Aug 10th, 12:00 AM

Towards a Digital Sleep Diary Standard

Lisa Schmitz
Reykjavik University, lisas@ru.is

Bjarki Freyr Sveinbjarnarson
Reykjavik University, bjarkis@ru.is

Guðni Nathan Gunnarsson
Reykjavik University, gudnig18@ru.is

Ólafur Andri Davíðsson
Reykjavik University, olafurd18@ru.is

Þór Breki Davíðsson
Reykjavik University, thord18@ru.is

See next page for additional authors

Presenter Information

Lisa Schmitz, Bjarki Freyr Sveinbjarnarson, Guðni Nathan Gunnarsson, Ólafur Andri Davíðsson, Þór Breki Davíðsson, Erna Sif Arnardottir, María Óskarsdóttir, and Anna Sigríður Islind

Towards a Digital Sleep Diary Standard

Completed Research

Lisa Schmitz
Reykjavik University
lisas@ru.is

Bjarki Freyr Sveinbjarnarson
Reykjavik University
bjarkis@ru.is

Guðni Nathan Gunnarsson
Reykjavik University
gudnig18@ru.is

Ólafur Andri Davíðsson
Reykjavik University
olafurd18@ru.is

Þór Breki Davíðsson
Reykjavik University
thord18@ru.is

Erna Sif Arnardóttir
Reykjavik University
ernasifa@ru.is

María Óskarsdóttir
Reykjavik University
mariaoskars@ru.is

Anna Sigridur Islind
Reykjavik University
islind@ru.is

Abstract

A sleep diary is an important tool to gather subjective sleep data, which provides key information for the diagnosis of a variety of sleep disorders. In 2012, an expert panel created a standardized sleep diary in pen-and-paper format. However, pen-and-paper has certain limitations, in particular, it is difficult to monitor participant compliance and memory bias. We improve upon these limitations with a digital design and identify benefits and drawbacks of the pen-and-paper format in comparison to a digital sleep diary in an empirical study based on an action design research project. The main contribution consists of five design guidelines: i) use the native environment, ii) utilize established input methods, iii) embed customization to minimize participant workload, iv) evaluate the application continuously using analytics, and v) integrate digital elements to increase compliance. Furthermore, we propose a mobile application design for a digital sleep diary that is in accordance with these guidelines.

Keywords

Consensus Sleep Diary, Digital Sleep Diary, Sleep Assessment, Information Systems, mHealth

Introduction

The digitization of sleep medicine is decades behind the digitization of society in general (Arnardottir et al. 2021). When a person experiences sleep-related issues, a sleep diary is an important tool to gather subjective sleep data, which provides key information for the diagnosis of many sleep disorders (Sateia 2014). A sleep diary can be used for monitoring severe sleep disorders and it can function as a valid assessment tool for insomnia with the potential of identifying candidates for cognitive behavioral therapy as a treatment trajectory (Buysse et al. 2006). A sleep diary is also a valuable research tool for sleep studies. Traditionally, sleep diaries are designed to be printed out and filled in on paper (Åkerstedt et al. 1994; Carney et al. 2012). However, the pen-and-paper format is rigid and neither scalable nor customizable. It provides limited possibilities to interact with the participants, personalize the diary, respond to questions or misunderstandings in time, or control the data during collection. To process and analyze the results, the pen-and-paper requires significant manual work (Văcăreţu et al. 2019), making it prone to errors. Moreover, an analysis of the data can only be provided after the data has been collected and transferred to the researchers (Islind et al. 2019b). In addition to that, sleep patterns can vary due to lifestyle changes, menstrual cycle, shift schedule and day-

light savings, and collecting subjective data over an extended period of time to capture these fluctuations would likely yield more reliable results (Óskarsdóttir et al. 2022). An extended data collection could allow for insight into new factors that influence sleep quality over time.

