Exploring the Potential for Improvement in Parents of Small Children's Sleep using Smartwatches and Embedded Sensors

Anders S. Lykkegaard, Henrik Lanng, Julian A. R. Dam, Søren Thornholm
Department of Computer Science
Aarhus University
{aslykkegaard, henrik.lanng, julian, soren.thornholm}@post.au.dk

ABSTRACT

In this study we explore how we can use smartwatches to improve the sleep experience among parents of young children. The first studies were made to explore different wake up methods and here we concluded that haptic feedback was the most appropriate method to our study. Afterwards qualitative and quantitative data was gathered in form of interviews and questionnaires. From the questionnaires and interviews we developed a prototype evaluated by using field studies through four iterations. During the evaluations a second questionnaire was captured the target demographics feelings on subjects such as smartwatches and electromagnetic radiation. With this data we discuss the possibility to potentially improve parents' sleep, avoid conflicts when the parents decide which parent should take care of the child, and lay a foundation which further research can be built upon.

INTRODUCTION

As the market for smartwatches is rising[1], empirical research shows that people still consider smartwatches as a second screen for their notifications[8].

We set out to explore other uses for smartwatches by focusing on it's inherent nearness and on-board sensors. The focus area of this exploration is in the context of traditional baby monitors.

Modern baby monitors typically work by using a one-way radio communication. If a baby is sleeping in a room separate from its parents, the parents will typically be woken by the sound of the baby crying originating from the baby monitor device.

In many cases, it is likely that just one parent needs to attend to the crying child. Consider the case where a child is still being breastfed. In such cases, it is likely that only the mother will need to attend to the child. By using the discreet haptic feedback made possible by smartwatches, we explore it's potential to eliminate the cases in which both parents wake up.

In situations where the parent that attends to the child is not of particular importance, we explore methods in which this technology can help in making an agreeable arrangement.

Many different wearable computers could potentially be used in such a system, but most of our current research is centered around smartwatches, as these are relatively accessible, programmable wearables.

RELATED WORK

While wearable computers have been around since the 1970s[2] in the form of calculator watches, it is only in the last decade that smartwatches have started gaining traction among consumers. The increased usage is likely owed to the increased functionality and lowered prices, as well as more capable batteries [6].

One of the major areas with potential for wide-spread smartwatch usage is mHealth, which concerns the use of mobile devices for fitness tracking, sleep tracking, and the like [6]. The area of mHealth offers potential improvements in the consumer's sleep experience'. While we will not be diving deeply into this area of research (although, we will be incorporating sleep tracking in the prototype), it should be noted that this is a very relevant area to our research domain. Instead, focus will be put on the empirical research on the target demography and current related products.

The babymonitor domain inherently has a very broad target demography. The parent demography has an age range of multiple decades and may come from vastly different cultural backgrounds, professional backgrounds, and socio-economic situations.

Three separate key articles, Martin[4]; Schirra, Steven and Bentley, Frank R[8]; and Cecchinato, Marta E. and Cox, Anna L. and Bird, Join[1], all focus on the current importance of the aesthetics of the wrist-worn wearables. Furthermore there also seems to be a big confusion as to where exactly current smartwatches find their place in today's world. The two latter articles are empirical studies among current smartwatch users. While presenting some interesting views, one should keep

[1] https://www.strategyanalytics.com/strategy-analytics/news/strategy-analytics-press-releases/strategy-analytics-press-release/2015/07/22/apple-watch-captures-75-percent-global-smartwatch-marketshare-in-q2-2015 *Retrieved 2015-12-28*

[2] https://www.media.mit.edu/wearables/lizzy/timeline.html#1977b *Retrieved 2015-12-28*

in mind that their sample sizes are very small. Schirra and Bentley have a sample size of 5, while Cecchinato et al. have a sample size of 10, most of which are within the IT profession. While current empirical studies on smartwatch usage is limited, it seems that aesthetics is a key factor to keep in mind when working with wrist-worn wearables, along with a sense of purpose other than just as an additional screen for the smartphone. Particularly the female demographic seems to be neglected in respect to the fashionable smartwatches, or as Schirra and Bentley put it: "Both women in our study discussed the challenges of finding a smart watch on the market that seemed designed with women in mind. As [female tester] noted, "It looks like they were [all] made for a traditional man to wear... it will just look like a big, huge man's watch."[8]. Since women represent a major part of the target demographic, we will later in this article compare our findings to see how they align.

Samsung has released a relevant feature - the so-called "Baby crying detector"[2]. This feature is present on the smartphones Galaxy S5 and S6, and only works coupled with their smartwatches, such as the Galaxy Gear Watch. The feature is placed in the hearing-accessibility tools which suggests that it is intended to be used by those hard of hearing. This is supported by the fact that the smartwatch must be within bluetooth range of the phone. Furthermore, this feature doesn't seem to be widely promoted (for instance, searching Samsung's website only returns support articles) and the massive disclaimer presented upon activation of this feature does not bolster confidence. Through desk research, we were unable to find any users of this feature. This system only supports one connected smartwatch, it requires the care-taker to leave their smartphone in a silent room within 1 meter of the child, and it uses bluetooth, which based on our empirical data may not be particularly well suited. If the parents do not wish to leave their smartphone with the child, there is no possibility of re-pairing the smartwatch to another smartphone without factory resetting the device. Nor is it possible to pair with other devices such as dedicated baby monitors.

Bellman & Symfon has a baby monitoring solution [3] that can connect to a wrist receiver. Their wrist receiver is targeted at those with hearing impairments and is capable of producing haptic feedback upon receiving notifications from devices in their Visit product line. These include notifications from their baby monitors, doorbells, and smoke detectors. The wrist receiver has a battery life up to 30 hours, and a charging time of about 8 hours. This product assumes that the wearer charges the device at night - while the device is charging, a separated bed shaker can be plugged into the wrist receiver. The device is limited to giving haptic feedback, a clock is not included. If one wishes to use the wrist receiver as an alarm clock, a separate alarm clock within the product line is required. Buying a compatible baby cry transmitter, alarm clock, and wrist receiver costs more than 390 GBP (> 3960 DKK) [4]. By using smarter devices, such as a smartwatch, we could potentially eliminate the use for a separate alarm clock, as well as utilize various sensors to give the option of of waking up a person while he or she is in the lightest sleep, provide simultaneous sleep tracking, offer greater configurability. Searching through Google Scholar and the ACM digital library we were unable to find any research that utilized more than one smartwatch in a single system.

