SleepWell: a personal information system for understanding the influence of temperature and humidity on sleep quality

Alexandros Dorodoulis
Technical University of Denmark
Kongens Lyngby
s182148@student.dtu.dk

Margarita Genova
Technical University of Denmark
Kongens Lyngby
s183214@student.dtu.dk

Nermen Ghoniem
Technical University of Denmark
Kongens Lyngby
s173583@student.dtu.dk

ABSTRACT

There are numerous studies proving the significant role sleep environment plays in the individual's sleep quality. Therein, temperature and humidity are key variables in finding the best sleep environment. Current consumer sleep tracking services do not consider temperature and humidity metrics in their analysis of sleep quality. Therefore, this paper will propose a personal informatics system, aiming at informing the users of the importance of sleep environment.

Sleep Well is a mobile application that works in companion with a prototype device to collect the temperature and humidity data of the sleep environment. The mobile application will allow the user the opportunity to enter data about their sleep quality, sleep duration and sleep habits. Additionally, data visualisations will be provided to allow the user to reflect on the importance of the right humidity and temperature values in their sleep environment for their overall sleep quality.

Author Keywords

sleep environment; sleep quality factors; health and quality of life; personal information system; insufficient sleep; sleep tracking; human-computer interaction system

INTRODUCTION (NERMEN)

Sleep is an essential activity for the development of the human organism, as important physiological changes occur during sleep (Chokroverty, 2009). Commonly, Rapid Eye Movement (REM) and the four stages of non-REM are employed to describe the habitual cycle of sleep, wherein the REM density is considered as an index of sleep satiety (Aserinsky, 1969). When humans are interrupted in their journey throughout this sleep cycle, they experience sleep problems affecting negatively their health and quality of life (Kuppermann, M., Lubeck, D. P., Mazonson, P. D., Patrick, D. L., Stewart, A. L., Buesching, D. P., Filer, S. K, 1995).

As traditional health care is facing changes and innovative patient-driven health care services are emerging; patients are becoming active responsibility takers and coordinators of their own health data (Swan, M, 2009). As a result, quantified self-tracking is greatly advertised to consumers to monitor and administer their sleep quality (Jeon, L., Finkelstein, J., 2015)

The bedroom environment has been proven to play a significant role in the individuals sleep quality, with experts providing clear guidelines to achieve the best quality sleep (Kay, M., Choe, E. K., Shepherd, J., Greenstein, B., Watson, N., Consolvo, S., Kientz, J. A., 2012). More specifically, the temperature is presented as a determinant factor for sleep quality, as the thermal environment has an effect on the mechanism regulating sleep (Okamoto-Mizuno, K., & Mizuno, K., 2012). Additionally, humidity has also been proven to have destructive effects on sleep, as it is an important variable for human thermal comfort (Manzar, M. D., Sethi, M., & Hussain, M. E., 2012). Yet, a majority of current services mainly assess the sleep quality by conducting an analysis over its duration, its depth as well as the registered sleep habits such as snoring, nightmares and sleep talking (e.g. Sleep Cycle, Pillow, Sleepzy, Sleep Better). None of the investigated sleep trackers take into account humidity and temperature data, as a factor of sleep environment.

This paper will address the problem by presenting the design, implementation and evaluation of a personal informatics and information management system for sleep tracking. The focus of the mobile application will be to make the user reflect on the impact of temperature and humidity on their sleep quality. The cycle of data acquisition, analysis, visualization, reflection and action will be reported for the complete description of the system (Li et al, 2010). Furthermore, the results of a feasibility study and compelling experiments on the input and data visualisation aspects of the solution led to iterations and refinements of the concept and design, which will be discussed with consideration to the further development of the personal informatics tool.

Problem statement: People are unaware of the impact of temperature and humidity on their sleep quality. Disturbed or insufficient sleep can lead to health problems and poorer quality of life.

Research question: How can a user optimally track and reflect upon the impact of temperature and humidity on their sleep quality?

METHOD

Participants (ALEXANDROS)

Six females and nine males participated in the experiments. The participants were volunteers from varied demographic groups (different age, sexual orientation, race, ethnicity, etc.). Moreover, they all have a different technical background. No participants were excluded in further analyses.

Materials (MARGARITA)

During the design process, the stage-based model (Appendix A) was used and different experiments were performed for the individual stages (Li et al, 2010). During the preparation stage, the Wizard of Oz method was used for validating the problem and identifying the users needs (Rouse, 2014). This concluded that the smartphone device would be required, alongside a customized device, which would collect details about the environment (temperature and humidity). The device was built using a wifi microchip esp8266 (Thaker, T, 2016) connected to a dht22 sensor (Ada, 2012) using a breadboard and wires. The Arduino programming language (Faludi, R., 2010) was used during development. The data collected are stored to a InfluxDB (Naqvi, S. N. Z., Yfantidou, S., Zimnyi, E., 2019) database. The connection between the database and the esp8266 was achieved with an internet API (Robins, N., Lau, V. , 2000), written in the Golang programming language (Pike, 2009).

