clock widget.

4.3.3 Forward looking functional and design ideas

During the interview, as the concept of a common element was re-emphasized particularly by the father, a slightly altered version of HAPPY was briefly shown to the parents that replicated this idea by making the emotion entry the sole primary input type and allowing access to the other input types through it. This was simply a version prepared beforehand, by making a minor edit to the configuration file the application was using, and the fact that it happened to come up organically was pure coincidence. The motivation for doing this, at that time, was because having such common element would actually help with realizing some of the ideas that had come up in the first interview and now again here, such as presenting all

entries in a graph or a timeline, that would require a common visual element to make it possible to relate the top-level entries to each other and to themselves.

When asked if this kind of approach would be a negative, the mother commented that in her case it would not be but that it might be because her way of using HAPPY was already mostly guided by emotions. Noting that other people might just prefer to record or write straight away. Having said that she also noted that it would be possible to just choose a neutral emoticon in those cases.

The father's comments, not unsurprisingly considering he also brought up the idea of this kind of common element, were positive. Noting that its even in the name HAPPY and that the emotion fits as a common denominator because of that. Likewise, he thought that this would make any debriefing or [human] analysis easier afterwards by allowing exactly the sort of visual representations he had also been thinking about.

It was mentioned that which of these two variants would be best probably depends on the exact use case, such as a potential research question or whether it was simply for personal use. Here the mother referred to HAPPY as, what can be translated as both, a mood or feeling diary, which is a relatively good description of these versions of the application.

When then asked about the range of icons or emoji available through the emotion input type, whether something was missing or redundant, the mother said that she did not feel like anything in particular was missing and that even though there may have been few particular emoticons, such as one with tears flowing, that she did not feel the need to use during the week if those were then removed there would now be a risk of not having them for someone using the application who could feel that way. The father's comment on this was that the simpler the better but mostly ok.

Before ending the interview, a few more clearly forward-looking ideas were discussed. Firstly, the idea of sharing information with others through the application itself, which was discussed during the initial interview, and following that whether the parents thought that there would be need for features to make the contents of the application more private using password or something similar. Regarding sharing, the comments were that it was not missed

in this version and that the entries might be private which meant it might not fit at all. The consensus was that if any sharing was to be facilitated by the application what was shared should be different than the offerings for your personal use and that the sharing should be controlled even as far as having to specifically select each time what to share and with whom. As for privacy features, both parents agreed that any privacy features already present on devices, such as an unlock code, and provided by the system would be enough and that it would likely affect how much the application would be used if it would, for example, ask for a password every time the user went to create and entry or use the application.

Right at the end the father also suggested changing the emotion views inputs such that instead of having them all on the same screen, as in the test version, the icon could remain at the top while changing the color or grading of the emotion could then appear separately one at a time below it. The mother commented that she preferred being able to see all the options and what was to be done next all on the same screen instead. Both of these approaches have their advantages and disadvantages which will be revisited through the full analysis in the next chapter (see Section 5.3.3 p. 39).

4.3.4 Looking at the test week through data

While the amount of data from this week-long test is not massive, together with the parents' comments from the follow up interview, a few interesting things should be noted about their use of this version of HAPPY. Of course, nothing should be generalized from here, however, the data in terms of numbers does correlate the comments during the interview and highlight what was said about titles or some additional support for descriptive information being helpful if available when viewing entries retrospectively.

Throughout the week the mother created top-level entries for 15 different timestamps, each containing 1 to 4 separate entries, for a total of 27 entries. One file for an entry was lost due to a software bug. Of the 26 entries, that were correctly saved, 7 were using the emotion input type, 3 the picture input type and 11 were audio recordings. The remaining 5 entries were short one sentence text entries acting as captions for preceding emotion or picture entries. All camera pictures included this kind of captions and usually when an emotion icon

was not elaborated by this kind of text it was followed by an audio recording, except for one solitary top-level entry that contained a single emotion icon. On one of the days all entries for that day were under a single top-level entry created in the morning, it is impossible to say for certain whether this was intentional or not, but it only happened once, and all other top-level entries were about self-contained events.

In the father's case, as he mentioned during the interview, he had used the application less making total of 11 entries under 4 different timestamped top-level entries. These were mostly from the start of the test period with one top-level entry from the end of the week. All his top-level entries contained an emotion icon entry and at least one other entry either text or audio. The remaining 7 entries consisted of 2 audio recordings, 2 pictures and 3 text entries. The text format entries were likewise used here as substitute for captions for preceding entries.

From both of their devices it is easy to see that because the picture input type did not support adding a discrete caption or title the text entries had to double for that purpose. With the previously noted counts the top-level entries, which are what appear initially on applications clock dial or when using the history view, would be what is normally considered the total number of their entries. However, in this more detailed count the respective total number of entries, 27 and 11, for each parent is the same as the number of times they had to specifically use the new entry dialog and select an input type, which could be considered slightly inflated due to lack of the caption feature for pictures referenced earlier. Although, the emotion input also technically has the same issue there the user can at least add a clarifying text entry before the resulting emotion entry is displayed to them for the first time.

5. The updated version of HAPPY

In this chapter, we will take a detailed look at the design of the updated version of HAPPY developed and tested during the co-operative design phase. Starting with a focus on the interface and overall application flow followed by a deeper look at the new emotion input type and the possibilities behind it.

The design thinking behind the features and changes in this version has its roots in the ideas behind Murray's models of interaction covered previously. Focusing on the idea of looking for possibilities to realize the improvement ideas and the concept design from the previous co-design phase in a way that allows for increased interactivity and expression by aiming to incorporate aspects of the *tool* model in particular.

5.1 Overview of the new User Interface

For its overall design language, the refreshed interface makes use of a combination of modern elements from Google's material design[5] initiative, such as the floating action button and cards, together with section dividers and titles reminiscent of older applications from the days of Android Ice Cream Sandwich (i.e. version 4.0). This results in a balance between old and new design elements which fits perfectly for a refreshed version of an application like HAPPY considering the final range of supported Android versions.

For the main view (see Figure 3 p. 30) a version of the original clock dial was retained and improved upon to give the application a clear continuity from its previous versions and effectively retain the original aesthetic and feel with the dominant use of the original bear graphic. This choice was also reinforced by the participating family's impressions of the original version of the application and their thoughts, from the initial interview, that this kind of aesthetic would make the application more approachable.

[5] https://material.io/ [Jun 26, 2018]

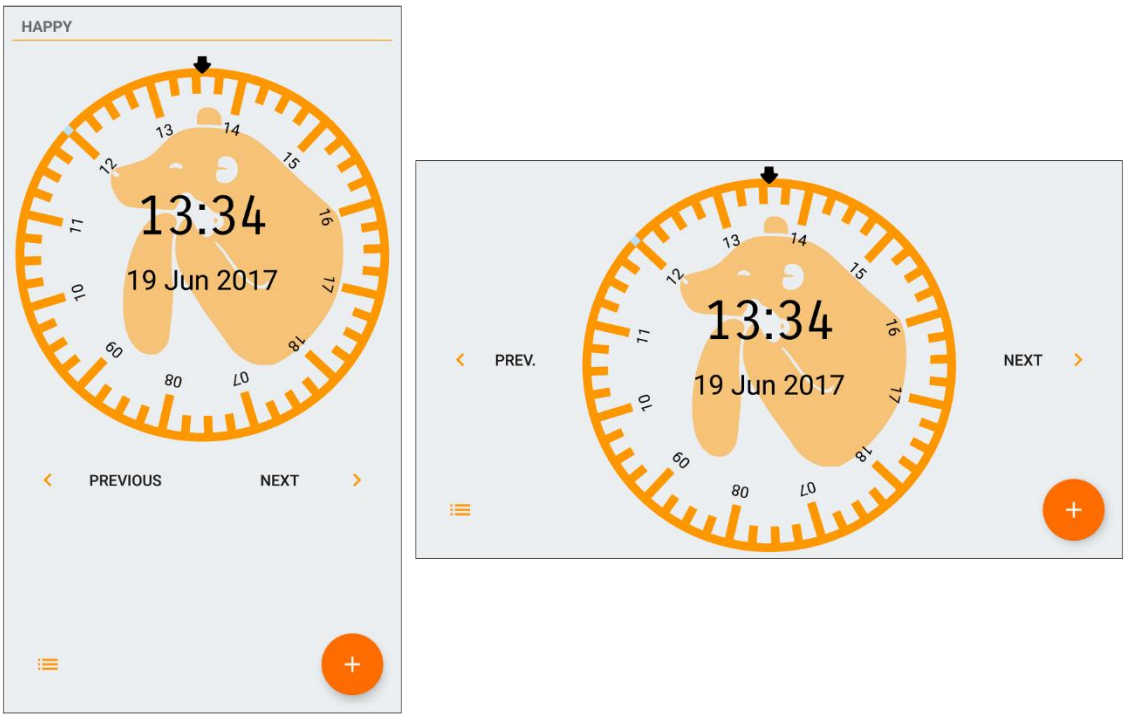

Figure 3 The new main view, for the version used during testing with the family, in both portrait and landscape orientations. Prominently featuring the floating action button and the clock dial.

Other than the clock dial, the main view is less busy with overall fewer interactable buttons. Having just one set of next and previous buttons for navigating between top-level entries on the clock dial and a focus on the floating action button, at the bottom right corner, for adding entries. To have the round floating action button, a material design UI component intended for the primary application task within an application view, is very characteristic of many modern Android applications. This includes majority of applications provided by Google installed on stock versions of Android. As such following this convention is worth it even if the application is not otherwise necessarily always following the principles of Google's material design, because of expectations and conventions it has created for users familiar with the platform, due to how dominant it is. Taking advantage of this kind of learnt patterns and user behavior in interface design is a concrete example of a broader concept referred to as "scripting the interactor" (Murray 2012) related to the models of interaction from Chapter 2.

The biggest challenge of working with the clock dial as an interface element is that it takes up a lot of the limited screen real estate on a device and in order to remain usable and

readable it has strict size requirements. Usually taking effectively at least half of the available screen space on the device screen, while being able to deliver limited amount of information to the user. So, it is also true that the reduction of interface elements, and introducing the floating action button, came to be just as much due to necessity than simply from being a useful interface convention. The input type selection itself simply had to be moved away from the main view, behind a single dedicated button, just to allow supporting a larger variety and number of input types to begin with. Without it, the result on lower screen sizes would have quickly been either hitting a theoretical maximum, potentially as low as two, input types due to elements quickly starting to overlap with the clock itself.

From Figure 3 we can also see that most of the buttons present in the main view are simply icons, lacking normal labels[6], to avoid the same issue of overlapping interface elements due to differing lengths when localized to different languages. Even the words next and previous have sufficiently different lengths when localized to encounter this issue, without abbreviated labels in landscape mode even in English and more so in Finnish which is the second language supported by this version.

Despite these challenges with clock dial for this version, in retrospect, apart from the feeling of continuity mentioned earlier the use of this interface element is also the main reason why the application retains its unique feel rather than becoming a cookie cutter Android material design application, which would have been the path of least resistance by far when it comes to designing user interfaces for the modern versions of the Android operating system.