The articles above conclude that the aesthetics of the wrist-worn wearables are very important and that most people often see the smartwatch as an additional screen for their smartphone. And as for existing baby monitors capable of alarming through wearables, numerous issues were outlined. We will therefore explore if the smartwatch can improve the existing baby monitors and try to explore what the possibilities are.

MOTIVATION FOR EMPIRICAL STUDIES

The goal of this section is to inform why this challenge is important to explore and what our motivation is for exploring the challenge.
Through research we discovered a gap between technology and the use of technology in parenting and how technology can improve the parents sleep pattern. The issue is that few have conducted research about this gap and to explore this we need to understand how parents with small children sleep, what conflicts are linked to this field and how parents handle the decision on which of the parents should get up to comfort the small child. These are the main points we need to explore before designing a prototype.

EMPIRICAL DATA

In our empirical studies we have investigated the current features in baby monitors and learned about the good and bad sides from the users. We then proceeded to distribute a questionnaire to users within our target demographic about the habits concerning who cares for the child when the child cries. Qualitative interviews with two mothers how they decide who does it, how its done, as well as their thoughts and emotions on various subjects such as how they feet about having smartwatches as alarm-receiving devices in the context of baby monitors. Besides the two interviews, we conducted a casual (unstructured, unrecorded) interview each time we delivered the prototype.

To learn more about current top trends in baby monitors, we went to a local baby product store (BabySam). An employee showed us their current selection and supplied us with knowledge on the popularity of the different monitors. Their selection consisted of 4 different baby monitors: 2 one-way radios; 1 two-way radio; and a one-way bluetooth monitor, that uses a smartphone application as parent-end receiver. Customers were initially most interested in the bluetooth monitor, making it the best seller. However, the bluetooth monitor also had the highest return rate. The high return rate was due to customers coming back claiming the bluetooth monitor was unable to

[3] http://bellman.com/en/our-solutions/baby-monitoring-solution/ *Retrieved 2016-01-02*

[4] First product retailer found through a Google search. Price shown in the text is the summed price include VAT (excl. shipping etc)
http://www.actiononhearingloss.org.uk/shop/bellman-wrist-receiver-inc-charger-product-b8101.aspx
http://www.actiononhearingloss.org.uk/shop/bellman-baby-cry-transmitter-product-b1491.aspx
http://www.actiononhearingloss.org.uk/shop/bellman-alarm-clock-visit-product-b1580.aspx *Retrieved 2016-01-02*

carry the signal through the walls in their homes. Adjusting to the return rate, it was a simple one-way radio that sold the best and stayed out with the customer.
The reason for the customers initial interest in the bluetooth monitor is not clear-cut. One possibility is the fact that it was slightly cheaper than the radio (1,199.- vs 1,299,- DKK)[5], but it could also be due to the extra features that could be acquired through in-application purchases, which consist of the ability to track the amount of sleep a child gets, the ability to listen to the child through the application, and checking the temperature in the room.
Most wearables today use bluetooth as the primary form of data transmission. Since customers are experiencing connectivity issues between their bluetooth monitors and smartphones, this additional challenge will be considered in our quest to introduce wearables as an aiding tool.
When describing the pitching process for the baby monitors, the employees would highlight the security in one baby monitor and the electromagnetic radiation level in another baby monitor. The fact that the best selling radio baby monitor was also the one with the best security (required to be within a short distance when pairing, and could only be paired with one receiver at a time) could be a reflection on the customers' security considerations.

After analysing the market, the attention was turned to the customers, or as we will call them henceforth, users. Two different approaches were utilized: A web-based questionnaire was made, two semi-structured [7], and four unstructured, unrecorded interviews. The semi-structured interviews were recorded over the phone and the unstructured interviews were done on delivery of the later system prototypes. Questionnaires were used to gather quantitative data, in order to gain insight to specific questions. The questionnaires were used to reach a broader audience than we had immediate relations to. This was achieved by reaching out to Danish forums for parenting. The semi-structured interviews allowed the users to answer all our questions openly and to let them express their thoughts of the situation. This also gives us the ability to explore the topic further if necessary. The two semi-structured interviews and the four unstructured were analysed by transcribing the key points and identifying patterns and themes[7].

The interviewees had two children each. The babies/small children had an age range of about 8 weeks to 3 years old. One of the things our interview showed was the fact that parents' stances on adopting new technology to replace their baby monitors were very polarized. While one family was a little more sceptical of the idea of using new technology, others, including the families from the casual interviews, welcomed the idea of a new form of baby monitoring. In response to how the interviewees felt about new technology, as well as smartwatches, replacing their current baby monitors, one mother responded: *"This would be fine with me, as long as it wakes me up"* [Appendix PI2 (5:12)] (translated from Danish), where the other interviewee answered the following: *"This would be fine with me, but one thing that concerns me when I am sitting with my baby and I have my own smartphone on me all the time is the electric magnetic radiation, because you do not know a lot about this... I think."*[Appendix PI1 (4:14)](Translated from Danish). There is also a big difference in when parents are using the baby monitor. One of the families did not use the baby monitor during the night because no door was closed between their room and the baby's. The another interviewee used the baby monitor at night instead of having the doors open. Lastly, one interviewee was questioned on who generally took care of the children to which she answered: *"That is me, but sometimes I can push my husband so that he can get up and take care of the oldest one of the children"* [Appendix PI1 (2:34)].

Questionnaires were answered by 78 people and analysed by inserting the data into a spreadsheet [7] and constructing categories for all qualitative questions. The spreadsheet is primarily used to categorize and visualize the responses.
Of our 78 participants 68 answered how often their child woke up during night. The distribution is seen in figure 1. The respondents' children were between 12 days and 4 years and on average woke up 2.9 times a night with some on average 0 times and some on average 8 times a night. As the respondents are woken up 2.9 times a night there is certainly room for improvement of parents' sleep if just one is to be woken up at a time. 50 of the participants said that it is the mother that primarily takes care of the child at night. In 19 of these 50 cases, this was a consequence of breastfeeding or maternity leave, and in 16 of the 50 cases were due to the fact that the mother was always the first to wake up and/or due to the fact that the father is unable to hear the crying. 16 participants said that it is primarily the father who takes care of the child and 13 said it is evenly split between the parents. In general the questionnaire lead to six different ways in which people chose who should tend the child. The first being the child was still breastfeeding, the second being that one parent was on maternity leave from work and therefore took the majority of incidents. The third being that the child only becomes quiet when the mother is nursing it, this was explained by families as "the child only finds comfort with its mother". The fourth was that they take turns at being the one to do it. We also had a category for families with one parent only, this was to see their thoughts on the subject. Lastly some families answered that the one partner simply required more sleep to function throughout the day.