During the past years, there has been a rise in digital symptom trackers in general (Lavallee et al. 2020) but little progress in digital sleep diaries in particular although the potentials are seemingly high. With a digital sleep diary, data can be directly transferred to and looked at by professionals, creating a bridge between these user groups. Such a connection would enable researchers to react and initiate communication with participants, even during the period of the data collection if needed. Most importantly, digital symptom trackers, such as a digital sleep diary can easily be provided to a broad audience (Lu et al. 2021). There would be no need to restrict the usage and analysis to a defined study period and participants could track their sleep habits for as long as they wish. This way, professionals (researchers and sleep specialists) would be able to collect more data and participants would benefit from a regular sleep assessment and feedback on how to improve their sleep.

This paper presents an action design research project with the aim of digitizing subjective data collection within sleep medicine through a digital sleep diary. We focus on analyzing the ability of a digital sleep diary to address the pen-and-paper format limitations and derive design guidelines for our design and development process of the digital sleep diary. Thus, we arrived at two research questions: *i) What are the benefits and drawbacks of the pen-and-paper format in comparison to a digital sleep diary?* and *ii) How can a digital sleep diary be designed and developed to increase compliance and counteract memory bias?* The main contribution of this paper is a set of design guidelines which target these aspects specifically.

Related work

In 2012, an expert panel created a standardized sleep diary in pen-and-paper format, called the Consensus Sleep Diary (CSD) (Carney et al. 2012), an important work to facilitate comparison across research studies and for clinical work . However, one major drawback of the CSD is that it comes in three different versions: i) a core, ii) an extended core, and iii) an evening version. The core version of the CSD was designed with the intention of creating a minimal viable set of questions for sleep diary research to be asked in the morning. The extended core version additionally tracks daytime data (CSD-M). The evening version was developed specifically to be filled out both in the morning and in the evening (CSD-E) and contains the same questions as the CSD-M with additional instructions on which set of questions to fill out in the morning and which set in the evening (Carney et al. 2012; Tonetti et al. 2016). The pen-and-paper format of the CSD made it necessary to treat each of these three variants as different versions of the sleep diary. However, to advance standardization further, it would be preferable to arrive at one single sleep diary standard. A digital sleep diary version could however allow for time controlled restrictions and due to that, a digital sleep diary could thus combine the three CSD versions into one (Arnardottir et al. 2021).

Monitoring participant compliance in a pen-and-paper design is difficult and counteracting it even more so. Participant feedback has shown that there is a tendency to forget to fill out pen-and-paper symptom trackers and try to make up for it by filling it out at a later time, also referred to as the ”parking lot syndrome” (Cerna et al. 2020). This is a major drawback, since symptom trackers and sleep diaries - when filled out in a close temporal proximity to the occurrence of the event - can contribute reliable and stable data for identifying certain sleep disorders (Arnardottir et al. 2021). For diagnosing sleep disorders, timely data has significant value in comparison to retrospectively collected data. Also, the retroactively entered data cannot be accounted for or filtered out, since it is difficult to detect it in a pen-and-paper format (Tonetti et al. 2016). Entering data retroactively also introduces another problem, namely, it gives rise to the potential problem of memory bias. Memory bias refers to the tendency to intentionally or unintentionally rely on recalling certain events and autobiographical memories and favoring those memories over others (Bianchi et al. 2020). This recall process often relates to significant events, including traumatic, unconventional events, or even to systematically selecting the most recent events in a series of events (Mogg et al. 1987). Memory bias and the tendency to rely on most recent events are especially visible when dealing with a prolonged condition period (Lindroth et al. 2018). Therefore, filling out the sleep diary based on recall can result in potentially biased and less reliable data. This is why finding alternative ways to increase compliance and counteract memory

bias, thus improving the integrity and quality of the sleep diary data, should be a priority (Maich et al. 2018). A digital solution can prevent these issues.