5.2 Changes to adding and creating entries

The process of adding a new entry in the updated version can be seen in Figure 4 (p. 32), where application flow is depicted from left to right. The leftmost view is what the user sees after interacting with the orange floating action button. The displayed modal dialog is

[6] All the buttons and other significant elements do in fact have hidden labels for accessibility services, such as screen readers.

context sensitive and the options given will depend several factors such as the application, device or server configuration (see Chapter 7) and whether the user has previously selected an existing entry on the clock dial using either touch navigation or the next and previous buttons in the main view.

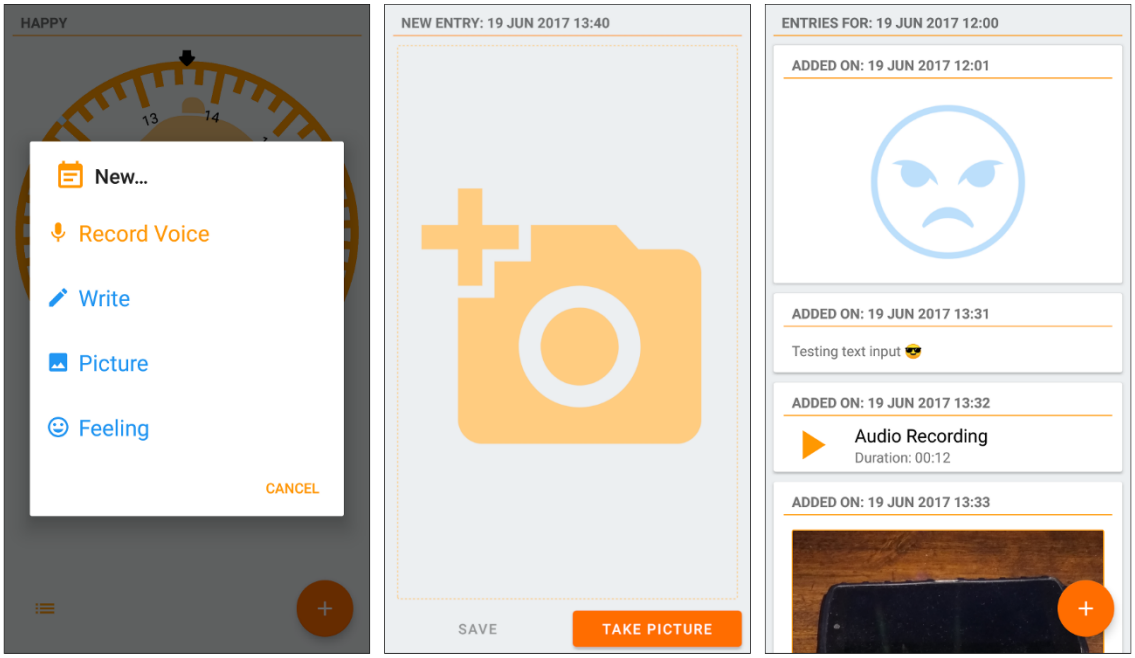

Figure 4 The application flow (left to right) for the new version when adding a new entry using the "Picture" option (center screen), other some other new input types are shown in Figure 5 (p. 33).

Figure 4 specifically depicts a case, as with the version tested by the parents, where all possible options in this version are immediately available to the user completely independent of each other. As noted earlier, the modal dialog was used to permit the list of available input types to grow freely regardless of available screen space. However, if for instance only one input type is configured for use by the application the modal dialog is skipped entirely, moving straight to the input type specific view (center screen) for creating the entry itself. This is equivalent to the alternate scenario presented during the follow-up interview where only the emotion input type is initially available through the floating action button.

The rightmost view in Figure 4 is the state of the application after the user has finished adding an entry. In this view the user sees the entry they just added along with any other

entries added for that specific top-level entry, i.e. point in time. The floating action button in this view functions identically to the one in the applications main view.

To immediately return to the application's main view the operating system provided "back" button may be used at any point. Additionally, while this is not depicted in the images, in the event of a view containing unsaved information the user is asked for confirmation to prevent loss of information due to accidental presses of the back button. Convention that is common in software across devices and platforms and thus speaks for itself but absent from the original version due to its unintuitive use, or lack thereof, of the system provided back button in regard to the navigation between different views of the application.

5.3 Closer look at the interfaces for creating entries

Next, to look more closely at the process of adding an entry, this section focuses on the interfaces displayed during the actual creation of the content for an entry. For this version of HAPPY the focus was the emotion input type, which can be seen again in Figure 5 side by side with the significantly simpler interface for recording audio entries.

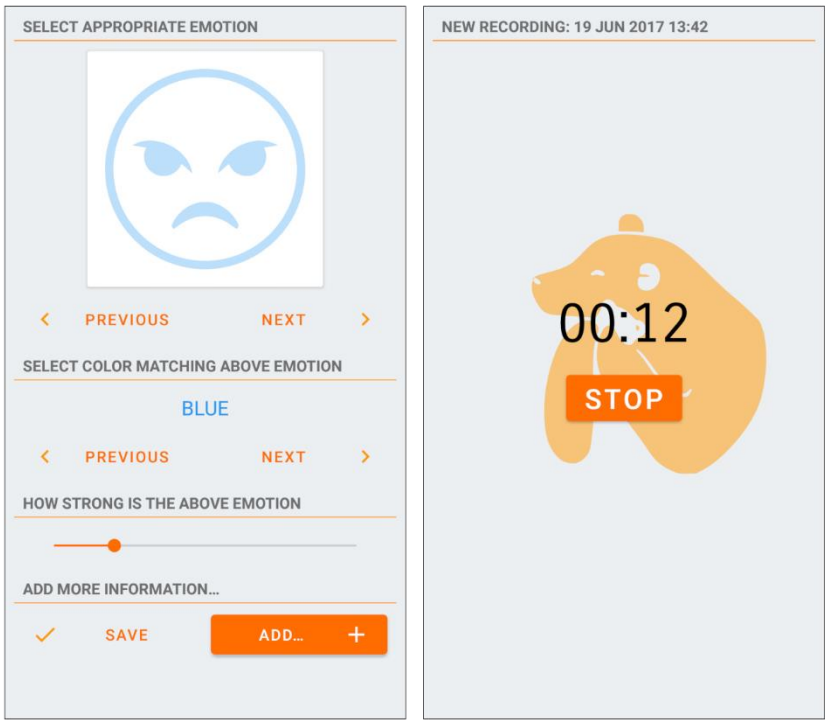

Figure 5 The new interfaces, from left, for the emotion input and audio recording side by side.

While the focus may have been the new emotion input type the simplified audio recording interface (Figure 5, right) still received minor changes to its original counterpart (Figure 2 p. 14, right screen). While these changes are mostly aesthetic the addition of the running timer making the view less static, for example, does reinforce to the user that the application is actively recording. This could have been further increased by creating an interface element that would directly react to the audio captured by the device microphone such as e.g. a stylized audio visualizer or a wave form. However, this was omitted in the interest of time during development settling for just the running timer and the caption changing depending when the recording is being made.

5.3.1 The emotion input type

The new emotion input type (Figure 5 p. 33, left) allows the user to create a visual representation of a feeling by asking three questions. First, to select an appropriate base emotion from a list of pre-existing icons, which is typically followed by selecting the color and grading the emotion based on how strong or intense it is. All these choices then manipulate the resulting icon by changing its base color and the shade used. Here, lighter shades are associated with less intense or "weaker" emotions and brighter more saturated shades with stronger emotions. Once the user has made their choices, there is an option to add additional information by combining the icon with any of the other supported input types, such as text or audio, or saving the entry as is. Adding some additional information is subtly emphasized by button placement[7], and later color, to encourage the user to give further information about the experience that resulted in them feeling the depicted emotion. The emphasis by color, while pictured in both Figure 5 and Figure 4, was added after the follow up interview.

In terms of the low-level design thinking, a few conscious choices have been made here. Firstly, the application tries to avoid giving explicit descriptions about the icons that

[7] Specifically, for onehanded operation in portrait view, by making the placement of the more desired action match the typical placement of the default action in system dialogs. The user must also specifically go out of their way to reach over the button to save the entry directly, but this aspect is specific to righthanded people and LTR (left to right) layouts.

are available this is done to avoid reinforcing associations between certain colors and a specific icon as well as to allow the user to decide the specific meaning of each icon rather than asking them to attach predefined labels such as “happy” or “sad” to the emotion they are feeling. For example, looking at the resulting icon from Figure 5, if the color chosen was red perhaps it would be easy to assume this icon represents anger but because the color is blue does it still represent anger or perhaps disapproval or dissatisfaction instead. Similar questions can be asked if the icon chosen was for example the heart, e.g. what the tone difference between a blue and a red heart is and are both equally positive.

The goal of not providing the user with explicit meanings for these colors or icons, by not labelling them in the interface as anything specific, is to actively try to avoid giving users the answers, that is the designers answers, to questions like the additional ones given here or those presented directly as part of the application interface. The thinking here being to make the user of the application think about what the icons and colors represent to them rather than tell them what they should think. This thinking is in line with Murray’s philosophy for the *tool* model of augmenting the expressive power of the interactor (Murray 2012). Ideally this would happen each time when creating an entry of this type, rather than just once when a particular combination of a color and an icon is used for the first time. For more on why this is significant, see Section 5.3.2, we need to look at how people associate colors with different meanings when prompted to think about or describe them.

In terms of the placement of the emotion input type in the flow of creating an entry, about a particular event or feeling, ideally it would be a precursor to using the other input types to provide richer contextual information. This input type can also be seen as an answer to the design goal presented earlier in Chapter 3 (see Section 3.2 p. 15) because it could be used as a replacement for the legacy bookmark input type while addressing its issue of not leaving a record on the server side. It also gives the user more than one point of reference if they want to return to elaborate on an entry at a later time, rather than just being record of the time of something of interest. The picture input type can of course fulfill a similar role; however, the emotion input type is also usable for this purpose in situations where taking pictures is not possible due to e.g. privacy reasons. The emotion input type could also in theory act as

a prompt of sorts for creating an entry since it asks users questions about how they are feeling so it might give users who have difficulty getting started some simple ideas on how to do so.

However, its front-loaded placement does place constraints on the complexity of it because causing the user to take too much time on this might affect their willingness to combine it with other input types. It is a delicate balancing act because it cannot be so simple as to be or feel trivial or excessively limit the user's choice, however, neither can it be so complex as to overwhelm or give the users so many options that by the time they are done with this part they automatically consider the entry finished or complete. In other words, the goal is to maximize versatility of this input type without negatively impacting the user's willingness to amend the entry further by adding more context through other supported input types.

5.3.2 The meanings behind the current icons and colors

For examples of all the available icons and colors in this version, see Figure 6. The theoretical total number of unique combinations of all icons, colors and shades in this version is 700 but we can get to a more realistic count of 210 easily distinguishable unique icons by roughly dividing the different shades in to three different categories of low, medium and high in the context of the strength or intensity of the emotion.