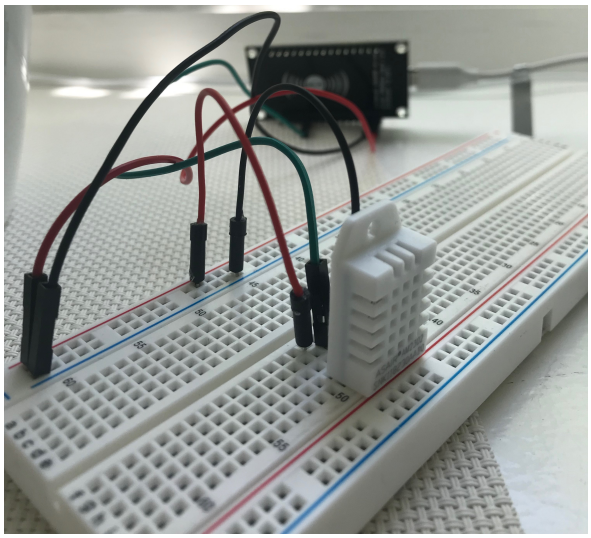


Figure 1. Prototype sensor for humidity and temperate

For the collection and reflection stage, it was decided that the most important aspects of the system to conduct experiments upon would be the input techniques and the data visualisation. Therefore, different wireframes were created on Figma, an interface design application (Bracey, n.d.) to conduct experiments with relevant participants. Three iterations of the design process were conducted with three respondents each, in which the respondents were first asked to perform A/B/C tests to evaluate three different input techniques (Battleson et al., 2001). Followingly, they were interviewed through a retrospective qualitative thinking-aloud usability test (Nielsen, 2012). Thereafter, the same participants were presented to four wireframes for feedback on the data visualisation, for which they were prompted to give feedback through a qualitative thinking-aloud usability test (Nielsen, 2012). The wireframes were presented through Figmas phone simulator on a Lenovo Ideapad 530s laptop in order to give the user the opportunity to interact with the prototype. All the experiments were conducted in a within-subject manner (Loftus & Masson, 1994).

These experiments and its feedback allowed the design and development of wireframes in the high-fidelity prototype, facilitating the planning and coding of the final mobile application. The application was developed using Flutter, a framework for crafting high-quality experience on iOS and Android (*Flutter (n.d)*, n.d.) and the programming language used was Dart (Bracha, 2015). The application was installed on iPhone 7 running iOS 12.1 and was presented individually to each one of the participants in the final interaction. After each experiment, the users were asked to fill in a Google survey to measure their perception of the usability of the system. System Usability Scale (SUS) was used for calculating the quality of the user experience (Nathan, 2015).

Procedure (ALEXANDROS)

As mentioned above, all the participants in the experiments were volunteers - family members or friends, who were interested in the topic. They were separated into five groups - one group for each experiment.

For the first iteration, the idea was presented individually to three potential users. They were asked questions regarding their familiarity with self-tracking devices and their interest in their own sleep quality. As mentioned above, the goal of the experiment was to validate the problem and identify the users needs. The questions asked were predefined before the experiments. Each respondent was requested to express honestly their thoughts. After the experiment, coffee was offered to the participants as compensation. The results are presented in Appendix B.

The focus of the following three iterations was to validate the different input techniques and the data visualisation of their collected statistics. For each iteration, three respondents were part of both experiments. The second iteration was conducted in Helsingr and Holte, whereas the third and fourth experiments were performed at DTU library. For each experiment, the three participants were, first, presented to respectively three wireframes with different input techniques, through a laptop. Then, they were individually asked to perform a series of tasks and verbalize out loud their thoughts throughout the execution of the tasks. Their comprehension test scores and error rate were measured as a result. Followingly, the participants were asked to choose their preferred input technique. Thereafter, the participants were presented to four wireframes with different data visualisations. Again, the respondents were prompted to perform a series of tasks and verbalize out loud their thoughts during the process. These

tasks were measured based on completion time and throughput. Before the beginning of each test, the participants were asked to give their honest opinions about the wireframes presented to them. After the completion of the experiments, the participants were asked if they had any questions and they were thanked for their cooperation. The tasks and results are documented in Appendix C, Appendix D and Appendix E.

The fifth iteration proceeded differently from the previous ones. A mobile device with the application installed was provided individually to three different potential users. The small sample size is due to the lack of time and resources to produce additional prototype devices for the collection of sleep environment data. Each participant was asked to take the device home and interact with it for one night. On the next day, the participants were asked to complete a survey, in which they were prompted to write their thoughts about the system. At the end of the evaluation, each participant received a piece of cake and coffee to show appreciation for their cooperation. The results have been included in Appendix F.

Design and Analysis (MARGARITA)

The data was collected throughout five iterations. Different tasks were performed for each iteration. The goal of the first iteration was to understand what people thought about this kind of solution and to understand what would motivate them to use the sleep tracking application. The following three iterations measured how the users interacted with the system to further understand what they liked or disliked about the solution presented. The final iteration measured how people felt about the system, after having the chance to interact with it for one night. The results from each iteration were analyzed independently.

RESULTS AND DISCUSSION

Results (NERMEN)

The results from the five iterations with, in total, fifteen respondents shed light on the diversity of the individuals interacting with the technology. Especially, feedback regarding the accessibility of the mobile application and the features were highlighted by 50% of the respondents as suggestions for further improvements. Visual improvements in the form of additional feedback and guidelines as well as easily distinguishable colors, icons and fonts were mentioned by all of the respondents.

This resulted in the addition of a getting started tutorial, which would facilitate the users understanding of the preparation stage of the personal informatics system. Familiar graphs and icons from similar or commonly used applications were also implemented into the design of the application. This would facilitate the users understanding of the different features, ensuring user readiness as well as error prevention. Additionally, for the later versions of the mobile application, there was an emphasis on designing for neurodiversity. Visual distractions on the wireframes were minimized to make the content concise and in line with the applications objective. 50% of the respondents found the former layout confusing and messy with regards to the information architecture. Thereby, the reading comprehension of the users was also enhanced. Navigation within the application was also made more intuitive in order to make the user feel informed and more in control throughout the process. The wireframes have been included in Appendix H and Appendix G.