There have been attempts to develop a version of a digital sleep diary. For instance, Tonetti et al. created a digital sleep diary based on the CSD in 2016. It was a mobile application (hereinafter app), developed for the iOS, specifically targeting iPads. They compared their digital sleep diary to the pen-and-paper diary, and received promising results; the digital version performed similarly to the pen-and-paper version. Unfortunately, the study does not detail the design of the app and since it was only developed to support iPads, this digital sleep diary version limits the potential uptake significantly. In 2019, another version of a digital sleep diary was developed by Văcăreţu et al. This version was not based on the CSD, but instead on another analog sleep diary, the Karolinska Sleep Diary (Åkerstedt et al. 1994). This analog sleep diary has a different format compared to the CSD and does not encompass the same set of questions. The research project focused more on studying the usability and user experience of this digital sleep diary and less on the sleep diary per se (Văcăreţu et al. 2019). In 2021, Vallo Hult et al. found that, when transformed appropriately, digital healthcare applications have a positive influence on professional practices as well as patient awareness and participation. Therefore, when turning an analog artifact into a digital one, it is important to take into account possible context changes like the ones described by Vallo Hult et al. (2021). The general need for further standardization of a digital sleep diary is illustrated in an overview of the current state of sleep measuruments by Arnardottir et al. (2021). What can be derived from these papers, is that there should be a three-folded focus in future research on digital sleep diaries: i) to create a standard for digital sleep diaries, ii) to take into account the context and multiple stakeholders, and iii) to design such digital sleep diary to fit multiple mobile platforms to increase potential uptake.

Methods

This paper rests on an action design research project (Sein et al. 2011) with the aim of designing and developing a digital sleep diary as an app (see Figure 1). The app was co-designed, involving end-users (those that will be using the app), researchers and healthcare professionals in the design process. The motivation behind this was to ensure that the respective needs of the different user groups were met (Islind et al. 2019a). For our action design research project we utilized several types of data sources. We used two questionnaires, a survey and application analytics to collect quantitative data and complemented that with the qualitative data that we gained from the semi-structured interviews.

The first out of the two questionnaires is the System Usability Scale (SUS). SUS was used to evaluate the usability of the digital artifact. The questionnaire was developed by Brooke in 1996 and is widely validated. It contains ten rating scales with a score from 0 to 10 and the ratings are added up and compiled to an overall usability score that ranges from 0 to 100 (Brooke 1996). We collected data from 52 participants (alpha prototype: nine participants, beta prototype: ten participants, artifact: 32 participants) in four rounds (alpha prototype: one round, beta prototype: two rounds, artifact: one round) during the entire project. The second questionnaire is the AttrakDiff. AttrakDiff is a questionnaire that was developed to measure experienced hedonic and pragmatic quality and user experience of interactive products (Hassenzahl et al. 2003). It measures four aspects, pragmatic qualities (PQ), hedonic qualities—identity (HQ-I), hedonic qualities—stimulation (HQ-S), and attractiveness (Att). We employed this questionnaire to assess the user experience of the digital artifact. We asked 32 participants to complete the AttrakDiff questionnaire for the digital artifact in two rounds. Additionally, we designed a survey for the purpose of this project to evaluate the alpha prototype. It contained twelve questions and was sent out to and answered by 59 participants. All that data, coupled with the analytics from the application, gave us insights into the use aspect. Furthermore, the semi-structured interviews were individual interviews, ranging from 15-40 minutes with the aim of gaining in-depth insights. In total, we conducted five rounds of semi-structured interviews (alpha prototype: three rounds, beta prototype: one round, artifact: one round) with 55 participants (alpha prototype: 23 participants, beta prototype: ten participants, artifact: 32 participants).

We went through multiple iterations utilizing these mixed method data gathering activities (Byrne and Humble 2007), moving from an alpha prototype in form of a web application to a beta prototype designed as an app which was then turned into a digital artifact. Each user test round followed the setup of the Five-Act

Interview (Knapp et al. 2016), where the participants were asked to make use of the think-aloud technique (speaking freely about what came to mind during their user tests).