The age of all the children of the questionnaire respondents is not known with certainty. This is due to an error in the questionnaire form, in which the field asking about the age doesn't specify which unit the age should be given in. A third of the respondents did not specify whether their input was in weeks, months, years. The age data used was interpreted from what sounded likely based on the other answers of the questionnaire - it is for instance unlikely that a 5 week or 5 month old child would have a younger brother.

In summary, the key findings of the empirical data was about different technologies that is used in baby monitors, especially the potential unsuitability of bluetooth monitors was interesting. There were three main findings from the interviews which are the subjective responses on when and how they used their

[5]Prices were gathered in-store.

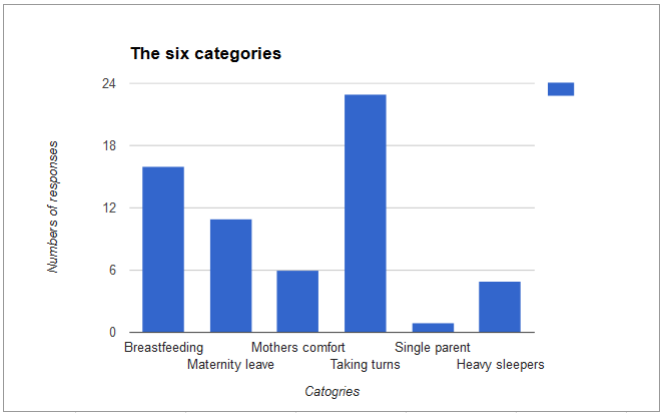


Figure 1: A graph showing how the one who takes care of the child is chosen

baby monitor, the concern about the electromagnetic radiation level, and the fact that one of the mothers needed to push the farther to get him up to take care of the child. The web-based questionnaire gave an overview on different categories inside the area of whom generally took care of the child when the child woke up at night. Furthermore it gave insight into how often the average child wake up at night and who takes care of the child.

Establishing requirements

Through our gathered empirical data we were able to establish some requirements for our further research. An important aspect of baby monitors are their reliability. The aforementioned bluetooth monitor was often returned due to losing connection, while these results are likely to vary from house to house, we chose not to go with bluetooth as our domain assumes that children are in separate room from the parents. Security is also a concern mentioned by the employee we spoke to and is also a requirement to consider. A baby monitor connected utilizing a secured WiFi connection, utilizes the same connection that is trusted with credit card details and so forth. This security is therefore likely to updated frequently as the Internet technologies evolve.

While electromagnetic radiation is a concern among parents, it generally seems to be little understood. By utilizing WiFi, we will be adding to the families' already existing WiFi network, and replace current baby monitors, which in themselves are sources electromagnetic radiation. No requirement addressing this is therefore set.

The questionnaire revealed that children woke up and needed parental attendance an average of 2.9 times a night. This, as well as the potential conflicts we discovered during the interviews, serves to reaffirm the original goal of waking up one parent rather than both. To avoid conflicts in situations where the parent needs to attend is not of importance, sleep tracking will be incorporated in an attempt to even out the distribution of who attends to the child, hopefully in way that is perceived as *fair*.

EXPERIMENTS

Based on our original visioned concept, the first prototype in the form of a video prototype was developed, even before most of the empirical data had been gathered. This was used to demonstrate the overall idea of using wearables as a baby monitor when sleeping. The video prototype was shown to fellow peers, and their feedback was taken into consideration.[6] The feedback on the video prototype was that the idea of waking one parent instead of two was good and they suggested that the parent who should be woken up is the person that slept the lightest, in which measured by the smartwatches. This was built into the design, since it was reasonable solution to improve parents' sleep pattern and a potential way to avoid conflicts between the parents since this method objectively decides which parent to wake up based on who will be disturbed the least.

Exploring wake up methods

The most natural choice for waking one out of two parents using smartwatches seemed to be vibrotactile feedback. To reaffirm this choice, various other alternatives were considered and evaluated.

We analysed three different ways of waking up: by sound, by light, by vibrotactility through a bed shaker, and by vibrotactility through a smartwatch. A normal baby monitor works primarily by reproducing the baby's crying to get the parents' attention. By using sounds it is very possible that both parents wake up when the child is crying. The sound could be increasing in sound level, but that would increase the time it takes for the parent to wake up and taking care of the child. Furthermore it could also wake up both parents. Then we looked into light as a wake up method as for example the Philips Wake-up Light[7]. The problem here is that the light will likely light up the whole room - this hypothesis was tested and confirmed by trying the Phillips Wake-up Light internally (that is, on ourselves). It could maybe work if the parents had one each and they lay back to back on each other such that the light will be primarily focused on one parent at a time. The Philips Wake-up Light primarily works by slowly incrementing the light over a 30 minute period which is definitely not preferable for a baby monitor, as there could be delay of up to 30 minutes before the child is taken care of. Vibrotactile feedback (vibration) was another possible wake-up solution. The vibration device would need to be at the side of the bed of the parent who needs to be woken up and it would need to be powerful enough to wake up one parent but low enough to keep the other parent from being awakened. Vibration could also be used through a smartwatch or a bedshaker as you can wear/use it at night. That way you could keep the haptic feedback focused on one parent at a time. Both the bedshaker and the smartwatch were tested to see how they worked out and see if it could be a possible solution to wake up one parent. The bedshaker did not work as expected because the vibrations spread through the mattress and did cause the partner to wake up at multiple occasions. Furthermore, another drawback identified through internal evaluation was that the bedshakers easily became uncomfortable due to their size and hardness, thus becoming an obstacle when rolling around in bed. The

[6] https://drive.google.com/file/d/0B7_cFEg96cc8cVgxNUo0c3pnNHc/ *Retrieved 2016-01-02*

[7] http://www.philips.dk/c-p/HF3500_01/wake-up-light *Retrieved 2016-01-02*

smartwatch did seem to work better, as it was a more local and lower strength of vibrations and it did not wake up the immediate partner.

After these considerations and evaluations, we concluded that vibrotactile feedback is the most suitable type of feedback for the concept described in the introduction.