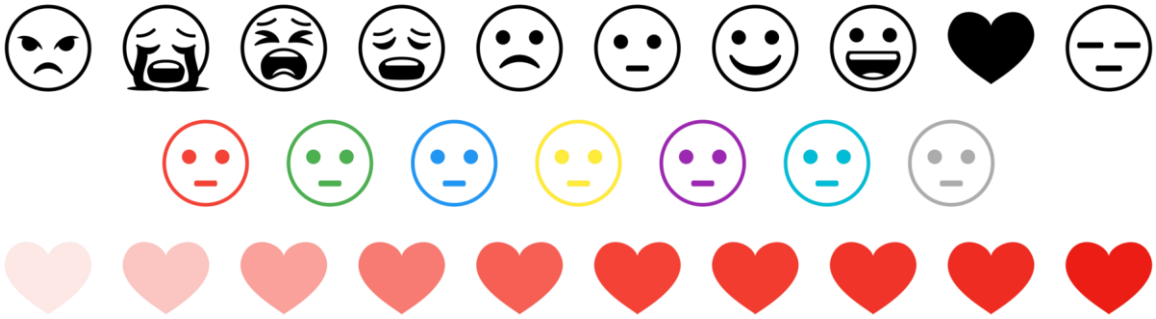

Figure 6 Icons and colors available in this version, top and middle rows respectively, along with the all the available shades of red as an example at the bottom. The original base icons that are manipulated by the application were a part of a set freely available under the terms of CC-BY 4.0 by the EmojiOne project[8].

[8] https://github.com/emojione/emojione/tree/v2.2.7/assets/svg_bw [Apr 10, 2017]

The simple answer to how many different emotions, can then be represented by these 210 icons that are available, based on simple mathematics, would be 70. Here we are simply omitting the strength of the feeling represented by the shading entirely and considering it as a simple modifier not unlike adverbs in a language, such as the words very, moderately and slightly in English for example. However, this assumes the user follows the given instruction precisely when considering the selection of the final shade of their chosen color for the selected icon over what they think about the given shade itself in combination with the icon. In other words, their interpretation of the final icon they come up with may actually be more than the sum of its parts. To understand why this might be the case and why we really should not make general assumptions about the meanings of the icons created by users like this lightly we can look at some previous studies in the field of psychology about how the human mind usually appears to create color preferences and associate colors with different meanings.

Palmer and Schloss have done research into color preferences between individuals, based on their Ecological Valence Theory (Palmer, Schloss 2010) which links the individuals like or dislike towards certain color with the past events and objects associated with it. They propose that even variance in preferences within individuals, i.e. at different times, has similar causes (Schloss, Palmer 2017). While their earlier research focused on object associations excluding abstract associations, such as emotions, they find that colors can have strong abstract associations that can be affected by the cultural context (Schloss, Palmer 2017 cit. Palmer et. al.).

The emphasis here being on multiple meanings, individual and cultural differences. It is all too easy to associate black with sorrow or sadness for example, but that is usually only in the context of mourning, funerary rites or someone passing away in general due to cultural associations. In much the same way it is possible to say that calm skies are blue and thus conclude that blue is a color that represents calmness, however, at the same time we have equally valid examples of cultural association between the color blue and melancholy, sadness or even depression e.g. through the expression “I am feeling blue”.

Similar logic can then be applied to these kinds of simple monochromatic symbols, such as the icons in in Figure 6 (p. 36), available to the user. This works because each symbol only uses a single color emphasizing the associations related to that color's different properties and thus adding more nuance or context to the otherwise basic symbols. Unlike with colors when it comes to the meaning of these specific base icons, however, in the case of this particular set, there actually is a standardized meaning for every one of them by the Unicode Consortium (e.g. Davis, Edberg 2018; and related), such as: expressionless face, weary face and so on[9]. Additionally, for each of those standard descriptions there exists multiple artistic expressions, e.g. the differences between software or platform vendors, of each emoji and different interpretations from users depending on previous use and current context which is easily more significant.

The base icon set is chosen here for its simplified style lacking in other distinct characteristics apart from the use of single color. Which aims to help in de-emphasizing some of these pre-existing meanings, from contexts such as social media, and allowing users to come up with their own as they see the same icon in different colors while making their selections. The resulting icons still retain their role as pictographs but hopefully by taking these steps in terms of aesthetics and allowing for user created variance in terms of color the meanings can remain more fluid. Due to the presented selection order, i.e. normally icon first, the color can be assumed to add nuance to the perceived meaning of the icon, however, if the order was changed or the user chooses to change the icon after selecting a color, then presumably the icon would be adding nuance to the color instead.

From a design point of view, it was desirable to actively look for this kind of ambiguity in order for the emotion input type to remain as versatile as it can, without too much added complexity or introduction of too many options. This is done specifically as a means to introduce an aspect of interaction closer to the *tool* model's ideal of increasing power of expression but without a complex interface. However, in order to better achieve this the application also attempts to take care not to deliberately reinforce meanings for the icons, colors

[9] https://unicode.org/emoji/charts/full-emoji-list.html [Aug 01, 2018]

and their combinations that would originate from the designer through design choices and details. To the same effect the application also has to avoid actively reinforcing meanings from previous instances of the emotion input type being used by the user themselves in so far as that is possible. In practice this results in a very deliberate absence of some convenience features such as saving of presets for example.

It should be noted that in attempting to introduce or maximize potential for varied interpretations by the user there are some inherent risks. As noted previously, the current design already forgoes convenience in some respects. However, this could easily be taken too far. As a concrete example during development it would have been possible to further reinforce the previous goals by introducing a level of randomness to the emotion input type in respect to the default values selected as well as in what order different options are cycled through each time the user initially starts interacting with it, however, this would have come at some expense to usability and in particular could have actively hindered the users' ability to learn to use the emotion input type effectively over time. As such this idea was ultimately discarded during development as counterintuitive on those grounds. The importance of this is highlighted even further when considering the possibility of the emotion input type replacing the old bookmark function for some use cases.

5.3.3 Evaluating alternative layouts for the emotion input type

During the follow-up interview the father brought up an alternative suggestion for the detailed layout and navigation flow of the emotion input type. For reference, see Figure 5 on the left (p. 33), for the current layout. The main difference in the father's suggestion was dividing the current layout into more discreet screens or views with just the current final icon or pictograph as the common element.

From the point of view of interaction, the difference between these two comes down to emphasis and perceived length or complexity. The current layout was chosen during development primarily for two reasons. First, it provides a holistic view of all the different attributes the user can change for the icon at once in a single view. The benefit of this is that the user may indeed change their previous choices easily, as they see the icon change, while

adjusting the different attributes. Second, it reduces the length, in terms of required actions, of creating the icon for the emotion input type by keeping everything as part of a single view and keeping the changes in that view's layout minimal for each action the user does. By comparison a layout where the different attributes the user can change would occupy part of the view one at a time would mean that the user is also effectively directed to focus on one attribute at a time and is thus less likely to change a value of a previous attribute, e.g. color or base icon, after moving past it simply because doing so would take an extra action.

In terms of raw number of actions in the current layout we have three attributes tied to one action each, that may be repeated to cycle through different options. Whereas with the proposed alternate layout we end up with the same three attributes, but each tied to at least two actions, one of which may be repeated to cycle options i.e. the same action as earlier to set the value and another to move to the next attribute. Additionally, for both of these control flows we have the shared action of either saving the entry or adding more information via one of the other entry types, the two options being mutually exclusive as adding more information is simply a superset of saving which moves you directly into creating what is actually just a second independent entry for the same point in time causing them to be grouped together[10].

Considering the previous points, the advantage of the alternative layout proposed comes clear in any instances where, as noted, extra emphasis on either each attribute individually, any accompanying or framing question or the specific order of presentation, and likely of answers, is seen as desired or significant. Notably, as a purely technical advantage, breaking up the view into what are effectively discreet sub views as described would permit a significantly higher number of potential attributes for the user to adjust so as such the alternative approach would be beneficial if additional attributes were to be added later. However, considering the current set of attributes is limited to three there is no obvious advantage in

[10] The grouping behavior of entries can be to a degree dependent on application configuration, which can be set by the study server.

changing the layout to the more complex one described here for the time being. Notwithstanding a more in-depth exploration of which of these two approaches would be likely to produce the best kind of balance between user experience and produced data for any potential research problem. As such currently it is safe to say the choice between these two approaches, from user point of view alone, appears to remain that of personal preference provided that the complexity of the emotion input type does not change.

While the next point of feedback is a more general one and applies to all input types, but especially the audio and picture types, it can also be addressed here when it comes to considerations related to interaction. This is the feedback about adding titles to entries which was suggested and would objectively be beneficial for discovery of previous entries, as noted by the family, due to giving context to an entry at a glance without necessarily having to start reading it or play it back when viewing them retrospectively. However, when it comes to the addition of titles if it is done the most important point to be made about interaction is that the title or caption should always be asked about and added after an entry has been completed and saved rather than during the entry creation process itself. This emphasizes the title being optional and ensures that the presence of a title does not needlessly limit or frame the contents of any entry before it has been created.

Considering the alternative scenario in which the user is asked for a title as the first step when creating an entry. In the worst case, the user might opt to not create the entry at all due to not being able to come up with a suitable title on the spot as it were. However, even in the best-case scenario, when asking the user for a title ahead of time we automatically frame that entry to whatever they chose rather than having them decide on the best title for an already created entry upon saving it.

For additional considerations regarding the addition of titles and captions, and why this is significant for the design aspects of the work described here, see Chapter 7. These have to do with factors related to the study server and its implementation specifically. In the end titles have yet to be added to the version of HAPPY discussed here for those reasons which are not apparent when considering the client application mostly in a vacuum, as separate from the practical data collection aspect of it, as is the case in this chapter.

6. Comparisons with similar software and concepts

This chapter looks at existing software and other related use cases of concepts similar to the ones present in the updated version of HAPPY, with focus on recording feelings and emotions. Leading with a look at a few variants of existing emotion diaries, or mood trackers, and concluding with a look at similar tools development and existing uses of pictographs, in particular emoji or equivalent, in research contexts.

It is not strictly necessary to do a similar comparison separately between the different versions of HAPPY because the newer version is effectively a superset of the original. Meaning that capabilities have not been removed, only added, and the development was thus an incremental iteration on the existing concept as opposed to radical change in direction.

Notably, however, this is not a systematic review, because conducting one would be extremely time consuming and challenging for several reasons. Chief among those reasons the simple volume of digital applications, even when limited to journal or diary style applications, particularly in the consumer space on mobile, that would be relevant in some capacity. On the other hand, many of these applications from the academic side are not easily accessible or likely even in a functional state where fair detailed comparison would be readily possible.

6.1 Comparison with existing emotion diaries and the diary archetype

Unsurprisingly the idea of an emotion diary is not an entirely new innovation. In the mobile space applications like these are sometimes also called mood trackers. To get started, however, we should first consider what makes a diary exactly. The two definitions of interest here are that of the stereotypical or generic diary as an object, e.g. a journal or a notepad with empty pages or an online blog, and the broader definition of a diary as any regularly kept record of contemporary events or experiences regardless of whether they are public or personal, freeform or structured (Alaszewski 2011).

The idea of the general-purpose freeform diary, which is not tailored for any specific kind of entries, intuitively would appear less common in the context of academic research as

such diaries are generated as an artifact of research for analysis regarding specific research questions. However, according to Alaszewski, the emphasis on structure can be the deliberate result of trying to avoid non-relevant data or ease analysis while a less structured format can produce more natural data not burdened by the research process itself making diaries a versatile format depending on specific goals (Alaszewski 2011 p. 112).