The respondents also went into the specifics concerning what type of content and statistics they would like visualised for them for further reflection. This allowed the inclusion of a sleep diary and additional sub-sections in the input section with regards to the sleep duration, habits and further comments. This also helped redesign the home screen, as the key elements of interest for the users were identified to be different. This resulted in a minimal version of the home screen, in which the current environments temperature and humidity data are presented as well as last nights sleep data.

Overall, the data visualisation saw great improvement as the comprehension test scores improved by 32,5%, and the error rate decreased by approximately 5%, leading to an improvement in accuracy by a similar percentage. Still, there is a room for improvement of the input techniques, like the scores in completion time and throughput scores improved by 15%.

In the last iteration, the focus of the survey was to evaluate the participants overall experience with the system, after interacting with it for one night. The small sample size resulted in different perspectives of the system experience. Still, in general, the participants responded positively to the usability of the system and marked that they would recommend the mobile application to a friend. Nonetheless, there is room for improvements. Some of the participants had difficulties in setting up and interacting with the device for collecting humidity and temperature, which is a problem for the data collection and visualisation in the mobile application. The results have been included in Appendix F.

Discussion (NERMEN + ALEXANDROS)

In light of the results, it is important to discuss the verification of the experiments and the conducted research. The verification index is measured on the basis of validity, reliability and generalization (Kvale, 1997). Validity measures whether the experiments accommodates the goals set for the experiment (ibid.: 233). This was ensured with pilot testings of the usability testings and the evaluation methods with 2 potential respondents at each iteration prior to its presentation to the participants. Still, as seen in some of the comprehension tests scores, some of the tasks were confusing or ambiguous to the participant. Furthermore, in the final summative evaluation, the questions might have been too leading to prompt the participant more directly to provide feedback on certain aspects of the system. Reliability refers to the degree to which the results of the experiment can be reproduced (ibid.: 231). The researchers followed rigid frameworks for the collection and analysis of the qualitative and quantitative data to ensure its reliability. Still, the reliability is exposed to external factors during the thinking-aloud tests, such as the presence of the researchers, which might have distorted the data (ibid.: 226). Longer observational studies would have provided factual results instead of anecdotal results based on personal experience. Nonetheless, the restricted time resources resulted in the choice of method. Finally, generalization indicates the

scalability of the results in terms of individuals (ibid.: 227). The current sample size is not representative and large enough to make statistically significant conclusions. Furthermore, due to restrictions with regards to time and resources, the sampling strategy was chosen to be a convenience sampling to make the most use of available resources (Blaikie, 2003). A random sampling strategy would have increased the generalization of the experiment results (ibid.).

For the evaluation of the different iterations, diverse UX methods could have highlighted different aspects for the improvement of the system. For instance, the AXE method is a qualitative evaluation method, in which participants are shown a visual stimulus to help them reflect upon a user scenario (All About UX, n.d.-a). Thereby, this facilitates the evaluation process for the participants, as they can easily reveal their attitude and values when positioned in a real-life context (ibid.). In this case, visual stimuli related to the bedroom environment could have facilitated the participants understanding of the system and altered their comprehension test scores, error rate and completion time. Another UX method could be the context-aware ESM, which is an experience sampling that prompts the participants to report their experience when the context fulfills the researchers criteria (Intille, 2003). The objective of the system is to make users reflect on their sleep environments impact on their sleep quality, with the hope of improving their sleep. Thereby, in this case, this method could be used to access participants with minor or major sleep problems, as it is of higher relevance to include them in the initial design process.

As mentioned previously, consumer sleep tracking services are increasingly advertised. Some of the most popular include the Apple Health app in combination with the Apple watch, the Xiaomi Mi-Fit in combination with the Xiaomi Mi band alongside the following mobile applications: Sleep Cycle alarm clock, Fitbit, Pillow, Sleepzy, MotionX, Sleep++, Sleep Time, Sleep Better, Beddit and Sleep Tracker+ Adhiya (2019). Currently, the most accurate sleep tracker in terms of data collection is considered to be the Apple Health application in combination with Apple watch (Tang, 2009). Still, none of the aforementioned mobile application track the sleep environment as a factor of sleep quality and only one of them consider sleep habits such as snoring, nightmares, freezing, etc., within their data reflection. Therefore, the proposed application could work in companion to the aforementioned applications, either as a product integration or an additional separate application. Also, the choice of data visualisation is minimal and repetitive amongst the popular sleep trackers, as all of them use basic visualisation such as bar charts. This suggests a lack of experimenting with the data visualisation. The dark theme is also commonly used amongst these apps, which was changed in the proposed mobile application, as respondents gave negative returns on the colors. Finally, as a competitive advantage, the proposed mobile application is the only application, incorporating elements of gamification. The moon and ghost are fun motivators, used consistently throughout the application as mascots to cast reference to respectively, sleep in general and sleep quality.

CONCLUSION (MARGARITA)

This project allowed the creation of a smartphone application that works on both iPhone and Android devices. The Minimum Viable Product (MVP) consisted of a mobile application for sleep tracking with insight into the sleep environments humidity and temperature data through a connecting sensor device. Early validation of the problem helped define the project scope of the MVP and the generation of wireframes for the prototype. Through iterative experiments, the prototype was tested and validated, which led to an improvement of the design material, the human-computer interaction and the information architecture. Moreover, further validations helped narrow down the implemented functionalities. For example, using familiar graphs and icons and adding intuitive navigation improves the users understanding of the system and allows the user to feel more in control.