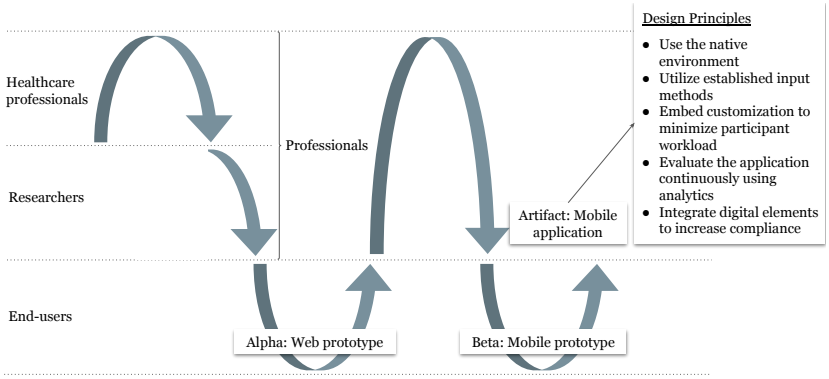


Figure 1. Action Design Research approach.

Alpha prototype

We started the development process by mapping the needs of our three primary user groups: i) end-users, ii) researchers, and iii) healthcare professionals. In total, nine participants took part in a semi-structured interview, three from each user group. Following the results of this interview round, researchers and healthcare professionals were treated as one group (hereinafter professionals) in the subsequent phases of the research project, as their needs turned out to be almost identical. Next, we handed out a pen-and-paper format of the CSD-E to 59 end-users. We collected their pen-and-paper format sleep diaries and evaluated their experience after filling out the CSD-E for one week. Our goal was to identify the key needs to consider and incorporate into a digital sleep diary. After the end-users had completed that task, we asked them to fill out a survey regarding their experience and how they would have like to have their data visualized.

Based on the results from this phase, we developed the alpha prototype as a web application. The initial designs were printed out on paper as low-fidelity wireframes and were tested with nine participants: three professionals and six end-users. They were asked to complete a list of tasks and do a think-aloud test using the prototype. From the feedback, we created an overview of the issues and prioritized according to their severity. The prototype was then updated in order to improve these critical issues. Subsequently, another evaluation was conducted for the refined version with the same participants using a high-fidelity prototype. In addition to completing tasks with the prototype, the same nine participants answered the SUS questionnaire to measure the usability of the alpha prototype.

Beta prototype

We decided to develop the beta prototype for a digital sleep diary in the form of a mobile application with optional notifications, because reminders like those in a typical mobile application notification style were a strongly requested feature by participants in the study phase with the alpha prototype. In order to digitize the CSD and turn it into an app, we held a workshop with sleep professionals, a clinical sleep specialist, clinical sleep researcher and an expert in patient-reported outcome measures, in the development process Based on the expert feedback, we developed a low-fidelity prototype and tested it with 5 end-users. We asked them to complete a list of tasks using the prototype. Afterwards, the participants filled out the SUS questionnaire followed by a semi-structured interview. Based on those results, we developed a high-fidelity prototype. We repeated the user tests in the same fashion as before with the same five end-users that tested the low-fidelity prototype, asking them to complete tasks, answer a SUS questionnaire and partake in semi-structured interviews.

Artifact

Next, we developed the digital sleep diary as a functional mobile application for Android and iOS. This prototype was handed out to 32 end-users that were divided into two groups. The first group consisted of 17 Reykjavik University students, which all had previous experience with pen-and-paper sleep diaries, whereas the second group consisted of twelve external end-users of which only one person had prior experience with filling out sleep diaries. The first group received an on-site introduction to the mobile application, while the second group received instructions via email. Before the end-users of both groups got to use the application, they were asked to fill out the AttrakDiff questionnaire. For the following week, the participants were asked to fill out the diary daily, both in the morning and evening. The end-users were able to give continuous feedback throughout this phase, which let to minor bug fixes regarding the text translation and notification cancellation. Additionally, we logged the end-users' activity in the application. After the end-users had completed the week of filling out the sleep diary, they were given the SUS and AttrakDiff questionnaires to answer. To conclude the study, we conducted semi-structured interviews with these end-users who had used the digital sleep diary.