Field Studies

Based on the original envisioned concept in the introduction, coupled with empirical data and the previous evaluations, we set out to make a prototype system. The prototype system was given to three families fitting the target demographic to be evaluated for 3-5 days over a total of 4 iterations (that is, one family had the tested system over two iterations). Along with the prototype system, a diary was provided with some predefined questions for them to reflect on and answer after every night of usage. This, along with casual interviews upon delivery and retrieval of the prototype system, was used for later evaluation.

SYSTEM DESCRIPTION

Here a brief description of an ideal system will be presented, followed by a detailed overview of actual system prototype implementation as used in the field studies.

A system notifying parents individually of a child crying, based on their sleep level, could be realized using two smartwatches (or other wearables) and a baby monitor. The smartwatches should be able to provide haptic feedback, as well as be able to collect and send sleep data upon request. A baby monitor in such a system would function as a traditional radio-based baby monitor (triggered upon noise exceeding a threshold), with additional connectivity capabilities. As discovered in the empirical data, bluetooth connections tend not to be sufficient in baby monitoring systems. Thus WiFi would probably be the better choice, considering its already established popularity.

Due to the test families not having smartwatches, and due to the technical limitations of the smartwatches (mainly the lack of built-in WiFi connectivity), the prototypes used for our field testing were considerably more complex, as detailed in the next section.

SYSTEM PROTOTYPE IMPLEMENTATION

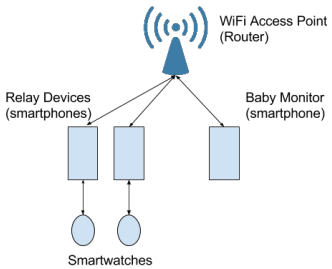


Figure 2: System Architecture

In the following section a system description will explain how the communication between the devices functions, and arguments for using different technology that the watches and smartphones contains. In the end of the section there is a quick summary of the system.

Two near identical prototype systems were prepared to allow for concurrent field testing. Each system consists of three Nexus 4 smartphones and two smartwatches. The systems only differed in the smartwatches that were utilized - one system used two LG G Watch R[8] smartwatches and the other two Moto 360[9] smartwatches. This difference is solely the result of being the smartwatches that were readily available.

One smartphone acted as the baby monitor while the two others were used to facilitate connections between the phone used as a baby monitor and the smartwatches, and to gather sleep data. Three Android applications were developed to realize the desired prototype.

Baby Monitor Application

An application that mimics traditional baby monitors was developed for the smartphones that acted as baby monitors. This specific smartphone in the system and application will henceforth be referred to as the baby monitor and the monitoring application respectively, as the phone for all intents and purposes acts like a baby monitor.

The baby monitor listens and accepts TCP sockets from the connection relay and the sleep tracking application found on the two other smartphones. When monitoring the noise level (amplitude) captured by the phone's microphone exceeds a specified threshold, a "trigger" event is sent to one of the connected devices. The device to send the event to is determined based on the sleep data it replies with: The one that responds with the highest movement average gets the trigger event.

The interface, as seen on the baby monitor in figure 3a contains the following: a status label informing of the devices currently connected; an address label; a button to toggle the monitoring; a sensitivity drop-down menu, which allows users to set the noise threshold; and a button to toggle the testing functionality, which allows users to test the noise threshold without triggering the alarms.

Connection Relay and Sleep Tracking Application

This application has two main responsibilities: To propagate the trigger event from the baby monitor to the smartwatch and to gather sleep tracking data. Which devices with this application installed will henceforth be referred to as the relay devices.

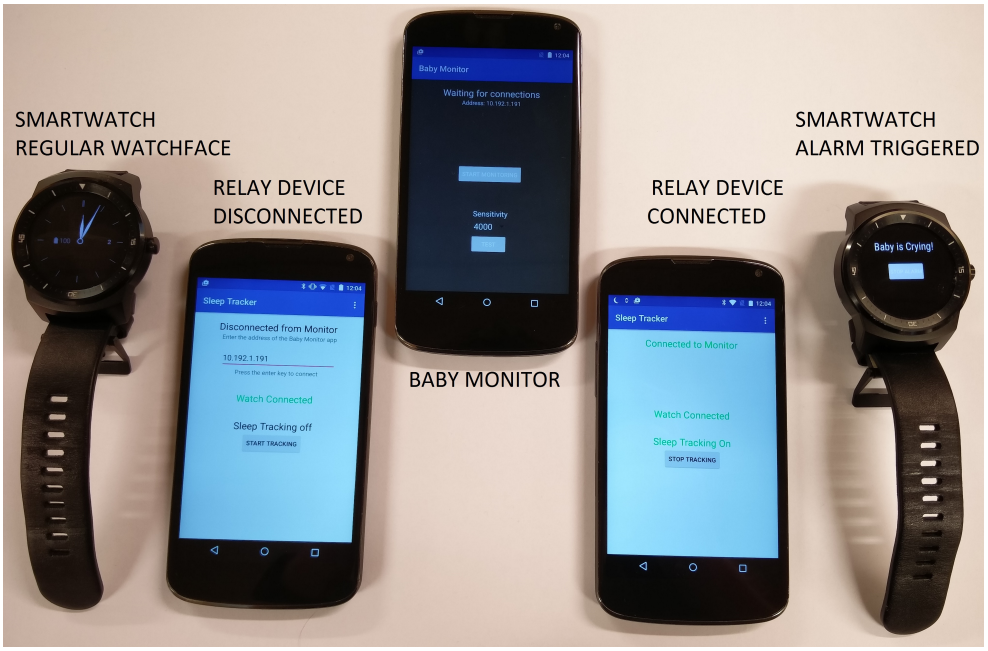
The connection to the baby alarm can be established through entering the IP address of the baby alarm. The sleep tracking data was gathered through Sleep as Android's API [10].

The application interface is seen on the relay devices in 3a. Upon opening the app the first time, the user should enter the address of the baby monitor to establish a connection. This is only necessary when the address of the monitor changes, such as when connecting to another access point.