In conclusion, the project went through five experiments and different validation methods were used such as Wizard of Oz, thinking-out-loud test and online questionnaires. All of the participants had previously used some kind of a self-tracking system to monitor their habits. Still, most of them were unaware of the impact of temperature and humidity on their sleep quality. Therefore, all participants were interested in the solution proposed and would like to participate in future development experiments.

INDIVIDUAL CONTRIBUTIONS

After considerable brainstorming, discussions and validations, the authors of this report have evenly participated and contributed to the creation of this project.

References

Ada, L. (2012). *DHT11, DHT22 and AM2302 Sensors.* Retrieved from `https://learn.adafruit.com/dht/overview`

Adhiya, D. (2019). *Best sleep tracking iphone apps of 2019 for optimal sleep cycles and health.* Retrieved from `https://www.igeeksblog.com/best-sleep-tracker-apps-for-iphone/`

Aserinsky, E. (1969). The maximal capacity for sleep: Rapid eye movement density as an index of sleep satiety. *Biological Psychiatry,*, *1*.

Battleson, B., Booth, A., & Weintrop, J. (2001). Usability testing of an academic library web site: a case study. *The Journal of Academic Librarianship*, *27*(3), 188–198.

Blaikie, N. (2003). *Analyzing quantitative data: From description to explanation.* Sage.

Bracey, K. . (n.d.). *What is figma?* Retrieved from `https://webdesign.tutsplus.com/articles/what-is-figma--cms-32272`

Bracha, G. (2015). *The dart programming language*. Addison-Wesley Professional.

Chokroverty, S. (2009). *Sleep Disorders Medicine: Basic Science, Technical Considerations, and Clinical Aspects. Third Edition.* USA: Saunders Elsevier.

Faludi, R. (2010). Building wireless sensor networks: with ZigBee, XBee, arduino, and processing. *OReilly Media, Inc.*.

Flutter (n.d). (n.d.). Retrieved from `https://flutter.dev/` (Retrieved on May 14, 2019)

Intille, R. J. K. C. A. I. B. L., S. S. (2003). *A context-aware experience sampling tool. in chi'03 extended abstracts on human factors in computing systems*. 1 Cambridge Center, 4FL, Cambridge MA 02142 USA.

Jeon, L., Finkelstein, J. (2015). Consumer sleep tracking devices: a critical review. *Digital Healthcare Empowering Europeans: Proceedings of MIE2015*, 458-460.

Kay, M., Choe, E. K., Shepherd, J., Greenstein, B., Watson, N., Consolvo, S., Kientz, J. A. (2012). Lullaby: a capture & access system for understanding the sleep environment. *In Proceedings of the 2012 ACM conference on ubiquitous computing.*, 226-234.

Kuppermann, M., Lubeck, D. P., Mazonson, P. D., Patrick, D. L., Stewart, A. L., Buesching, D. P., Filer, S. K. (1995). Sleep problems and their correlates in a working population. *Journal of general internal medicine*, *10(1)*, 25-32.

Kvale, S. (1997). *Interview. en introduktion til det kvalitative forskningsinterview*. Hans Reitzels Forlag.

Li et al. (2010). A Stage-Based Model of Personal Informatics Systems. *In Proc. of CHI2010*, 557-566.

Loftus, G. R., & Masson, M. E. (1994). Using confidence intervals in within-subject designs. *Psychonomic bulletin & review*, *1*(4), 476–490.

Manzar, M. D., Sethi, M., & Hussain, M. E. (2012). Humidity and sleep: a review on thermal aspect. *Biological Rhythm Research*, *43(4)*, 439-457.

Naqvi, S. N. Z., Yfantidou, S., Zimnyi, E. (2019). Time Series Databases and InfluxDB. *Studienarbeit, Universit Libre de Bruxelles.*.

Nathan, T. (2015). *How to use the system usability scale (sus) to evaluate the usability of your website.* Retrieved from `https://usabilitygeek.com/how-to-use-the-system-usability-scale-sus-to-evaluate-the-usability-of-your-website/`

Nielsen, J. (2012). *Thinking aloud: The 1 usability tool.* Retrieved from `https://www.nngroup.com/articles/thinking-aloud-the-1-usability-tool/`

Okamoto-Mizuno, K., & Mizuno, K. (2012). Effects of thermal environment on sleep and circadian rhythm. *Journal of physiological anthropology*, *31(1)*, 14.

Pike, R. (2009). The go programming language. *Talk given at Googles Tech Talks*.

Robins, N., Lau, V. . (2000). U.S. Patent No. 6,115,744. *Washington, DC: U.S. Patent and Trademark Office.* .

Rouse, M. (2014). *Wizard of oz prototyping.* Retrieved from `https://searchcio.techtarget.com/definition/Wizard-of-Oz-prototyping`

Swan, M. (2009). Emerging patient-driven health care models: an examination of health social networks, consumer personalized medicine and quantified self-tracking. *International journal of environmental research and public health*, *6(2)*, 492-525.

Tang, L. C. (2009). *U.s. patent application no. 29/325,588.* Retrieved from `https://patentimages.storage.googleapis.com/40/3e/4f/99942fc489e458/USD589375.pdf`

Thaker, T. (2016). ESP8266 based implementation of wireless sensor network with Linux based web-server. *In 2016 Symposium on Colossal Data Analysis and Networking*, 1-5.