Results

We will now present the findings from user tests with the alpha prototype, the beta prototype and the artifact.

Alpha prototype

The results of the final survey, which was conducted after the end-users had filled out the pen-and-paper CSD for a week, showed that 81.4% of the end-users preferred the sleep diary in form of a website instead of on paper. Overall, the majority of the end-users had a positive attitude towards the sleep diaries, almost 60% found it "entertaining" or "interesting", while more than 20% found it "okay". However, more than 17% had a negative experience while filling out the sleep diary as shown in Figure 2 and 4.6% of the end-users did not complete their pen-and-paper sleep diary.

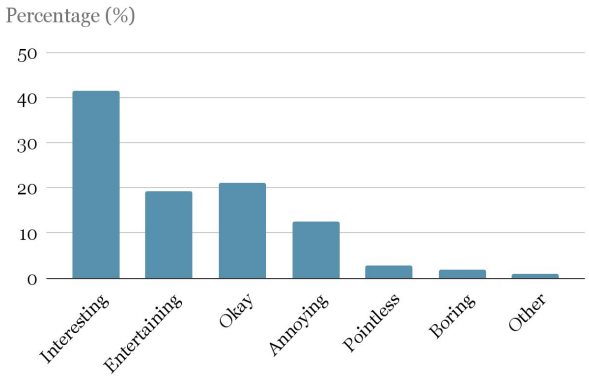
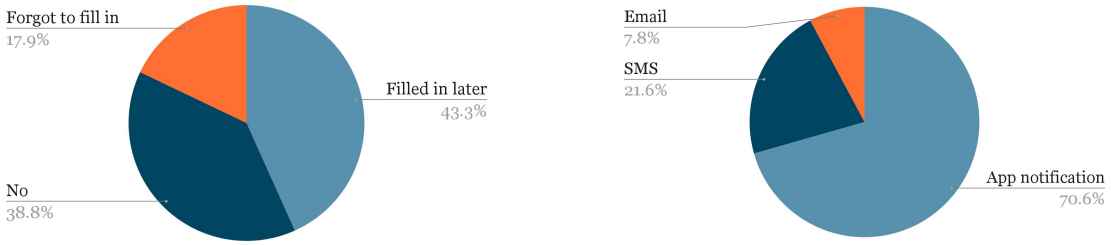


Figure 2. Survey answers to: "How did you find filling in the sleep diary?"

Over half of the end-users reported trouble remembering to fill out the sleep diary in time. As shown in Figure 3 (a), more than 40% filled in the sleep diary later on more than one occasion and almost 18% forgot to fill out the diary completely at least once. One important topic for over 80% of the end-users was that they wanted to be able to receive reminders and over 70% preferred to receive those in form of native app notifications (see Figure 3 (b)).

Beta prototype

In a workshop with professionals, we derived requirements for the context change of the sleep diary from pen-and-paper format to an app, these requirements were then included in a new beta prototype. We decided to use native iOS and Android input elements to collect the sleep diary answers. Furthermore, we



(a) "Did you ever forget to write in your sleep diary?" **(b)** "How would you like to receive reminders?"

Figure 3. Survey answers.

agreed to include the instructions for answering the evening questions on a separate instruction view. This view is accessible through an information button positioned at the top right corner of the screen 5. We also allowed the participants to fill out entries only for the current day, and not for any of the previous or following days. In this phase we also worked on the requirement of notifications. The participants had the option to receive notifications in the morning and in the evening at a chosen time. By default, the notification times were set to 9 AM and 9 PM. However, we derived the need for this to be customizable (which was implemented in the following phase).