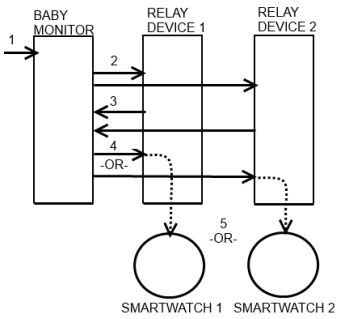
[8] http://www.lg.com/global/gwatch/index.html#gallery
Retrieved 2016-01-02

[9] https://motorola-global-portal.custhelp.com/app/answers/prod_answer_detail/a_id/107206/p/2815,9141
Retrieved 2016-01-02

[10] https://sites.google.com/site/sleepasandroid/doc
Retrieved 2015-12-28



(a) The 5 devices of the system prototype. Screenshots can be found in the appendix (PSC1-4), as well as a video (V).



(b) Device Communication Sequence: (1) The baby alarm is triggered by a crying baby, (2) the baby monitor requests sleep data from the two connected relay devices, (3) the relay devices respond with the sleep data, (4) the baby monitor compares the sleep data and orders one of the relay devices to trigger an alarm, (5) the relay device that receives the order propagates the order to the connected smartwatch - the smartwatch starts vibrating.

Figure 3: The System Prototype

When "start tracking" is pressed, Sleep as Android starts logging aggregated acceleration averages over the X, Y and Z axises from the connection smartwatches. When the relay application receives a trigger event, the average over the last 30 minutes of the acceleration data logged by Sleep as Android is sent to the baby monitor, so the baby monitor can determine which watch has moved the most, and therefore is sleeping the lightest.

Baby Alarm Application

This application is installed on the smartwatches. Its main responsibility is to open and start the vibration when receiving the trigger event from the relay application. When an alarm is triggered, the wearer is presented with a screen as shown on the right side in figure 3a. The wearer has the option of stopping the alarm, or closing the application. It will however be re-triggered every time the noise threshold is exceeded, which if the baby is crying, happens in intervals of 5 seconds. Furthermore, standard watch notifications are displayed upon connection establishment and disconnection with the baby monitor.

The application may also be opened to test the alarm. This may be used to make the wearer aware of what to expect when the actual alarm is triggered.

Summary and Argumentation

The components are connected as shown in figure 2. The baby monitor was placed near the baby in a separate room from the parents. The parents wore the smartwatches while sleeping. The relay devices had to be close to the smartwatches due to the range-limited bluetooth connections.

The use of the selected smartphones and smartwatches was decided based on the facts that these devices were readily available and all use the same platform, allowing for better programming-time efficiency. This was of great importance as we needed to be able to quickly implement new and rework features between field tests. Furthermore, as we did not know what devices the test users would have at hand, for easy installation, and for privacy concerns, the system was delivered on a set of test devices, rather than installed on the users' devices (which in many cases wouldn't be possible due to incompatibilities).

In summation, the high-fidelity prototypes were built with the three principles of prototyping as defined by Y.K. Lim et al. [3]:

- *The Fundamental Prototyping Principle*: The design space chosen to investigate was the core of the system concept-attempting waking one, and only one, parent based on sleep level when their baby is crying.
- *The Economic Principle*: The devices used were readily available for educational and research purposes through the university. Devices with the same platform were used (e.g. as opposed to building and programming custom baby monitor hardware).
- *The Anatomy of Prototypes*: The prototypes focus on the functionality of the core concept of our imagined system. Other features such as the application interaction (with the exception of the deliberately designed haptic feedback upon alarm trigger) and aesthetics were filtered out.

PROTOTYPE LIMITATIONS AND SHORTCOMINGS

The prototypes used for field testing had a number of limitations. The biggest one perhaps being the sole amount of

devices. In an ideal system, the relay devices would not have been required. The amount of devices within the prototype system, especially the three smartphones, was a source of confusion for the first test family: they found it hard to keep track of the paired prototype devices. Later test families were given devices that were clearly labeled with the labels "MOR" (mother), "FAR" (father), and "BABY", but even this is not necessarily a big improvement, as the families still needed to keep track of which devices needed to be charged and so on. Furthermore, the users would also be using their own devices. This factor contributed to confusion among some users, as they were not necessarily familiar with the basic operations of the devices.
The primary shortcoming of the prototypes was the lack of ability for the users in selecting which parent should be awoken upon the baby monitor being triggered. The motivation for this is that the one of the focuses of the evaluations was to test the multi-user sleep-level tracking functionality. However, a workaround was to simply connect one smartwatch/relay device to the baby monitor.

Status Quo Limitations
Our research builds upon the assumption that parents are willing to sleep with the wearables on at night. Preliminary interviews have suggested that it is common to not sleep with the smartwatch on, but rather leave it to charge at night. Despite this, we believe our assumption is fair, as (1) it is not unreasonable to expect that battery life of wearables will increase over time, and (2) users that currently view their smartwatch as an additional screen might not see much purpose in sleeping with their smartwatch on. In providing a purpose we believe that this pattern might change.
Addendum: The smartwatch we are using is the LG G Watch R, and has a battery-life of 1 to 2 days, and takes around 30 minutes to fully charge - thus charging it over a whole night is far from necessary.

EVALUATION

This section covers evaluation of the system prototype described earlier. A second questionnaire on user experience in relation to the parents in the target demography subjective stances on new technology. In the end the key points from the evaluation will be summed up.

Our evaluation is a field study where the users are in a natural setting [7]. This study was chosen due to the time-consuming nature of the concept and due to the subjective experiences parents' might experience when it comes to technologies involving their children. The test persons were equipped with a diary in which they were instructed to answer a set of questions on daily basis. The question was: did the child wake up, and if yes how many times; did you wake up because of the test system, and if yes was it both of you, just one; was there any problems with the system, and if yes please explain further; how was your (singular and plural) sleep; and finally: question own notes/comments (translated from Danish). The reason why diaries were utilized is that they do not require many resources and they are useful for gathering subjective data in long-terms studies. Another advantage is the fact that they encourage users to write down their impressions day by day. This lessens the risk of initial impressions to be overridden. In addition to the diaries, the second iteration of the baby monitor application also logged events such as application start, selected sensitivity level, device connection and disconnection events, as well as alarm trigger events and sleep tracking data received. All log entries were time-stamped and no personally identifiable information was logged by the application. This data was used for evaluation and to iteratively improve upon the system prototype.