Appendices

Appendix A - Stage - based model

<table>
<tr>
<td>Preparation

- Smartphone device
- Smartphone application
- Apple HealthKit or Google Fit
- Custom device
- Optional: smartwatch</td>
<td>Collection

Collection of:
- personal information
- amount of sleep (in hours)
- environment details
- quality of sleep
- date/time
- gender-related conditions (periods)
- amount of coffee/alcohol/cigarettes consumed

Frequency (sampling): collection of user-entered data once a day</td>
<td>Integration

Distance between collection and reflection:
- Long: Monthly statistics on sleep patterns trends after one month of self-tracking
- Short: Daily statistics after one night of sleep

Efforts needed by the user at this stage: Effortless

Maintenance of device: synchronization of data from Apple HealthKit and Google Fit through our server</td>
</tr>
<tr>
<td colspan="2">Reflection

- New understanding of their sleep patterns through visualizing statistics based on user-entered information
- Long-term sleep trends and patterns
- Motivate user to track action through system messages</td>
<td>Action

What to do based on reflection:
- Set new goals (E.g. Change the ideal amount of sleep hours)
- Behavior change (suggestions based on data):
- E.g. change sleep environment, fresh air before sleep, read before sleep, drink more water
- Direct actions:
- E.g. Make an appointment with the doctor for potential symptoms of insomnia or other sleep disorder</td>
</tr>
</table>

APPENDIX B - WIZARD OF OZ - EXPERIMENT 1

Question	Testperson 1	Testperson 2	Testperson 3
Are you using any self-tracking devices? If yes, what do you track?	I do. I am tracking how many calories I am taking per day.	Yes, I am tracking my water consumption.	Yes, I track my daily steps.
Are you happy using the service?	No, not really.	Not really. I get too many notifications per day. Sometimes I get annoyed.	Sometimes I don't understand the statistics available on the application. I get confused about what exactly the specific results are.
What is your goal of using this application?	With the current one I am using, I would like to reduce the amount of calories I obtain per day.	I feel really dry at the end of the day if I haven't drank enough water. That's why I want to be reminded to drink more water during the day.	Nothing specific. I am just curious how many steps I make per day.
If you knew that temperature and humidity play big roles for your sleep quality, would you be interested in using application where you can track your sleep?	I actually heard this claim once. I would say that i might be interested, because I can see the difference after I have slept in cooler environment than in warmer one.	It depends how much effort I have to spend for using it.	If it motivates me enough and I don't have to spend a lot of effort while using it, I would say that I will use it.
How often do you want to measure?	I am not sure. Maybe once or twice per day.	One per day. Either way, I will get annoyed quickly.	Maybe once per day.
What type of data do you think you should put in?	Maybe the temperature and the humidity while I was sleeping. And how much time I have slept.	How much time I have slept. The quality of the sleep maybe.	I would like to see if there is some correlation between the time I have slept and the environment, so I would say how much time I have slept, temperature and humidity.
If you are using the app how long do you want to keep track of your sleep quality?	Maybe several months, so I can be sure that the sleep environment really does affect my sleep quality.	Couple of months. I would say not more than 2 years.	Not more than couple of years.
Do you want the application to provide any analytics?	Yes, of course. I want to see what my progress is.	Yes. I want to see how my sleep changed during the different time periods.	Yes, but nothing fancy. I want it to be as simple as possible. I want to be

			able to see how my sleep quality changed since using the app.
What features would like the application to have?	On the current one I am using, I am not able to change the calories to certain product I have consumed. I would like to be able to change the records.	When I started using the water tracking app, I didn't have any "getting started" tutorial, so I was confused how exactly the application worked at the beginning.	I want to be able to add comments under each record. I may think that something I did during the day, might affect my sleep quality.

APPENDIX C - THINK - ALOUD USABILITY TEST- EXPERIMENT 2

Evaluate different input techniques:

Tasks		Respondent 1	Respondent 2	Respondent 3
Initialize application	Comprehension test scores: 50%	It is nice and simple - you can either go through all the steps or skip them. I'm missing back button. If I want to go back and see a certain step again, I am not able to.	I like it. It is nice and clear about what you need to do in order to use the application.	I like that there is a skip button. It would be nice if I can go back to it after skipping.
Enter sleep data about last night	Comprehension test scores: 30%	Maybe put different icon than the + button. It is nicely placed and it's easy to find. Colors are too similar. I would like to be able to see some chart with my progression on the home page, not only when I demand it.	The colors are too similar. I like that you show the current temperature and humidity. (The user needed more explanation about this task and hints in order to click on the right button.)	It is not very intuitive. I would have clicked on the + sign if I had to update my sleep data or add a new one. Maybe the task is not well assigned.
Enter sleep quality as "OK"	Comprehension test scores: 50%	Intuitive and easy to understand. Don't like the colors. They are too similar.	Maybe make them little icons. Also, change the colors. They are too similar.	I like the bar.