We had to make changes to some of the CSD-E questions in order to keep them understandable in the new context. All changes to the sleep diary questions are shown in Table 1. Follow-up questions like 12b and 13b are displayed isolated on the screen like all the other questions, which is why they needed clarifying updates. Conditional questions like 6d were simply omitted if the condition was not fulfilled. Question 14 expects two different answer inputs which is why we decided to separate it into two questions in order to keep the diary entry layout consistent.Questions 16-20 were added as a five point visual analog scale for additional information on end-user daytime functioning. Additionally, we asked the end-users in the onboarding of the beta prototype if they drink alcohol or caffeine, or take sleep medication. This was a need derived from our interaction with the end-users on the one hand and the professionals on the other hand. The end-users can adjust the answers to these questions later on in the application settings if needed. If one of these setup questions is negated by the end-user, the corresponding diary questions are not displayed. This way the end-users do not need to answer questions that are irrelevant to them. Since the beta prototype is supposed to be used shortly before the end-users go to sleep, we tested both light and dark theme and found out that the dark theme was preferred, with the argument of it being good to reduce the light exposure through the app as much as possible moments before sleep. All of the aspects that worked well during user testing in this phase, were fully implemented in the artifact phase that followed. The beta prototype went through two rounds of user tests. In the first round, the prototype achieved a mean SUS score of 91.0, with a minimum score of 85.0 and a maximum score of 100.0. After we incorporated the user feedback from this round, the mean SUS score of the next iteration of the beta prototype improved to a mean score of 97.5, with a minimum score of 85.0 and a maximum of 100.0. Based on those results, we moved on to the artifact phase.

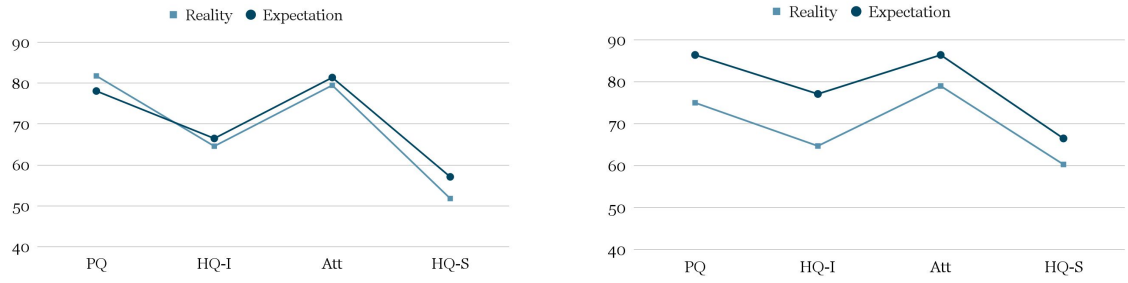
Artifact

The data collected from the user tests with the digital artifact shows that 81.4% of the end-users preferred to use an app rather than a pen-and-paper format. The participants gave reasons for this in the interview; they found the digital sleep diary more engaging, accessible, and easier to comply with than pen-and-paper. On average, the morning questions of the digital sleep diary were answered 6.66 out of 7 times and the evening part was filled out 6.47 out of 7 times. It took the average end-user 397 sec (SD = 2449) to fill out the morning sleep diary, with a median of 121 sec. The evening sleep diary was filled out in an average of 75.4 sec (SD = 57.2), with a median of 61 sec. Since there were occurrences of outliers, most likely caused by opening the diary in the evening and then keeping it open until the morning, the median time gives a more accurate estimate than the average. 22 out of the 32 end-users filled out the SUS questionnaire that was handed out during this phase, eleven participants of the first group and eleven end-users from the second group.