For example, in the demonstrated use cases of the original version HAPPY the diary still manifested as fairly structured due to the classification of entries either as “closeness” or “separation”, which are mutually exclusive, even if audio recording as a chosen format is very open ended otherwise. For the updated version, on the other hand, the emotion input type can be used, both to create similar added structure or simply to act as prompt or a starting point for the user or participant which could be compulsory or optional. If configured as compulsory part of all entries, which was one a possibility discussed during the second interview, then the user has significantly more structure because of this. However, the setup chosen for the test of this version, as described in Section 5.2 (Figure 4 p. 32), deliberately avoided this because the research questions here are related to the application itself and its use so any unexpected ways to use and combine the provided features were seen as a positive. Additionally, considering the user centric approach to the design the relevancy of an entry is dictated simply by the fact that the user chose to create it in the first place. Whether the entries adhere to a specific format or concerns a certain topic exclusively is as a result somewhat less important in this specific case.

6.1.1 Example: pen and paper emotion diary

Before taking a closer look existing offerings of other emotion diaries in the digital space it is important to note that the same thing can also be done easily without a digital solution. The analog emotion diary chosen here as an example is specifically published by THL (Finland’s National Institute for Health and Welfare), for use as an exercise in student mental healthcare, as a part of a guide on treatment models and methods intended as examples for staff and persons in charge (Haravuori, Muinonen et al. 2017). The actual template consists of three independent exercises in Finnish. The mood diary portion is the second exercise (Haravuori, Muinonen et al. 2017 p. 232) and it provides the student with a table to fill in

that has columns for: “Date”, “Situation, Event, Activity” and “Strength of the feeling (0-10)”. The scale for the last column is additionally defined as just: “(0 = nonexistent, 10 = strong)”. The inclusion of 0 on the strength scale is worth of note, because it in theory suggests also recording the absence of specific emotions, for example, perhaps those present previously in the same situation. However, the same table also comes pre-filled with one example day and in that example the feeling column contains multiple feelings or moods for each situation, e.g. “Happiness, uncertainty” but only one number in the third column that presumably corresponds to both emotions.

Unfortunately, the previous publication does not go into too much detail about how to apply this particular exercise type, so we can only guess whether the idea of also noting absence of expected emotions or moods was specifically intended by the authors or not. Regardless of their intent, however, it is at least something the new version of HAPPY is unable to represent in an explicit way using the new emotion input type. On the other hand, an analog diary like this obviously loses the benefit of being literally in the palm of your hand which a solution based on a mobile application has simply due to the platform it is on.

6.1.2 Example: digital diary and lifestyle application with emotion focus

On the digital side for consumer applications, the main challenge is not a lack of potential applications to choose from and compare, as noted, but rather finding a point of comparison that has been through at least some level of real verifiable scrutiny. However, as the barrier to entry in the mobile application market is low so the best metrics available in this regard are number of installs and user reviews.

For this application the one chosen is a lifestyle application Daylio[11] described as a mood tracker and a micro-diary, by its developer, with a byline of being able to do all of this without having to write a single line. Judging by the public metrics this seems to resonate with Android users, with an average rating of 4.8 out of 5 based on ca. 180,000 reviews and

[11] https://play.google.com/store/apps/details?id=net.daylio [Sept 21, 2018]

over 5 million recorded installs on Google Play in the three years since its initial release in August 2015 (statistic as of September 2018).

In terms of the application's actual feature set, the main function remains quite simple: the ability to create timestamped entries that consists of an emotion pictograph, with a customizable label, combined with a set of activities, from an editable list of presets, and an optional line of text (Figure 7).

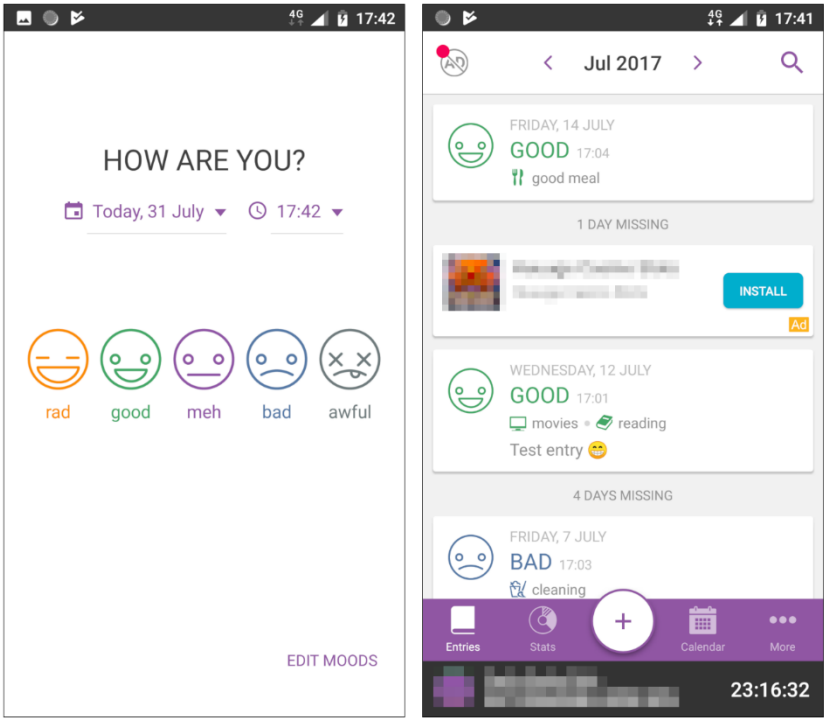

Figure 7 The Daylio application with the view for creating new entry on the left (using default values) and the listing for the day's entries (main view) on the right. When creating an entry, the selection of activities would happen in between these two views.

At first glance the basic idea seems similar to the emotion input type of the new version of HAPPY. However, the questions the user gets asked are simpler and quite different, e.g. "How are you?" and "What have you been upto?" While the initial focus is on the user's current emotion or mood the focus is then immediately shifted to something less abstract. Additionally, the default emotion options presented are clearly, if sometimes colloquially, labeled. As for the use of colors, unfortunately, upon deeper inspection they are just visual

flare turned revenue stream[12]. That is to say, as far as the colors are concerned, they are not a recorded part of the entries but a part of a theme that the user can change at any time in the application settings also changing the colors for any previous entries retroactively.

In terms of auxiliary features Daylio offers visualization of the data over time in different ways and some gamified aspects, such as achievements and streaks, which here are in practice simply digital badges for doing specific things and a record of most continuous days in a row of creating at least an entry. Additionally, it has the ability to search for specific types of historical entries. As noted in Chapter 2, gamification serves as a means of introducing additional extrinsic motivation for the user. However, in this case the implementation is subjectively so basic that it appears, for the lack of a better word, textbook like. This is as a reference to the fact that the gamified elements chosen to appear here appear quite superficial to the use of the application and also somewhat generic. Some of them are also direct implementations of common examples that are sometimes used when discussing the concept of gamification itself, which definitely contributes to the previous impression.

Likewise, the provided graphs and other statistics are also quite basic and not transformative in any significant way, based purely on quantity over time. This on its own is non-issue, however, the application does appear to do some kind of averaging of multiple recorded moods for a single day when it has to decide on a color for that day when looking at data in terms of just days, e.g. on a calendar view. As a result, considering the moods user can select from are completely malleable by them, it is uncertain as to what the logistics of this are and it appears more random than anything else. Especially considering the colors used to represent everything in the app, including these visualizations, can be changed by the user at any time if they change their theme. Regardless, it is undeniable that Daylio is successful in its own category of apps, appearing frequently on the top grossing apps on Google Play under the Lifestyle category. Likewise, it has certain features that in the long term the

[12] The application uses a freemium business model where the application itself is free but additional features, in this case more color sets among several others, can be unlocked at a cost via an in-app purchase (for more on the freemium model, see e.g. Seufert 2014).

new version of HAPPY would be likely to benefit from. Such as in particular the ability to search for different types of entries and offer a overviews as visual graphs, however, for the latter it is important to note that a graph or visualization in itself does not add much value unless the user is able to interpret its contents on their own without the application telling them what they should think. While techniques such as artificial intelligence, in particular e.g. machine learning, could be used to facilitate deeper analysis of the data for the end user the ever-present risk here is the application appearing smarter than it actually is which is an unfavorable scenario, as discussed in Chapter 2, particularly when considering the contexts HAPPY is expected to be used in.

6.1.3 Some other mood tracking applications briefly

While the previous section focuses on a single application, it should be noted that, ostensibly, many of the different mood trackers available are not that different from Daylio. The differences are mostly in the minutiae of creating entries or how they are visualized, graphs being a common auxiliary feature. However, some additional examples are briefly given here for going beyond the basic concept of mood tracking as previously discussed.

The first of these is Moodtrack Social Diary[13] for turning the act of tracking moods into a social activity through open sharing of entries with other users. However, due to the choice of business model here, privacy has become a commodity, in so far as the visibility of the user's entries to others is concerned, even though the application is otherwise anonymous.

The second example is aiMei[14], for combining a mood tracker with a facsimile of a digital assistant, whose role is primarily to provide reminders and onboarding with using the application as well as administer a set of fixed personality tests. It also allows the user to set a goal to areas measured by these personality tests and see if it has been reached upon retaking the test later.

[13] https://play.google.com/store/apps/details?id=com.moodtrak.diary [Sept 10, 2019]
[14] https://play.google.com/store/apps/details?id=bpu.aimei [Sept 10, 2019]

6.2 Digital diaries and data collection in research contexts

For the academic side of digital applications, an in-depth first-hand review of a single, or even couple of applications would be preferred. However, as noted earlier, this is generally not possible because many of these applications are either not readily accessible for easy first-hand evaluation or even necessarily in a functional state. Therefore, what is left are second-hand accounts, in the form of published articles, that detail the development and use of new tools for research taking advantage of the digital and mobile space. Additionally, in this instance attempting to focus just on digital diaries that deal with emotions specifically proved too narrow of a focus. To address this additional problem a separate look at how emotions and their depictions, in particular the related use of pictographs, have been previously measured and used more broadly in research proved necessary (see Section 6.3) and here the focus is more on data collection using mobile or smartphone as a platform.

For an overview of different approaches available and considerations involved in creating digital systems that collect data from research participants, purely from researchers' point of view, of particular interest are articles dealing with the development of tools for ESM/EMA or similar research methodologies due to the noted similarities with both versions of HAPPY. One such article by Hofmann and Patel (2015) focuses specifically on their efforts to refine the popular approach of using SMS text messages as a signaling or reminder mechanism. As part of their article they also provide an excellent comparison of different approaches that is largely methodology agnostic (Hofmann, Patel 2015 p. 237-239; especially Table 1 p. 238). While the following might be a bit of an oversimplification, it is easy to identify few compelling main points from their comparison: the importance of the idea of function over form, reusing existing technologies and arguably most importantly finding the lowest common denominator, as it were, while keeping any costs for both the researcher and respondents at minimum.

Additionally, in regard to the point about having a separate device to use, as was with the testing of the new version of HAPPY, being a negative is one Hofmann and Patel (2015) also make and that is in line with the feedback gathered during the follow-up interview. Therefore, it is important that when the dedicated application approach is considered for tools

development the end goal is always to make the application such that it can be installed and left on the participants' personal devices. Ideally, the application should also have some inherent value outside just the base functionality of collecting data for research, so that it might be useful for participants even after the research has concluded.

Considering the mobile phone as a device the features they have that are device and network agnostic, regardless of carrier, are obviously SMS text messages and voice calls and for smartphones specifically the ability to browse the internet in some way. This common set of features is thus the absolute lowest common denominator referenced earlier which is important for a researcher to maximize the number of potential participants. However, the obvious downsides include a fragmented user experience, as rather than one system multiple independent systems are effectively combined, such as in the case with Hofmann and Patel SMS messages and online survey platforms, and the user has to navigate through them. Additionally, with SMS in particular there are significant limitations to the kind of entries or content that it can handle, and these limitations may vary from device to device.