The first test family had two children, one which was under a year old and the other one was around three years old. The mother and the father had split the task between them on which parent took care of which child during the night. The mother took care of the youngest child because she was breastfeeding. The father of the family did take care of the older child, however the mother often took care of this child as well, because the father often did not wake up. This family had the prototype two times. First time they had the prototype for around five days. The first time the family had the prototype there were troubles with connecting the equipment and even though we had tested in our environment the prototype was very unstable and therefore the test went less well than expected. The first iteration still yielded some results. The findings here was that they found the smartwatches uncomfortable to wear, because they thought the watchband felt cheap in materials, and they found it big to wear, which also aligns with the related work on the importance of aesthetics and size.
The second time they got it was around a week after the first time.
After the test we went back to improve the system and came out with a new version of the software and some different watches, to the same family. This time the system was much more stable and well working. The discovery the first night was that the father claimed that he did not wake up by the vibration on the watch, which was the first and only instance of this problem we have had during these tests. Other discoveries was that the woman did not like to wear anything around the wrist and especially the smartwatch because she was afraid to scratch the children with the watch due to its size. Additionally they found it hard to keep track of all the devices and to connect them. Furthermore the watches' screens lit up for no reason during the night which they found very annoying.
Afterwards we interviewed one of the test persons to compare the two prototypes. The findings here was that the watches were much better looking, but still to big to wear. The improvement from prototype one to second prototype was a much better system, although she found it very difficult to use. This was ude to the numerous devices and chargers (the prototype system consisted of 3 smartphones, 2 smartwatches, and 5 chargers in total).

To accommodate some of these challenges we did improve the prototype with the following features: we made a guide with screenshots included, including an instruction on how to turn off the screen on the smartwatch until the tactile button is pressed; made it easier to see when the smartwatch is connected to the baby monitor, and made it possible to turn sleep tracking on and off within the application, rather than through a 3rd party application.

The third iteration of the field studies was conducted with another family. This study was meant to be over a couple of weeks, but as we discovered at retrieval the system had only been used once even though we had contact with the family a couple of times during the testing period. We did however get some good feedback based on the family's experience with the system. One big challenge for the test family was to get the system to work properly and keep track of all five devices and their battery levels. They tried to set the system up twice but the second time one device was out of battery. No further attempts were made. The one night where they tested the system the mother was woken often because of the father's watch lit up even though *theatre mode* was enabled. As we primarily stated that the system was to test in a two parents situation they used only one smartwatch as the mother was still breastfeeding the child. The system could also be used by when only one device was connected. There was also some confusion of what the sensitivity is and how it should be understood. The family said that the system could also be useful in the daytime as there are places where it would be more acceptable to just have vibrations instead of sound. They said that the system might also be useful for older children - if the parents got more than one child they may divide the nursing of the children between them. Then the system could be used to wake up the right parent and not both. One question they asked was if a fitness band could be used instead. This feedback will be discussed later on.

The fourth iteration of the field studies was with a third family. In this iteration the system was used for two days and an interview was conducted afterwards. Overall the family was happy with the system and found it easy to set up. During the test the father was woken up six times and the mother was woken up once. They commented on the type of sleep tracking we are using and questioned the feasibility of this type of tracking. One note they had was that it would be nice to have a snooze possibility, not to snooze the alarm, but when they are up and taking care of the child. Another note was that they would like to be able to decide if the sleep tracking should decide or a time based system where the father for example took the first part of the night and the mother the last part. A third note was that the alarm once got activated because the child coughed. A fourth note was that the family tried to disconnect one device during the night because the mother was very tired, but the system stopped working and did not alarm the father when the child cried. The family even used it once during the daytime when the child was napping and found this useful and nice, as they did not have to remember to carry the parent-end of the baby monitor, as it was already on their arms. This test the family was not disturbed by the light from the smartwatches like the other two families were. As a last comment we were told that even though the mother was breastfeeding, it should not be underestimated how much it was appreciated that the father could bring the child to her.

After the first two evaluations and discussing our empirical data we fund a lack in our empirical studies. The lack is data on how users subjectively feel about new technologies in baby monitoring. To explore this we developed a questionnaire asking people about their feeling and position on the use of new technology on the topic of baby monitoring. The questionnaire was answered by eighteen people and put on the same four forums as used before. The questionnaires used open questions, so they could elaborate their answer [7]. These are the findings: we found that fourteen out of eighteen people were positive about trying new technology in their baby monitoring activity. What was important was the reliability of the system along with the user's need of confidence that the system will work. Furthermore eight out of seventeen do not believe that a vibrations from a smartwatch would wake them up or they did not want to sleep with the smartwatch. To secure that one will be woken up there needs to be some sort of back-up solution if one did not wake up from the vibrations or is out of range of the phone which was another concern. One expressed that the system had to either be better or at least at the same level as the existing solutions. Another consideration is the background noise around the baby which also have to be handled in some way. A question we tried to get an answer for in the questionnaire is if parents considered electromagnetic radiation levels upon purchasing their baby monitor. Nine out of eighteen people did some research on the electromagnetic radiation the baby monitors emit. One thought that if the electromagnetic radiation was a problem the devices would not be approved. One answered that they used test from the Danish consumer council (Tænk)[11] or similar tests. One answered that they did check the electromagnetic radiation levels of the alarm, "but without actually knowing more about it. It just seems logic that less electromagnetic radiation is better..." (translated from Danish). Electromagnetic radiation is of some parents' concern and should be examined, also such that we are able to compare to other baby monitors.

The key findings during the evaluation, was that the father did not wake up to the vibration during the night which we found problematic since the system should secure that the parent is woken up, that most people were positive to the use of new technology and some people do not believe that they can be woken up by the vibration or do not want to wear the smartwatch at night. The findings from the questionnaire were that people generally were positive to the use of new technology in baby monitoring situations, however there was concerns about that some did not believe that vibration is enough to wake them up, that they do not like to sleep with the watch, that there have to be some security if the devices disconnect or are out of range, and at last that the electromagnetic radiation. These key points will be discussed either in the Discussion section or in the Further Work section.

DISCUSSION

The discussion will focus on the key findings from the empirical data and evaluation. Thereafter we will discuss some of the findings in context of the related work.

This study builds upon the assumption that the target demographic is willing to sleep with smartwatches on, given a purpose to do so. 2 out of 17 respondents stated that they would not sleep with a smartwatch in response to the question: "how would you feel about being woken up by vibrations from a smartwatch" (translated from Danish), while 6 out 17

[11] http://taenk.dk/ *Retrieved 2016-01-02*

were sceptical to whether the vibrations from a smartwatch is enough to wake them up. How many of these respondents have actually tried or own a smartwatch is unknown, however even if they have, the fact that such a sceptical stance was shared among almost half of the respondents may represent a significant issue in the target demographic's will to adapt systems such as our envisioned system.