Question	CSD	Our App
6d	If yes, how much earlier?	How much earlier?
12b	What time was your last drink?	What time was your last alcoholic drink?
13b	What time was your last drink?	What time was your last caffeinated drink?
14	Did you take any over-the-counter or prescription medication(s) to help you sleep? If so, list medication(s), dose, and time	a) Did you take any over-the-counter or prescription medication(s) to help you sleep? b) List medication(s), dose and time taken
15	-	Was today a workday? (If not, then it was a freeday)
16	-	How fatigued were you today?
17	-	How sleepy were you today?
18	-	How stressed were you today?
19	-	How was your mood today?
20	-	How much did you exercise?

Table 1. Changes to the CSD-E questions.

Overall, the answers resulted in a mean SUS score of 88.9 (SD = 9.0). In the first group the mean SUS score was 85.2 (SD = 9.6), while it reached a mean of 92.5 (SD = 8.4) in the second group. Furthermore, we used AttrakDiff to determine the digital sleep diary's user experience compared to end-users' expectations.



(a) First group **(b)** Second group

Figure 4. Results from the AttrakDiff questionnaires.

Figure 4 shows the calculated mean values for each AttrakDiff scale among the participants. The first AttrakDiff questionnaire was filled out by all 32 participants, whereas only 21 of the participants chose to answer the second one. The digital sleep diary received the lowest score for the hedonic quality for stimulation (HQ-S), meaning that the users did not find the app very stimulating but other aspects were well received. In the interviews, the end-users commented positively on the setup of the sleep diary, finding it easy to understand and learn. They also mentioned that it was quicker to fill out than the pen-and-paper version. The notifications were generally appreciated and praised frequently by the end-users. They found that it helped to remember to fill out the diary and even described them as "motivating". However, some difficulties came up when filling out the diaries. The end-users found it hard to estimate times and sometimes did not understand the questions. Unfortunately, they often overlooked the instruction button which could have helped to avoid this issue. Figure 5 shows selected screens of the digital sleep diary including the onboarding, the diary and the data visualization screens.

Discussion

In this paper we address two areas of interests. Firstly, we focus on the main benefits and drawbacks of the pen-and-paper format in comparison to the digital sleep diary. The findings from that part outline our requirements. Secondly, we use those requirements to design and develop a digital sleep diary with the aim

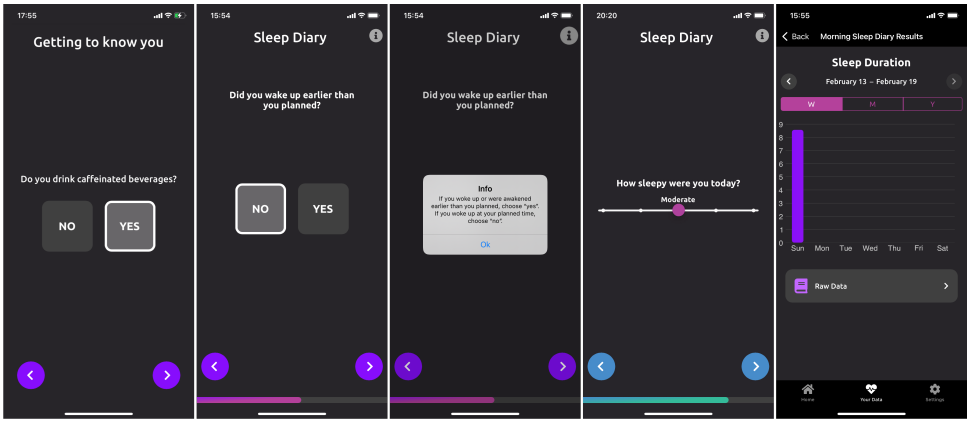


Figure 5. Selected views of the digital sleep diary in the developed mobile application.

of increasing compliance and counteracting memory bias, which were the main issues found while using the pen-and-paper version. We found that on a piece of paper, it cannot be controlled when and how participants enter information. Examples include various time formats that are used to enter data, wrongly entered data and unprompted comments (Văcăreţu et al. 2019). We found that these issues can be avoided with input control in a digital sleep diary. Choosing native inputs for time, numbers, booleans (true/false) or text prevents the participant from entering incorrect data and ensures data homogeneity.