However, arguably this long standing lowest common denominator has begun to change in recent years with the availability of ever cheaper entry level devices, developing technology and increasing coverage. Naturally, there are still significant differences in user penetration between different geographical areas when it comes to phones and smartphones especially. However, globally as of 2019, almost half of the total population are connected to the mobile internet, with smartphone adoption of around 60%, and in many parts of the world, such as e.g. in Europe and North America, the numbers are expectedly significantly higher than that and these numbers are projected to keep growing worldwide (GSMA 2019). Additionally, the obvious homogenization of the software platforms used by most of these devices, i.e. either Google's Android or Apple's iOS, is making the approach of using a dedicated application more appealing[15]. This is particularly true in such cases as with HAPPY when the goal is to increase the value to the participant also during the participation rather

[15] Depending on source between 75-85% percent of devices are on Android with generally less than 1% listed as others and the rest as iOS (IDC 2018, StatCounter 2019).

than only after the research has concluded and the value is evident in the form of results and findings.

6.3 Pictographs (emoji) and emotions in research contexts

In the earlier chapters, the focus was largely on the emotion input type and what it produces from the point of view of the user and in particular the interaction that can happen and one would hope to happen when the user uses this new input type, i.e. the design thinking and rationale. However, it is also necessary to consider how similar symbolic representations, such as emoji in their more static forms, have seen previous use in research contexts.

To start off with a simple example from an everyday context, that also came up in the initial interview with the parents, Figure 8 shows the use of emotion related pictographs in commerce analytics as data being gathered by stores. The stores may use machines, not unlike the one pictured, with a selection of buttons for collecting customer feedback e.g. about the quality of cashier service in stores or fast food restaurants.

Figure 8 A customer feedback terminal from a retail store using pictographs for input. The terminal itself is a service provided by a third-party company. User feedback to an input is dim light at the top of the unit.

The idea of use and integration of emoji into scholarly research methodologies has naturally also seen an uptake in academic circles, due to the significant presence of social

media sites and services that use these icons, in recent years (Kaye, Malone et al. 2017). This seems a particularly relevant topic when considering the methodology in studies involving younger children (e.g. Fane, MacDougall et al. 2018, Schouteten, Verwaeren et al. 2018, Swaney-Stueve, Jepsen et al. 2018). When considering emoji use in research with adults a study found that there are not many significant differences in peoples' interpretation of commonly used emoji due to: age (=18-60), frequency of use or gender; but that some may still exist on a case by case basis (Jaeger, Xia et al. 2018). However, Jaeger, Xia et al. also duly note that their work does not give special consideration to cultural differences.

As the previous examples show, the ability to find such common meanings for some emoji, which researchers can now take advantage of, have extended the possibilities of using them in research. This could afford researchers the possibility of replacing some existing verbal or older pictorial scales with ones based on curated sets of emoji (Kaye, Malone et al. 2017, Swaney-Stueve, Jepsen et al. 2018).

Commonality between all of the previous research presented here, however, is that they do not consider the possible subtle differences in people interpreting the meaning of an emoji differently based on aesthetic choices made by the artist, i.e. that the meanings attached to an emoji are not verified against multiple artistic representations of it. For example, this could prove significant if labels are not used or the question does not guide towards a specific kind of interpretation such as e.g. the difference between a feeling and an attitude. Although conversely it should also be noted that something like this is unlikely to be of significance particularly if the emoji are part of a closed set or ordered scale as is the case in many of the examples here. The same can be said about any CATA/RATA surveys, assuming the emoji related to a given question are displayed to the respondent all at once. After all, in both cases the reasonable expectation would be that the respondent is predisposed to treat the presented emoji as a related set in the context of the given question.

By comparison the new version of HAPPY, with its emotion input type, does as much as it can to make those subtle differences between individuals and different instances of the same icon matter more rather than less. So, in some sense what many of these previous researchers worked to find, refine and verify is at odds with what the new version with its

current implementation is hoping to allow the user to do. Rather than thinking of the potential different interpretations as a problem or a challenge we can also afford think of them as an opportunity for the user to do more with less. We should not, or could not, use or need tens of completely distinct icons to represent equal number of different moods or feelings with various subtleties for example. At the same time, we now have a way to avoid overwhelming the user with too many options at once by breaking down the different possibilities to more manageable sets.

Of course, in the end there can be no guarantees that the users will think about their choices in a way that results in significant variance between the interpretation of an icon and color combinations between different users. However, the application gives them the affordance to do so if they wish to give their choices a deeper meaning and those meanings can change from instance to instance because there is possibility for significant variance in color even if the base icon is the same.

6.4 The preliminary impact hypothesis for using the new HAPPY

The goal of these relatively short looks at related concepts and software is to gather some preliminary insights on whether the results of the previous user-centric design thinking, with the aim of giving the user more value and expressive freedom, is in theory usable in a research context. If so, in what role and what considerations may need to be taken before possibly conducting a new pilot and/or feasibility study with this new version of HAPPY to bring the three cycle model of design research to its conclusion, i.e. effectively closing the relevance cycle (Hevner 2007).

First of the two main insights we can come to here is that the new emotion input type could in theory be used, as is, if the researcher considers the produced pictograph to simply be an optional visual aid or a prompt for the participant (user) and looks for any deeper meaning and context from the entries combined with and added to it. This leaves the researcher with the burden of instructing participants to try to avoid the use of the emotion input type without adding any additional information, e.g. audio or text, to it unless it is used with the explicit intent to add some later, i.e. as with the bookmark use case with the original

version. However, even in this first scenario the produced pictograph is not without value to the researcher because it can still be used for anonymized visualization of data from single participant over time and it may be possible to identify patterns from it.

The power of visualization was also discussed during the interviews in the co-design phase, brought up in particular by the father, as a result supporting this capability in some way both in the user application and perhaps the study server (see Chapter 7) would be a logical next step in terms of how to get most out of the new version in the different contexts. However, as noted earlier, the value in such visualizations is inherently dependent on the ability to read them and draw conclusions which is something that may be difficult for participants to do alone. Because of this it may be tempting to consider using e.g. various aspects of color theory and the models developed for predicting and measuring color preference and harmony (e.g. Ou, Luo et al. 2004) as tools for some automated analysis such as through e.g. predictive rating or grouping. However, as noted by Weingerl and Javoršek (2018), even when using much more recent and complex models based on significantly larger sets of data, generalizing from the results of such models has some risks and may not be suitable for all contexts. As such, at best a coarse rating or high-level grouping could be produced, based on color in this case, that would still have to be independently verified for best accuracy based on other available information. While this kind of data may prove interesting for future research or valuable to go over with a professional, as specifically noted during the interviews in regard to even just simple visualization over time, there is no obvious benefit to providing too much raw data to the participant directly without much more scrutiny and further work on how it could best be presented and what effect it should have in the user application specifically to avoid the risk of the user drawing wrong conclusions because of presentation. This also highlights that placing too much weight on just one aspect of an entry can be problematic in this context.

The second possibility is simply curating the set of emoji, much like the researchers from some of the earlier examples have done, and potentially labelling them or limiting color options to allow more explicit universal meaning or meanings to be attached to each pictograph. This can and should in this case be reinforced by making the labels visible to the

participant. Thus, allowing each pictograph to have a meaning such that they could, if necessary, stand on their own and be more easily analyzed. The main downside of this approach is that it would simplify the interaction opportunities the participant has with the application in a way that does not necessarily provide significant gains in ease of use or usability for the participant in the moment. In other words, in terms of interaction design and the models of interaction covered in Chapter 2 it would simplify the user interaction back closer to the *machine* model by limiting the power of the interactor in exchange for allowing the pictographs and the related recorded metrics to be more directly analyzed. In this case the alternative layout discussed in Section 5.3.3 could be beneficial especially if some elements of the emotion input type should be specifically emphasized or weighted differently as the current design considers the different attributes that make the pictograph equal, or with slight emphasis on the base icon at most, through its presentation.

Both ways to use, or refine if necessary, the new version of HAPPY have their merits when considering the use of this version for research purposes. However, in the context of this thesis and in the domain of interaction design specifically the latter has some significant downsides from the users' point of view. The former on the other hand would probably require follow up interviews with the participants, after a data collection period, to be most effective. This would also be in line with idea of having some concrete recipient that came up during the interviews with the participating parents on more than one occasion in different contexts. Based on that discussion we can assume that having such a follow up interview or arranging for another way to feel like there is a clear recipient, another type of response or more generally some kind of human element, might be an extra motivator for some people. How this would be easiest to realize in practice is somewhat out of scope of this thesis, however, one possible alternative to a researcher organized interview, which was brought up during the interviews, is to better integrate the application with the routines of the NICU such as e.g. the weekly meetings. The feasibility of this being dependent on co-operation from the hospital naturally and any potential privacy concerns they might have about the software.

Additionally, irrespective of any potential decisions made regarding the two scenarios presented here, the obvious opportunity to include the remaining element from ESM, i.e. a

signaling mechanism, should not be excluded. This means that the participants would still retain the ability to record experiences whenever they choose, providing immediate value for the participant in the moment, but those experiences would be complemented by others recorded as a result of engaging the participant through signaling them at appropriate points decided on by the researcher. The signaling could be e.g. time or location based or any combination the two, considering the affordances of the common smartphone as a platform.

7. The updated study server and client refinements

This chapter will mainly focus on the data collection or study server aspect of the new version of HAPPY and how it influenced the realization of the design ideas and concepts, as presented in previous chapters. Additionally, any functional changes and behaviors previously referenced, but not discussed in-depth, in those chapters will be included as part of this chapter.

These additional changes include client functionality and interface elements directly related to the study server. There are also some behaviors that are otherwise transparent or unable to be tested during the normal operation of the client application, such as in the context of the previous co-design phase and its associated testing period, which are likewise addressed here.

7.1 Client and server overview

While the previous chapters have viewed the client application as a largely independent entity this section will relate those chapters to the final distributed system where the study server is also present. Naturally, the need to support data collection as an eventual use case was a persistent concern during development. However, due to practical reasons noted in Chapter 4, there was no need to support remote data collection during the co-design phase itself. For this reason, a positive side-effect of minimizing the dependencies between server and client in the completed system was achieved effectively at no extra cost, because the lack of network connection was the presumed default state during design and development of the client itself.

Furthermore, in a situation like this where the starting point for the project is an existing software solution containing an older version of both client and server components, the obvious consideration early on had to be whether to start from scratch or iterate on the existing components. For the client application the choice to use the original version as a starting point was an easy one because one of the main goals was to add additional value, so it would not have been that intuitive to start by tearing everything down. This was of course also done

in order to expedite progress early on, so that no time would have to be spent initially completely re-creating user facing functionality that already existed in some form or another in the pre-existing client.

The previous reasons do not generally apply for the server component, however, because it could be developed entirely after the client design was created and implemented. In the end the opposite choice proved most beneficial for the study server for mostly technical reasons related to futureproofing the server component as well as allowing the client to evolve freely during the co-design process itself. However, it should also be noted that early on maintaining backwards compatibility with the original server as an option was considered, even though that eventually changed once work on the server began in earnest, due to potential security considerations. Briefly, taking any measures to improve security and integrity of the server and its communication with the client would be a pointless exercise if a compatibility layer could be used to partially bypass any such measures.