Enter nightmare s as sleep symptoms	Comprehension test scores: 60%	Not sure which picture reflects to what. Maybe write them instead of visualizing them.	Intuitive and easy to understand. Not sure what does “apple health kit suggests..” means. Does it mean that I need to sleep around these hours every night?	I like that you ask me if I had any symptoms before entering them, because I might not have any.
Average comprehension test score:		47.5%		
Error rate:		16%	16%	15%

Evaluate different data visualisation:

	Respondent 1	**Respondent 2**	**Respondent 3**
Age:	26	29	25
Gender:	Female	Male	Male
Occupation:	Nurse	Student	Student
Location of test:	Helsingør	Helsingør	Holte
Date of test:	30.03.2019	30.03.2019	31.03.2019
Time slot of test:	1 pm	1:30 pm	1 pm
Time spent on task 1:	1 min	1 min	.3 sec
Time spent on task 2:	1 min	2 min	1 min
Time spent on task 3:	.3 sec	.3 sec	.3 sec
Time spent on task 4:	1 min	.3 sec	.3 sec
Completion time:	210 sec	240 sec	150 sec
Throughput	45%	40%	60%

APPENDIX D - THINK - ALOUD USABILITY TEST - EXPERIMENT 3

Evaluate different input techniques:

Tasks		Respondent 1	Respondent 2	Respondent 3
Initialize application	Comprehension test scores: 80%	I really like it. It is nice and simple.	I like the idea of having "Skip" button, so I can skip the tutorial. I also like that there are pictures and little text related to it.	I like the colors a lot. It is simple and easy to interact with.
Enter sleep data about last night	Comprehension test scores: 50%	It is hard to understand how exactly I have slept last night. I would like to have an overview of my sleep quality so far. I think "Did you know" section is not necessary in this type of applications.	I think "Current Sleep Environment" takes too much space on the Home Page. It is nice to have it, but maybe not taking that much space.	I don't like the colors and "Did you know section". Colors are too dark and I am using the app to see my progression, not to read different facts about the sleep.
Enter "Ok" as sleep quality	Comprehension test scores: 80%	That was a easy one. I like the slider bar. It is something commonly used and familiar.	I think that the time of sleep should be on the top with the date.It is nice that I have an option to choose different levels of sleep quality.	I like the slide bar. Most of the applications have couple of options and not letting the users to identify their moods with more than 3-4 options.
Enter nightmares as sleep symptoms	Comprehension test scores: 70%	I like that you ask me first If I had any symptoms. I think that you should have a section where I	Colors are too dark. I like that you asked me first if I have any symptoms.	It would be nice if there was a comment section where I can add different comments

		can put some comments in.	I also think that it looks a little bit weird that you have separated the different sections. I am also not able to see clearly how many hours I have slept last night.	related to this specific record.
Avg. comprehension test score:		70%		
Error rate:		15%	12%	13%

Evaluate different data visualisation:

	Respondent 1	**Respondent 2**	**Respondent 3**
Age:	25	24	24
Gender:	Male	Female	Female
Occupation:	Student	Student	Student
Location of test:	Lyngby	Lyngby	Lyngby
Date of test:	06.04.2019	06.04.2019	06.04.2019
Time slot of test:	1 pm	1:30 pm	1 pm
Time spent on task 1:	.4 sec	.3 sec	.3 sec
Time spent on task 2:	.3 sec	.5 s	.4 sec
Time spent on task 3:	.3 sec	.3 sec	.3 sec
Time spent on task 4:	.3 sec	.3 sec	.3 sec
Completion time:	130 sec	140 sec	130 sec
Throughput	65%	60%	65%

APPENDIX E - THINK - ALOUD USABILITY TEST - EXPERIMENT 4

Evaluate different input techniques:

Tasks		Respondent 1	Respondent 2	Respondent 3
Enter sleep data about last night	Comprehension test scores: 80%	I like it better than the previous one. I think I put some useful information on the first screen. It is clear and easy to understand.	I like the colors on the Home Page. It is clear where exactly I should click in order to add data about last night.	It is nice displayed. I get information about how many hours I slept last night, what is the current sleep environment and my sleep quality.
Change your sleep time record	Comprehension test scores: 80%	It is nice that you have displayed both date and time. I feel more in control now.	I like the colors. It is easy to understand.	It is something familiar, so I understood quickly where I have to click.
Enter note about your last sleep	Comprehension test scores: 80%	I like that you can add notes. Sometimes is nice to add more information about a specific record.	It was easy.	-
Save your changes and check your "Sleep Diary"	Comprehension test scores: 80%	It is nice that I can check previous added records.	-	-
Avg. comprehension test score:		80%		
Error rate: 10%		10%	9%	10%

Evaluate different data visualisation:

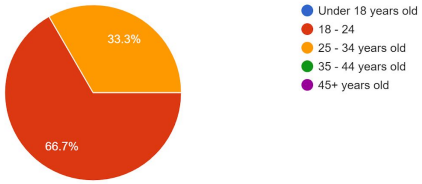
	Respondent 1	Respondent 2	Respondent 3
Age:	25	29	26
Gender:	Male	Male	Male

Occupation:	Student	Student	Developer
Location of test:	Lyngby	Lyngby	Lyngby
Date of test:	13.04.2019	13.04.2019	14.04.2019
Time slot of test:	1 pm	1:30 pm	1 pm
Time spent on task 1:	.3 sec	.3 sec	.3 sec
Time spent on task 2:	.4 sec	.5 s	.4 sec
Time spent on task 3:	.5 sec	.3 sec	.5 sec
Time spent on task 4:	.3 sec	.3 sec	.3 sec
Completion time:	150 sec	140 sec	150 sec
Throughput:	60%	65%	60%

Appendix F - High - Fidelity Prototype - Experiment 5

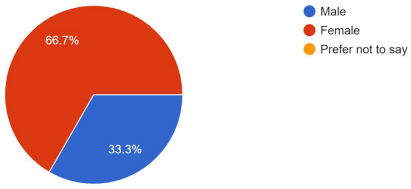
What is your age?