As seen in the evaluation, further negative feedback was received on the appearance of the smartwatches from one mother. This aligns well with the empirical studies in the related work section. Perhaps inherently to screen-based smartwatches is the issue with incorporating a screen of a usable size into a small and fashionable watch. A solution to this problem could be to utilize other wearables that do not include a large screen, such as fitness bands.

Despite the previously mentioned scepticism, the rest of the respondents seemed fairly positive about the idea of letting new systems replace their current baby monitors, including through the use of vibrotactile feedback. Furthermore, empirical data shows there is a potential to improve the sleep experience by waking one parent up instead of both like analog baby monitors do. From our empirical studies there were different ways to how people agreed on which parent should get up to take care of the child. In some situations the mother needed to get up because she was still breastfeeding the child. Sometimes the parents' scheduled who should take care of the child or else it is just one parent who primarily does it. In these situation we can see a potential to wake up one parent with the haptic feedback from the smartwatches instead of sound coming from regular baby monitors. Another main point from our study is the situation described in the Empirical Data-section, where the mother pushes the father because he needed to get up. With this system there might be a chance to avoid these kinds of conflicts where people get mad at each other by letting technology decide who takes care of the child.

When participants of the web-based questionnaire were asked about their level of concern for electromagnetic radiation levels were, we used the same term that's often used when promoting baby monitors here in Denmark: *stråling* (Literal translation: radiation). The word on it's own might be a source for concern *radiation* might bring up connotations to nuclear radiation and so forth. It could be the case that the levels of concern would have been levels if the radiation was specified to be of the electromagnetic kind. More empirical research into parents' concern of electromagnetic radiation from smart devices could therefore provide valuable insight into the feasibility of for instance using a smartphone as a baby monitor.
Another concern was that one parent did not want to use the smartwatch when dealing with small children due to a fear of the smartwatch causing scratches. Further research into how widespread this concern is therefor warranted.

Compared to the systems presented in the related work section, the built system prototype further extends the Samsung's "Baby crying detector", as well as changes the connection between the child and caretaker from bluetooth to WiFi, which is likely to be more reliable. Compared to Visit's wrist-worn wearable, our system design has the benefit of adding to potentially pre-owned using general purpose smartwatches and smartphones. Furthermore, the smartwatches used in the system prototype have a charging time of 30 to 60 minutes, compared to Visit's 8 hour charging time.

In field studies, the first and second test families had issues with the prototype due to the amount of devices needed to test the system. The 5 devices needed to test the prototype also needed a charger each. By having 5 devices we excluded the compatibility problems with the smartwatches and the phones, as we only programmed for Android devices. The watches do work with IOS devices, but are limited in functionality compared to Android's support. We also eliminated the process of "on site" installation of third-party applications as errors might happen and people might not want unknown software on their phones. As we discovered this led to confusion among the test families. If the system prototype was to be optimal, the smartwatches would have been their own, and would have been connected to the families own phones.
While having 5 devices is not part of over envisioned system, these issues are definitely something that should be kept in mind when creating future prototypes.

FURTHER WORK

Our research builds upon some assumptions, mainly those that people will be willing to sleep with smart watches in the future, and that sleep tracking can be considered a valid measure of comparable sleep-levels. Furthermore, in our studies we have discovered various questions worthy of further research, as well as design issues in our system prototypes. These assumptions, questions, and design issues are presented here, along with proposed action plans.

From the evaluation and the second questionnaire there concern that people did not wake up from the vibrations arose, such as in the situation from our evaluation where the father did not wake up. The problematics lies in what is going to happen if the parent does not wake up and how we can ensure that one will get up to help the baby. There will always exist a risk of a user not feeling the vibrations from the smartwatch. Possible solutions to accommodate this problem are proposed here. The smartwatch could detect movement in the parent receiving the alarm. If it is detected that the parent is not reacting (e.g. getting out of bed), then the alarm could be sent forwards to the other parent. Another proposition is to wake the other parent after a certain amount of time has elapsed with no reduction to the noise level. In cases where there's only one watch connected to the baby alarm, we propose using alternative non-rhythmic vibration patterns, brightening the screen as much as possible, and using audio output. This is currently not possible on many smartwatches, including the smartwatches used in our system prototype. In such cases connected smartphones could possibly be used as an alternate source of audio output.

Another potential issue is that the initial assumption that parents will be willing to sleep with smartwatches on, as argued for earlier might not hold. Multiple respondents in our questionnaires stated that they don't sleep with their watch on. It would be optimal to test out the different designs of smartwatches to see how it could affect the sleep. Furthermore, it

would be interesting to investigate if a smartwatch would be preferable at all or we should look in another direction.
In fact, this was something the second test family also suggested. Specifically, they enquired about the feasibility of using fitness bands instead. It might be possible to use some of the fitness bands which already have some kind of sleep-tracking tools but our limitation might be the possibilities to install apps on those bands. Fitness bands that can provide vibrotactile feedback, can communicate in real-time, wirelessly with other devices, and can track movements, should be able to realize the proposed concept as well. A similar system using fitness band instead of smartwatches could yield interesting results, as the exclusion of relatively large watch-screens, can potentially more comfortable to sleep with.
The lack of a screen on fitness bands also serves to prolong the battery-life, and (depending on the band design) removes bright lights that some of our field study testers found bothersome. Particularly, the mother in the second family was woken up multiple times the testing night because the fathers smartwatch was lighting up even though *theatre mode* was enabled. We therefore propose that similar field studies using fitness bands instead of smartwatches can provide insight as to whether more people are willing to sleep with such bands, and can also serve to explore potential problems with using fitness bands.

The prototype system had some functionality and interface issues. One such functionality issue was discovered by the third, and final, family. This family experienced the alarm triggering when their child coughed. The cough was apparently loud enough to make the average noise-level in a period of 5 seconds (sampled every 500 ms) go above the alarm-trigger threshold.We propose a possible solution to this. The vibrating could be done in real-time, proportionally to the noise level that exceeds a specified threshold - that is, a cough would only vibrate in one short interval, if at all, depending on the sampling-rate.
The inteface of the baby monitor application was a source of confusion for the second test familiy. The sensitivity could be selected in a drop-down menu (with selections from 2.000 to 20.000). The sensitivity started out as a threshold for our own testing of sound levels. It was later relabeled as a *sensitivity* setting. This is admittedly an error in the design. Intuitively, when something is more sensitive, it should be easier to trigger - in the system prototype, the *sensitivity* was a threshold, so a higher value actually made it harder to trigger (that is, required more noise).
Originally, we considered a *snooze* button for the smartwatch to be a danger, as this would make it easy for parents to just mute the alarm in a groggy state. However, as the third family, which did not normally use a baby monitor neither during the day or the night, expressed wishes for some kind of snoozing ability. If the snooze instead was at the baby monitor the parent would have to get up to snooze the alarm.