7.2 The study server and remote client configuration

This section will detail how the study server can influence the client as well as the general functionality of the server itself. The original server component for the old version of HAPPY was used as a reference in so far as being able to guarantee feature parity between the old and new server implementations. Making the new server component essentially a superset of the old one in terms of features, although not backwards compatible with the original version of HAPPY as previously noted.

7.2.1 User interface for the new server

Figure 9 shows the interface of the new server after authentication with full privileges. Notably the new server allows optionally exporting the collected data with text-based files being wrapped in HTML for consistent display of special characters, including international characters and especially emoji. There is also an option for exporting entries with times adjusted to the server time, rather than the individual clients' time zones, this option is a by-product of addressing edge cases related to dealing with daylight savings times and potential

entry collisions due to changes in time zone, e.g. due to travel. The last functional changes apparent from the interface alone are the option for using preset configurations, which was added to expedite testing by making creation of new Research IDs (accounts) take less time, and the functionality to send push messages as notifications to clients. The latter of the two was added to provide an experimental implementation for a signaling mechanism as mentioned in Section 6.4 previously, although in this case there is deliberately no automation in its current form since a third-party service is used for message delivery.

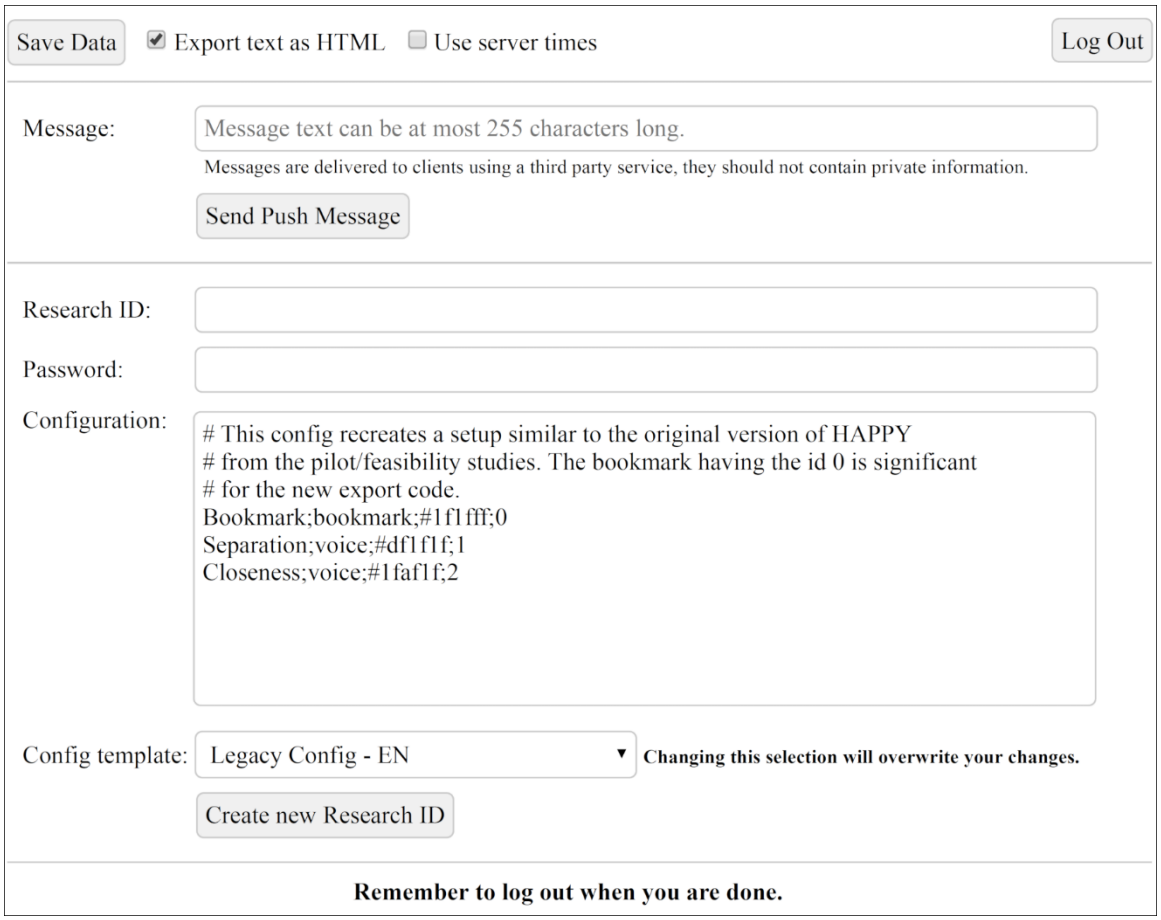

Figure 9 New server interface, after logging in with the master ID.

The ability to create new accounts depicted in Figure 9 is exclusive to the first initially created Research ID that acts as the master account in the system, this behavior is functionally identical with the old server, however, the authentication and other implementation aspects of it were entirely redone. For regular accounts the only actions available are saving of the

data, with the related options, and logging out. Notably the new server implementation requires users to actively log out of the system, whereas with the original closing the browser was the way to achieve the same effect. The former is the more typical approach for web applications that require user authentication. This is preferred because it requires deliberate contextual action by the user, so the state in which the system is left in is always explicit when the user leaves it and allows system tasks to be performed when this happens.

As for the process of exporting the collected data it was changed to be a fully asynchronous process, as opposed to being attempted during a single HTTP request as with the old server, which allows larger datasets to be exported correctly. This became increasingly relevant with the addition of photos as a possible input type for clients which can quickly increase the total size of an export. The web interface will report the progress of a running export operation relative to the number of the total files processed so far.

7.2.2 Server options for changing client behavior

Previous chapters reference the server's ability to affect client behavior and in this section the ways in which this is possible with the new server are discussed. Listing 1 shows an annotated format description of the configuration passed to and parsed by each client when registering or checking in with the server when the client application starts.

```
# Lines beginning with a hash (pound) character like this are comments, which are not
# sent to clients.

# The following line will set a configuration option "key" to value "value".
value;key

# The next non-comment line will create an item in the input selection dialog, where:
#   label = any string literal, used as item label in the client interface
#   type  = string literal: (voice | text | picture | emotion | bookmark)
#   color = string literal: (primary | primary_complement), or any valid HTML
#           hexadecimal color value (e.g. #000000)
#   tag   = integer >= 0
label;type;color;tag
```

Listing 1 Annotated listing of the configuration format parsed by the client.

The additional literals listed, `primary` and `primary_complement`, for defining the item color are constants that the client will map to its locally defined primary interface

color and its complement respectively, i.e. for the current client these would map to the specific shades of orange and blue unless changed in later versions.

All but one of the available input type options in Listing 1 have been previously covered. However, as for the remaining legacy `bookmark` type (Figure 10) which can now effectively be superseded by the new emotion or picture input types, if it is still used a client should, and the current implementation does, create a hidden text file for the server that notes a bookmark having been created at a specific time as part of its entry. Additionally, from user point of view the bookmark input type also still requires the use of the clock dial in the main view of the client application in order to navigate between different bookmarks as they have no other user visible presence on client-side.

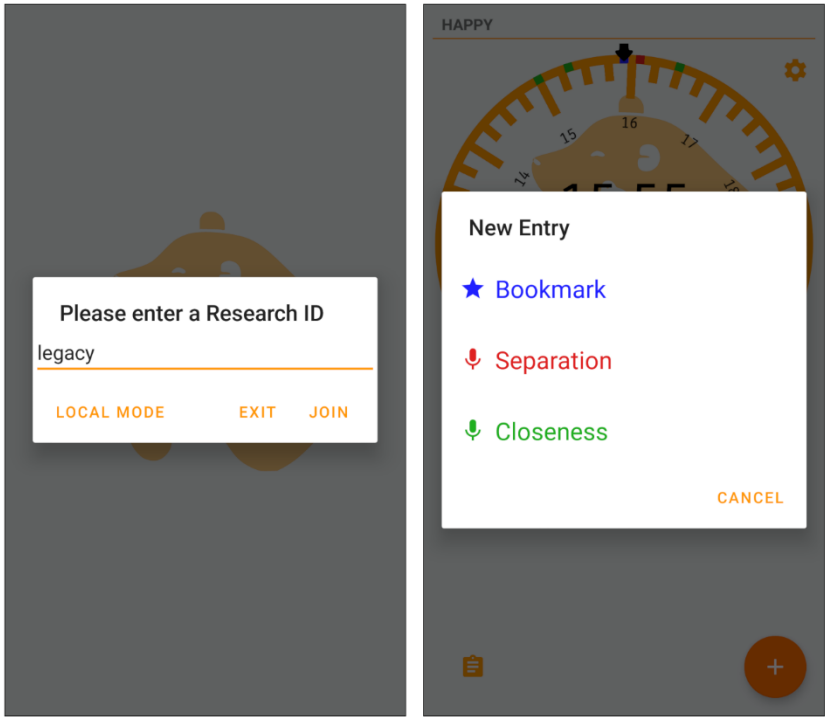

Figure 10 The original research setup recreated with the new version, left screen shows the simplistic server registration dialog with an option to set the application for use locally in "local mode".

So, while the old-style bookmark remains a supported option, and addresses the original problem with it noted in Chapter 3 in this way, using it still cannot be recommended when better alternatives now also exist. This is likewise indirectly supported by the feedback from the interviews where the necessity to directly manipulate the clock dial was brought up as a

negative element in terms of usability which is something the plain bookmark input type as it exists still requires and would bring back with it as a hard requirement if it were to be used.

The client also supports several new options that the server can set through the key, value pairs (see Listing 1) which are documented in Table 1. Notably these options include some that control how the client interacts with the device camera and upper limits on entry lengths for specific types. The option to disable the use of camera completely is particularly beneficial in instances where emotion input type is used but the inclusion of photos through it is not desirable. It can also help with low memory devices in general or devices on metered connections, where the option to scale down pictures might not produce desired results, if it is known that all or most active clients are in this type of a situation remotely disabling the applications camera functionality is worth considering.

Table 1 Additional client configuration options

Option (key)	Description	Default Value
compress_photos	Suggestion for clients to scale down camera images. *Note: the precise level of scaling is up to clients.*	0 (false)
disable_camera	Forces the device to behave like it had no camera. *Note: emotion input provides access to all input types.*	0 (false)
audio_length	Limits the maximum continuous audio recording time in seconds.	600
text_length	Limits the maximum number of characters in text entries seconds.	2000

Finally, as noted in previous chapters, if there is only one input type configured by the server this effectively skips the input type selection in the client when adding entries. However, in this instance that effectively also means the provided label will not be visible to users. Additionally, multiple input type definitions with an identical or different `tag` number will mainly affect how entries are grouped for exports from the server. However, in some edge cases, the `tag` can also affect how entries are grouped when viewed through the client. Specifically, on client-side, the `tag` does matter when two entries with different `tag` values are created for the exact same point in time.

For example, the configuration which recreates a comparable setup to the one used in the previous studies with the original version of HAPPY (Figure 9, p 58) with this new version makes use of this special case. This is done in order to keep the files automatically generated for bookmarks separate from real user entries while making sure that the client will not allow bookmarks to contain other files thus guaranteeing that the user created entries for those bookmarks are self-contained in server exports. On the other hand, the default all-encompassing configuration used for the test during the co-design phase did only use a single `tag` value internally, however, due to the local nature of the test itself this did not have any user visible impact.