3 responses



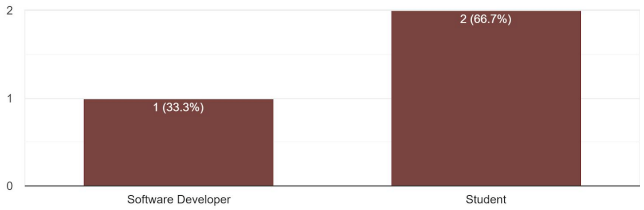
What is your gender?

3 responses



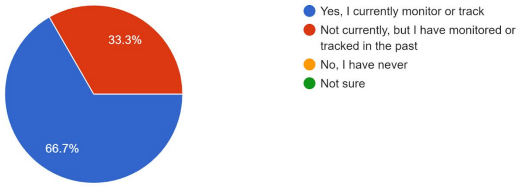
What is your occupation?

3 responses



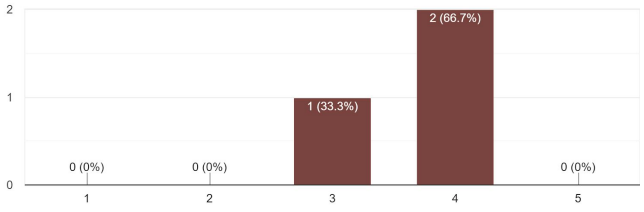
Have you ever used any self-tracking devices?

3 responses

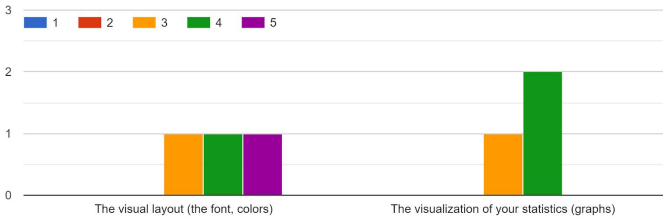


How would you rate your overall experience with the device?

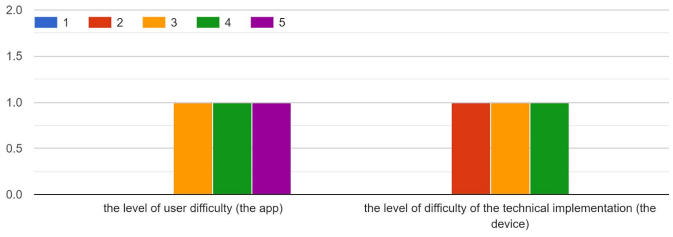
3 responses



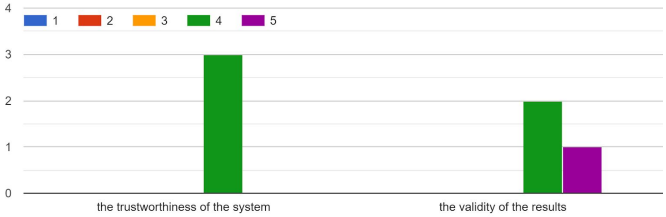
From a scale of 1 - 5, where 1 is Not very good and 5 is Very good, how would you rate:



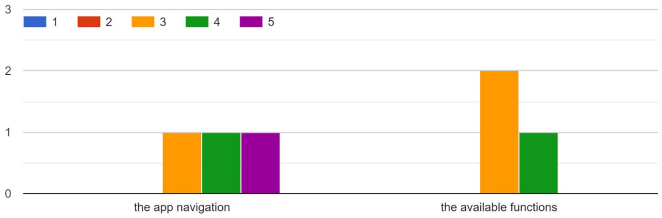
From a scale of 1 - 5, where 1 is difficult and 5 is Easy, how would you rate:



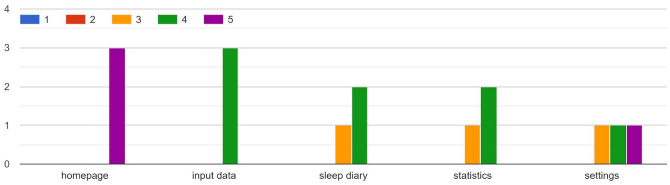
From a scale of 1 - 5, where 1 is Not very good and 5 is Very good, how would you rate:



From a scale of 1 - 5, where 1 is Not very good and 5 is Very good, how would you rate:

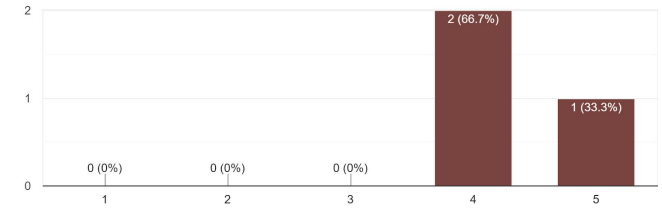


From a scale of 1 - 5, where 1 is Not very good and 5 is Very good, how would you rate the different functions for:



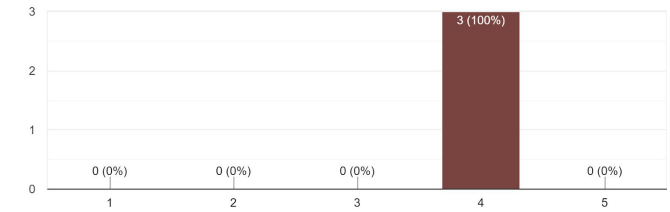
How would you rate your understanding of the system?