The system we designed can not gather measurable data as to whether the parents improved their sleep or not. Due to the design choice of utilizing sleep tracking not being based on any evidence of mobile sleep-tracking providing any fairly comparable sleep-level measurement, and the due to the lack of qualitative testing as to whether the sleep tracking based choice system actually improved the sleep experience of users. Our implementation of the alarm system is very naive as described in the System Prototype Implementation section. While actigraphy (frequently used in sleep-tracking, including by Sleep as Android [12]) has been shown to be provide valid measurements of sleep and wakefulness[5], we have yet to see whether these actigraphs actually provide any reliable comparison between multiple trackings. Further research is very much needed to analyse this.

CONCLUSION

Through the evaluations, discussion, and proposed further work, this paper has attempted provide a foundation upon which further research can be built. A baby monitor system consisting of smartwatches and a device similar to traditional, analog listening device. This system was built on the requirements established from gathered empirical data in the form of web-based questionnaires, an semi-structured interview with an employee at a baby product store, experiments with wake up methods, semi-structured interviews, casual interviews, and existing related works.
Prototypes of the proposed system were built and evaluated through field studies. The parents' within the target demography (that is, parents using baby monitors) concerns are among the discovered challenges. These can be summarized as the following: (1) Multiple parents are sceptical to whether vibrations from a smartwatch can wake them up, (2) electromagnetic radiation is still an important factor in baby monitors and in smart devices, (3) and a possibility that parents' are worried about handling their child when wearing large watches.
Further challenges identified for the use of smartwatches include: bright screens that disturb sleep for some, a lack of comfort (especially when going to sleep), and a lack of female-friendly smartwatch designs. The prototypes of the built system also provided some insight into the difficulty of testing multiple devices within a family. When families are given the entire set-up, including smartwatches and smartphones, the number of devices, coupled with the learning-curve of the basic device operations, sometimes becomes an issue.
We propose that further similar studies utilizing fitness band, or similar wearables, may provide valuable insight into remedies for the previous identified issues. We also propose that further research into whether actigraphy has any face validity providing comparable sleep/wakefulness measurements should be considered. Alternatively, providing a flexible system which allows the users to chose which person should be woken when the baby is crying at specified hourly intervals of the night.
Whether the proposed system indeed improves the sleep experience of the target demography is too early to conclude, our empirical data does however suggest that parents are willing to try smartwatches as an alternative to ordinary baby monitors, and further research has been proposed.

[12] http://sleep.urbandroid.org/documentation/core/sleep-tracking/ *Retrieved 2015-12-29*

REFERENCES

1. Marta E. Cecchinato, Anna L. Cox, and Jon Bird. 2015. Smartwatches: The Good, the Bad and the Ugly?. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '15)*. ACM, New York, NY, USA, 2133–2138. DOI: http://dx.doi.org/10.1145/2702613.2732837

2. Engadget: Jon Fingas. 2014. Samsung's Galaxy S5 doubles as a baby monitor, if you have a Gear watch. Article. (30 March 2014). Retrieved September 18, 2015 from http://www.engadget.com/2014/03/30/galaxy-s5-baby-monitor-feature/.

3. Youn-Kyung Lim, Erik Stolterman, and Josh Tenenberg. 2008. The Anatomy of Prototypes: Prototypes As Filters, Prototypes As Manifestations of Design Ideas. *ACM Trans. Comput.-Hum. Interact.* 15, 2, Article 7 (July 2008), 27 pages. DOI: http://dx.doi.org/10.1145/1375761.1375762

4. T.L. Martin. 2002. Time and time again: parallels in the development of the watch and the wearable computer. In *Wearable Computers, 2002. (ISWC 2002). Proceedings. Sixth International Symposium on*. 5–11. DOI: http://dx.doi.org/10.1109/ISWC.2002.1167212

5. Charles P Pollak, Warren W Tryon, Haikady Nagaraja, and Roger Dzwonczyk. 2001. How accurately does wrist actigraphy identify the states of sleep and wakefulness? *SLEEP-NEW YORK-* 24, 8 (2001), 957–965.

6. Reza Rawassizadeh, Blaine A. Price, and Marian Petre. 2014. Wearables: Has the Age of Smartwatches Finally Arrived? *Commun. ACM* 58, 1 (Dec. 2014), 45–47. DOI: http://dx.doi.org/10.1145/2629633

7. Y. Rogers, H. Sharp, and J. Preece. 2011. *Interaction Design: Beyond Human - Computer Interaction*. Wiley. https://books.google.dk/books?id=b-v_6BeCwwQC

8. Steven Schirra and Frank R. Bentley. 2015. "It's Kind of Like an Extra Screen for My Phone": Understanding Everyday Uses of Consumer Smart Watches. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '15)*. ACM, New York, NY, USA, 2151–2156. DOI: http://dx.doi.org/10.1145/2702613.2732931

APPENDIX

RESOURCE LINKS

(PI1) Phone Interview 1: https://drive.google.com/file/d/0B7_cFEg96cc8TlhDZ3V6Y2tPbk0
(PI2) Phone Interview 2: https://drive.google.com/file/d/0B7_cFEg96cc8ZzlDbWRjYWZOdm8
(PSC1) Prototype screenshots: https://drive.google.com/file/d/0B7_cFEg96cc8NzY0TTFiZzl4Mnc
(PSC2) Prototype screenshots: https://drive.google.com/file/d/0B7_cFEg96cc8bW5rcXVZMzRycG8
(PSC3) Prototype screenshots: https://drive.google.com/file/d/0B7_cFEg96cc8aHJDVWt3Q2FVdUE
(PSC4) Prototype screenshots: https://drive.google.com/file/d/0B7_cFEg96cc8UGZtTnRJTlVNdkE
(V) Study Summary Video: https://drive.google.com/file/d/0B7_cFEg96cc8RVZaM3JjSGQ2ejg