7.2.3 Server impact on implemented client features

This section will give an overview regarding the absence or omission of certain features discussed in previous chapters, namely titles for entries and automated transcription, from the version of HAPPY presented here as it relates to the study server. The former ended up not being implemented because it was desirable for the implementation of the server to provide similar or better guarantees than the original study server regarding how the data is stored on it[16]. The latter, on the other hand, ended out of scope because it does not interface with the end user of the client directly, in the context of the discussions had during the co-design phase, the client having no feature that would directly require it.

Regarding titles specifically, implementing them on client-side would be quite simple for the current version of HAPPY, so in this case the study server's existence added, in regard to interaction and usability, additional complexity because it made addressing gathered feedback more difficult. This is also the case for looking up past entries, because of the flexibility afforded by the study server. Although, while under specific configuration, as was also shown and discussed during the follow-up interview a common element that would be usable for looking up past entries can be guaranteed this guarantee is not universal. After all the server

[16] Briefly, all user generated content needs to be encrypted and stored separately, behind different authentication schemes (file system and database access), to the cryptographic keys required for decryption.

may currently configure clients in a way where this guarantee does not exist. Resolving the technical problem of how the study server should handle user generated metadata for files of arbitrary type[17] would remove the blocking issue on titles and allow them to act as this type of universal element, even if an empty title should be considered valid, regardless of server configuration.

The technical documentation related to the practical work will go into more detail regarding titles, transcription and the limitations of the study server. However, in the interest of the subject of this thesis, presuming the blocking issue was resolved, Section 5.3.3 has already addressed the issue of titles in terms of interaction and placement in the flow of using the client application. It should be noted that while the emotion input type, ostensibly, already stores additional metadata for an entry the approach taken for it is not something that should be generalized for other input types because it is quite inflexible.

7.3 Addressing accessibility concerns in the client

This section briefly evaluates accessibility of the updated client. This was previously referenced in a Footnote in Section 5.1 (p. 31). Specifically, the new client ensures that basic accessibility is addressed, in so far as ensuring the interface remains functional and logical when it comes to using the accessibility services provided by the operating system. This is done by providing accessibility descriptions or labels for all interface elements. However, the emphasis with what is provided is on it being at minimum functional, as opposed to necessarily usable or user friendly, level and these two main points are what will be discussed here further.

Specifically, the most significant reason for why this kind of consideration is needed is to highlight the downside of such designs that rely so heavily on visual feedback, without an alternative, and indeed the concept of color especially in this case. The emotion input type in particular is the obvious pain point here, because if we remove the visual element from it,

[17] The case of storing additional metadata for text-based files being a trivial one.

which would be the case e.g. for most people using screen readers, then it becomes just a short survey with two or three questions[18] and the entire case being made earlier, chapters 5 and 6, effectively ceases to apply in many respects (see specifically sections 5.3.2 and 6.3, as well as the related points in Section 6.4).

The only solution in the short-term to this issue was to change it so that for the emotion input type each base icon had a defined meaning attached to them for accessibility services only and the question of color could be set as blank, i.e. not applicable, if necessary. However, as is proposed in earlier chapters doing this will significantly limit the breadth of options available for the user and this is even more pronounced when the visual element of the choices made being interconnected, by their combined influence on the final pictograph, is missing entirely.

Additionally, there is an obvious case to be made for people who perceive colors differently, e.g. because of different varieties of color blindness for instance. In these cases, there are two things that need to be considered. Namely whether the user is using one of the color correction modes provided by the operating system or not and whether the set of color options given to the user should change and should the previous point about color correction modes be considered separately or together with this second point. In any case it would certainly have an impact on any potential conclusions drawn both by the user and any drawn by a likely researcher from the raw data about user selections. This in the short-term will place even higher importance on the contextual information given by any entries amending the emotion input type, as discussed in Section 6.4.

7.4 Additional client interface elements

As the development for the new version of the client and the associated server component wrapped up notably one additional interface element was added to the client. This is pictured in Figure 11 (p. 65). The additional view, to a degree, does address the point made

[18] Depending on whether the user has any concept of color at all, e.g. the difference between someone who has been blind part of their life and who has never experienced colors at all.

in Chapter 5 regarding the clock dial taking proportionally too much screen real estate for the actual information it delivers to the user. It does this by demonstrating how additional controls can be placed into the main view and highlighting at least two previously underused areas of free screen space. However, it is worth noting that these areas have different properties in the sense that one is much harder to access than the other in some situations. This may or may not be desirable, more on this below.

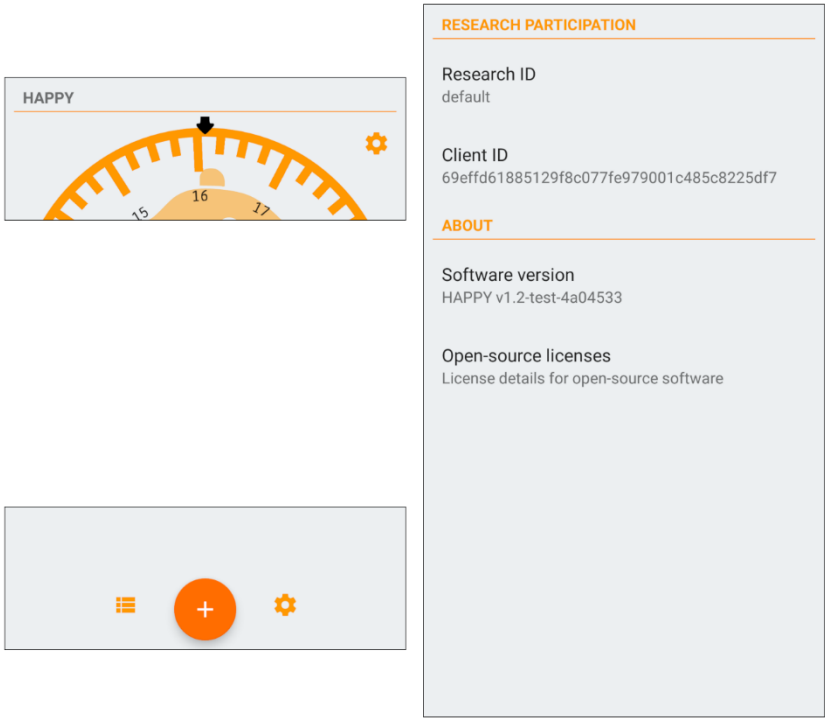

Figure 11 Pictures of the current informational settings view (right) with two potential placements for the related button (cogwheel icon) in the application main view shown on the left.

However, the added settings view (Figure 11, right) currently only used to deliver secondary information to the user about the state of the application such as: whether there is currently an active link with a server or not, the current application version and user readable copies of related open source licenses[19] and any required attributions. Additionally, there are two viable options presented in Figure 11 about where the access to this new part of the

[19] These are distribution licenses, so an end user does not need to agree to them explicitly, however, many of such licenses carry a requirement of being distributed with copies of the application. The approach taken here is a common one for applications on Android.

interface should be placed in the main view, the placement shown top left of Figure 11 is the one in use with the current version. After all, there really is nothing in the settings view currently that would be of significant interest to the user during regular usage of the application so the more central positioning of the second option (Figure 11, bottom left) is hardly warranted on that basis.

This new view does, however, now offer a pre-existing place for any user configurable options for the client added in the future such as possibly e.g. the option to wait for wireless network being available, or more generally to not use a metered connection, to send files to the server. The previous example being one rather common convenience feature for applications that use network connections but do not constantly require one to function, which is true for the updated version of HAPPY as well. If this or other such configuration options are added later, then the recommended position for the navigation elements in main view would be at the bottom center edge of the screen in portrait mode, as shown with the second option depicted in Figure 11 (p. 65).

Typically, the kind of information included in this new view as well as any user configurable options, in an Android application would be accessed through toolbar and e.g. an off-canvas pop out menu, however, due to the space constraints set by the clock dial this would not be a viable approach on most devices. This is the case specifically when used in the landscape orientation where the available screen space can be even more limited (see Figure 3 p. 30, right screen). While it would be possible to lock the orientation to portrait, as indeed the interface is first designed to be used in portrait mode, it is always preferable to allow the user to choose the orientation freely. This is the case particularly whenever text input is involved so that the user may opt for landscape mode in case that is their preferred typing experience using a touch screen.

8. Conclusions and discussion

This thesis has presented the evolution of a new version of the HAPPY application and related work, such as the study server, based on the theoretical concepts of design science research and practical considerations related to the software development process itself. While some previous chapters, such as sections 6.4 and 7.3, implicitly contain certain conclusions and discussion of a more practical nature, including forward-looking statements and future work, both are considered self-contained parts of the literature review and the analysis of the HAPPY system.

Foundationally the work done focused on two of Hevner's (2007) three cycles, specifically the design and rigor cycles with significant focus on the former, with arguably a quite limited form of the relevance cycle present, if there is one at all, because the field testing done was exclusively for the purposes of using co-design as a major part of the design process inside the design cycle itself. In order to present the work done here, the literature review was conducted in two parts quite literally on either side of the design cycle.

8.1 Interaction design

In Chapter 2, the first part of the literature review, Murray's models of interactivity were defined, and in parts redefined, to provide us with a means to evaluate the practical work. The redefinition in the context of this thesis refers to the deliberate exclusion of one of the models and the increased emphasis on both designer and user intent over the aspects of the detailed interaction process or specific actions taken by the interactor. In the second part of the literature review, in Chapter 6, the focus was on comparing the results of the design cycle with previous work by other people. It is easy to identify this as one of the halves as it were, or grounding as Hevner names it, in the rigor cycle.

Based on this we find that the models can be both highly context and inter-dependent and propose the *machine* model as a sort of default model of interaction for many digital systems and artifacts. This is quite intuitive, and most likely not that novel of a conclusion, because the platforms on which many digital interaction systems exist on are often physically

or semantically machines themselves, or at the very least require a degree of abstraction, and as such expressing interactions that fit the *machine* model is inherently easy within those contexts.

However, the more significant theoretical concept for interaction design presented and used in the context of HAPPY is the idea of treating multiple interactions, and in our case both the process and result, as meaningfully more than just the sum of their parts. The design of the emotion input type as realized inside HAPPY demonstrates this in practice and aims to directly express it through its output. However, it is important to note that the current work still leaves this as quite theoretical because of the lack of well-defined and closed relevance cycle. So significant future work would still be required, aspects of which previous chapters discuss in more practical detail.

Briefly returning to Murray's three discussed models in general, as they are originally defined by her, they are simultaneously quite abstract but also rigid in the sense that gaps must be filled in with concepts such as the idea of an expressive *machine*. Chapter 2, and specifically Figure 1 (p. 8), deliberately chooses perhaps somewhat controversial example for the *tool* model in that if the readers first analog frame of reference for a word processor is a typewriter it would not be wrong to argue that Figure 1 actually gets it wrong. However, when there are two things, in this case word processor and a voice recorder, that are ostensibly similar yet quite different in the way they are interacted with and they need to be fit in those models just assigning the same one to both would be equally problematic because the degree of manual effort and control is completely different. The reason why we return to this here is to emphasize that something like Figure 1 or the examples given in Chapter 2 are indeed context dependent and specifically that Figure 1, a visual aid, is just that and not some innovative representation that could be generalized in objective terms.