3 responses



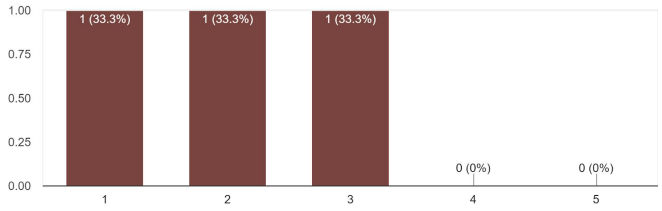
I think that I would like to use the system frequently

3 responses



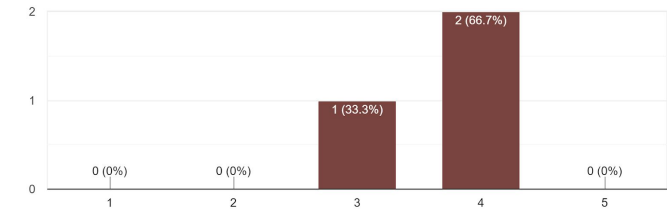
I found the system unnecessarily complex.

3 responses



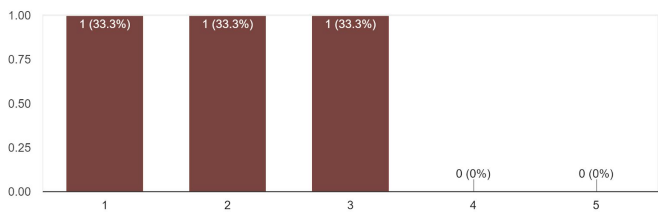
I thought the system was easy to use.

3 responses



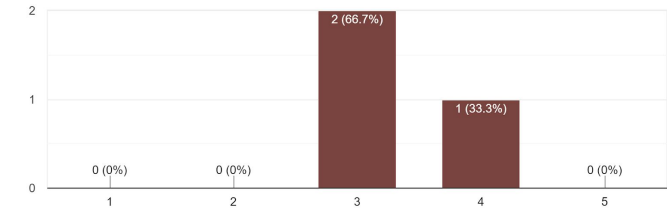
I think that I would need the support of a technical person to be able to use this system.

3 responses



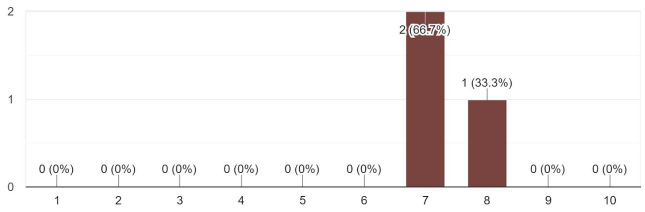
I felt very confident using the system.

3 responses



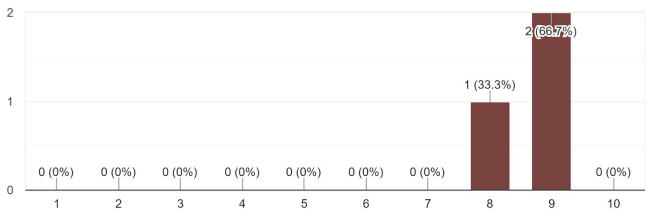
How likely are you to recommend it to a friend, family, colleague?

3 responses



How likely you are to use the system again?

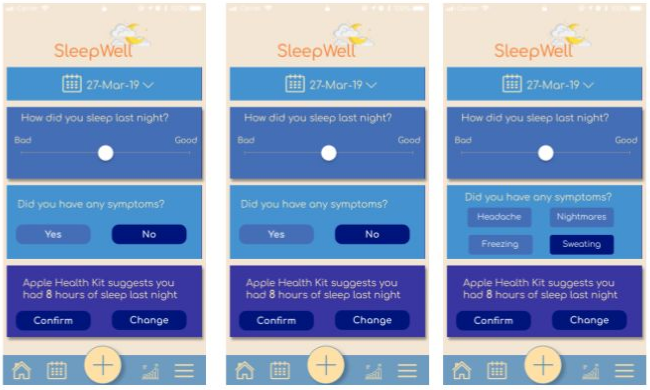
3 responses



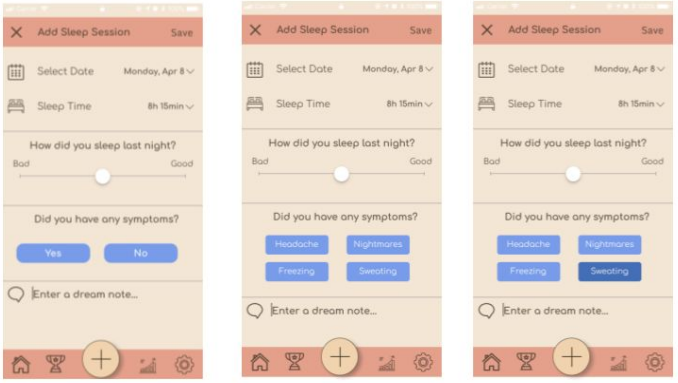
APPENDIX G - WIREFRAMES EVOLUTION - INPUT TECHNIQUES

Input techniques (version 1)

Input techniques (version 2)



Input techniques (version 3)



APPENDIX H - WIREFRAMES EVOLUTION - DATA VISUALISATIONS

Data visualisation (version 1)

Data visualisation (version 2)

Data visualisation (version 3)