8.2 Co-design process

The co-design phase as a design process and part of this thesis work proved fruitful for the development of HAPPY. In this case its function as an effective objectivity check and a

framing context for the work was invaluable. This is because no matter what happened between the conclusion of the co-design phase and the development after that the core concept created and tested through this process would always act as the first point of reference for any designer originating changes. Specifically it helps in cases such as this, where the person responsible for implementation is also in part involved with some design details, because while it does not eliminate the risk of implementation concerns negatively affecting the design itself it does, however, significantly reduce it as well as the impact of designer bias in general.

One of the inherent risks related to the use of co-design is that the initiator, in this instance the author, but more generally the party responsible for executing on the produced design ideas could effectively, unintentionally, sell an idea to the participants. More specifically, whenever examples have to be given, they should be broad or general in nature focusing on what could be used and not so much how it could be used and especially not if something should or should not be used. An effective example of this being done correctly is how the initial example of an additional element during the co-design phase was just using color as an element or something else more abstract like it, while this quickly transitioned into discussion about pictographs, the example referenced by Figure 8 (p. 50) that started discussion in this direction came from one of the parents. On the other hand, as an example of a slight misstep, how the presence or lack thereof of a title as very specific structural element was mentioned in advance of the test week effectively ensuring it would come up later (see also Section 4.3.1). It is impossible to say whether titles specifically would have come up as strongly as they did, without this mention, but looking at other discussion had about the need for better support of retrospective browsing of entries it at least appears likely that something similar would have come up.

In retrospect, if the design cycle would be iterated further some form of co-design should be considered again. Especially when the process could now be refined with the existence of the new study server, which was developed after the conclusion of this phase, by for example allowing users to use a specific entry type to document their usage experiences during tests or even to collect design ideas. Particularly the part about documenting their

usage experiences with the application itself is something that could e.g. be applied as an opt-in style option to any future research project using HAPPY even if that research project would not be directly aimed at continuing development of the application as such. This could effectively create a varied dataset, if so agreed, that could then be used as part of any later work or during such future research that does aim to develop the tool further.

8.3 Research questions and results

This section seeks to provide some direct easily accessible answers for questions posed in Section 1.2. The answer to the question of additional value to the users using it directly is a straightforward one, yes. For examples of such added value one need only look at the discussions had during the co-design phase of the development. Some of these examples are even just about the circumstances where the application is used, rather than a specific feature or capability.

Specifically, about different formats and added value, we can see that the new, if simple, ability to visually combine different input formats under a one entry alone already does allow, due to the added flexibility, even to overcome some shortcomings of the current implementation through creatively combining entries. For example, the use of picture and text to express emotions the participant feels are not currently represented by the new emotion input type. One particular entry from the test week comes to mind, where the camera was used for a photo of a gesture that clearly expressed an emotion in a case where presumably the user felt the emotion input type did not provide strong or varied enough options as is.

In retrospect, the research and results presented here are not entirely dissimilar to those of an exploratory case study (Yin 2003). In effect what is being discussed here is a largely isolated case, in a specific context, namely the idea of applying co-design into what is basically a single iteration of the design cycle in Hevner's three cycle representation of design science research. Specifically evaluating the produced artifact and how it evolved as a result. Naturally, the question of bias becomes ever more relevant when considering what has been presented from this type perspective, no matter how self-critical one aims to be. However, as noted in the previous section, while co-design does not eliminate the problem of bias it does

offer relief from it by introducing multiple views and a solid recorded point of reference[20]. The takeaway from this comparison is that the conclusions drawn here are indeed dependent on their context and whether they can or should be generalized is a question for different further research, likely, using additional or expanded methods (see Section 8.4).

8.4 Future work

In addition to what has already been covered in the earlier sections noted previously in this chapter. In broader terms, any future work concerning HAPPY could include some of the following considerations. Additionally, as noted in Section 8.2, in some instances it may also be possible to employ the use of HAPPY to gather data about itself, however, this is something that should be carefully considered on a case by case basis.

Firstly, how to better integrate HAPPY into the proposed target environment. This means to say that currently HAPPY does not make a strong case for being an application for use in NICU's specifically, except through its previous association. This can also be considered a good thing, particularly in terms of HAPPY being useful as a more general-purpose research tool, especially when considering the degree of configurability offered by the study server. However, adding more targeted experiences could also provide significant added value to the users in the NICU context.

Secondly, the research and data collection side, i.e. the study server or HAPPY from the researchers' point of view. While this thesis comprehensively covers the user and client-side interaction and features of HAPPY the server component as presented in this thesis only qualifies as a starting point for future work focused on improving the capabilities of HAPPY as a research and data collection tool specifically. This is evident particularly through some of the scoping decisions done here, such as the choice to not focus on automated transcription or other speech to text functionality due to it not interfacing with the end users of the client application itself at present. Likewise, the same is true for the relatively limited changes in

[20] In so far as how co-design was executed here, relative to general software development, it is atypical to have full verbatim records of what is effectively requirements and feedback gathering.

the user facing functionality of the new study server and its focus on more technical improvements.

Lastly, regarding the emotion input type specifically, this thesis took an approach that is, for a lack of a better description, almost risk free. This is, naturally, in reference to the fact that due to the flexibility afforded by the study server as long as the emotion input type being available does not actively decrease the amount of collected data it should not be able to negatively impact the results of any data collection efforts using the new version of HAPPY, outside of potentially having extraneous data available. However, both risks can be minimized if not completely mitigated by proper instruction in the use of the application, as has been discussed for example in Section 6.4. This, however, does not mean that it would not be possible to further research the differences in the types of data collected with different configurations of the study server, including but not limited to the presence or absence of the emotion input type. This could be done for example by a series of A/B-style tests[21], presumably with qualitative elements since the problem of effectiveness in this context is more nuanced than measuring differences in levels of engagement, where the emotion input type could be either required, optional or completely unavailable. Further precision could be achieved by, for example, changing how the inputs inside the emotion input type interact with each other and the produced pictograph, by having it as a behavior present in one version but not the other[22].

[21] A/B, or split, -testing is a testing methodology where the impact of a change can be evaluated by providing two versions of something, in this case an application, one with and one without said change to two comparable groups of people (e.g. Johari, Pekelis et al. 2015).

[22] While the previous three examples of A/B tests could be executed with the new HAPPY as is, without any changes, the last one would require specific versions to be created because it would change the functionality within specific input type, not just the used configuration.

REFERENCES

ALASZEWSKI, A., 2011. *Using Diaries for Social Research.* London: SAGE Publications Ltd.

COOPER, A., REIMANN, R., CRONIN, D. and NOESSEL, C., 2014. *About face : the essentials of interaction design.* Fourth edition edn. Hoboken: John Wiley & Sons.

DAVIS, M. and EDBERG, P., May 21, 2018-last update, Unicode® Technical Standard #51 UNICODE EMOJI. Available: https://www.unicode.org/reports/tr51/tr51-14.html#Identification [Aug 01, 2018].

EKBIA, H.R., 2008. *Artificial Dreams : The Quest for Non-Biological Intelligence.* New York: Cambridge University Press.

FANE, J., MACDOUGALL, C., JOVANOVIC, J., REDMOND, G. and GIBBS, L., 2018. Exploring the use of emoji as a visual research method for eliciting young children's voices in childhood research. *Early Child Development and Care,* **188**(3), pp. 359-374.

FEELEY, N., GENEST, C., NIELA-VILEN, H., CHARBONNEAU, L. and AXELIN, A., 2016. Parents and nurses balancing parent-infant closeness and separation: a qualitative study of NICU nurses' perceptions. *Bmc Pediatrics,* **16**, pp. 134.

GSMA, 2019-last update, The Mobile Economy 2019 [Homepage of GSM Association], [Online]. Available: https://www.gsmaintelligence.com/research/?file=b9a6e6202ee1d5f787cfebb95d3639c5&download [Mar 20, 2019].

HARAVUORI, H., MUINONEN, E., KANSTE, O. and MARTTUNEN, M., 2017. *Mielenterveys- ja päihdetyön menetelmät opiskeluterveydenhuollossa : opas arviointiin, hoitoon ja käytäntöihin.* Helsinki: Terveyden ja hyvinvoinnin laitos.

HEKTNER, J., SCHMIDT, J. and CSIKSZENTMIHALYI, M., 2011. *Experience Sampling Method.* Thousand Oaks, California: SAGE Publications, Inc.

HEVNER, 2007. A three cycle view of design science research. *Scandinavian journal of information systems,* **19**, pp. 87-92.

HOFMANN, W. and PATEL, P.V., 2015. SurveySignal: A Convenient Solution for Experience Sampling Research Using Participants' Own Smartphones. *Social Science Computer Review,* **33**(2), pp. 235-253.

IDC, 2018-last update, IDC - Smartphone Market Share - OS. Available: https://www.idc.com/promo/smartphone-market-share [Mar 21, 2019].

JAEGER, S.R., XIA, Y., LEE, P., HUNTER, D.C., BERESFORD, M.K. and ARES, G., 2018. *Emoji questionnaires can be used with a range of population segments: Findings relating to age, gender and frequency of emoji/emoticon use.*

JOHARI, R., PEKELIS, L. and WALSH, D.J., 2015. Always valid inference: Bringing sequential analysis to A/B testing. *arXiv preprint arXiv:1512.04922*.

KAYE, L.K., MALONE, S.A. and WALL, H.J., 2017. *Emojis: Insights, Affordances, and Possibilities for Psychological Science.*

MURRAY, J.H., 2012. *Inventing the medium : principles of interaction design as a cultural practice.* Cambridge, Mass.: MIT Press.

NIELA-VILÉN, H., FEELEY, N. and AXELIN, A., 2017. Hospital routines promote parent–infant closeness and cause separation in the birthing unit in the first 2 hours after birth: A pilot study. *Birth,* **44**(2), pp. 167-172.

OU, L., LUO, M., WOODCOCK, A. and WRIGHT, A., 2004. A study of colour emotion and colour preference. Part III: Colour preference modeling. *Color Research & Application,* **29**, pp. 381.

PALMER, S.E. and SCHLOSS, K.B., 2010. An ecological valence theory of human color preference. *Proceedings of the National Academy of Sciences,* **107**(19), pp. 8877-8882.

SCHLOSS, K.B. and PALMER, S.E., 2017. *An ecological framework for temporal and individual differences in color preferences.*

SCHOUTETEN, J.J., VERWAEREN, J., LAGAST, S., GELLYNCK, X. and DE STEUR, H., 2018. *Emoji as a tool for measuring children's emotions when tasting food.*

SEUFERT, E.B., 2014. *Chapter 1 - The Freemium Business Model.* Boston: Morgan Kaufmann.

STATCOUNTER, 2019-last update, Mobile Operating System Market Share Worldwide. Available: http://gs.statcounter.com/os-market-share/mobile/worldwide [Mar 20, 2019].

SWANEY-STUEVE, M., JEPSEN, T. and DEUBLER, G., 2018. *The emoji scale: A facial scale for the 21st century.*

WEINGERL, P. and JAVORŠEK, D., 2018. Theory of Colour Harmony and Its Application. *Tehnicki Vjesnik,* **25**, pp. 1243-1248.

YIN, R.K., 2003. *Case study research : design and methods.* 3rd ed edn. Thousand Oaks (CA): Sage Publications.