Our interviews with the participants showed that they were interested in their data and liked to see it visualized. Based on that, we theorize that insightful data visualizations combined with actionable feedback can further increase compliance and engagement. Instead of only collecting data with the sleep diary, a visualization of the sleep diary data can offer an analysis and feedback to the end-users. In this way, the sleep diary can be transformed from a pure data collection device to a corrective tool with the potential of continuously improving the sleep hygiene of the end-users. A digital companion app also has the potential to gather additional data through, for example, cognitive tests and can potentially include data from wearables, such as fitness and tracking devices. This data can be used to complement data collected through the diary, automatically fill out relevant sections of the sleep diary and evaluate the accuracy of the sleep diary data (Arnardottir et al. 2021). Along those lines, the CSD paper by Carney et al., shows the value in comparing the collected data with data from a medically valid sleep assessment such as polysomnography data. Even though it cannot be expected that a sleep diary can lead to similar conclusions as a polysomnography, it could still be insightful to compare subjective sleep quality with an objective sleep score. Therefore one of the next research steps is a clinical validation of the final version of the digital sleep diary in comparison to polysomnography data. This can also allow us to identify inconsequential or redundant questions and thus further narrow the core of the sleep diary questionnaire.

The results of the user experience tests show that the vast majority of participants prefer a digital sleep diary over a pen-and-paper format. However, the participants were comparatively young and some of them were even students of computer science. A certain bias towards digital artifacts is probably reflected in the results. Nevertheless, since the results were clearly in favour of the digital sleep diary, it can be concluded that the general tendency towards a preference of a digital sleep diary is still substantial. However, particularly among older participants, the sleep diary in analog format was sometimes preferred over a digital sleep diary. Moreover, it turned out to be a good decision to move from a web application to an app. Considering that the notifications were mentioned as well liked by a significant number of participants during the evaluation of the digital sleep diary, we want to highlight them as one of the key features of any digital symptom tracker. By utilizing a notification system to remind participants to fill out the digital sleep diary, we can influence them far better than with a pen-and-paper version. Notifications can be set up at individual times for each participant, thereby tending to different daily routines and minimizing the participants' workload. Most importantly though, a digital sleep diary provides options of restricting the times that the diary is filled out and makes it possible to log the exact date and time of a data entry, hence, accounting for memory bias,

which is clearly shown in our results.

The low score that the mobile application achieved in the HQ-S category of the AttrakDiff (see Figure 4) stand outs as the most negative aspect of this version of a digital sleep diary. It agrees with the feedback that we received from participants that filled out a pen-and-paper diary: more than 38% did not enjoy filling out the sleep diary (end-users who found filling out the diary "okay", "annoying", "pointless" or "boring"). Therefore, the digital sleep diary still has similar issues with user engagement as the pen-and-paper version. This can be explained with the fact that the sleep diaries, which the participants tested, functioned first and foremost as a clinical sleep assessment tool. The current design does not offer any additional value to the participants. However, the participants have shown interest in their data and augmenting the data visualization with actionable feedback could be a valuable addition to the mobile application. Based on our combined findings in this paper, we summarize our main contribution into five design guidelines for designing and developing a digital symptom trackers in general and for designing and developing digital sleep diaries in particular:

1. **Use the native environment**: It is important to use the advantages that come with a native app design. Native input control and notifications are clear advantages from the digital format. However, it is important to evaluate changes that are made to the original question setup to ensure that they do not affect the way a participant answers the questions.
2. **Utilize established input methods**: It is important to use native control elements instead of coming up with new input ways. Users have shown to be most comfortable with familiar input methods.
3. **Embed customization to minimize participant workload**: It is important to utilize customization and onboarding to omit questions that are not relevant to particular participants. Since feedback indicated that filling out the symptom tracker is not enjoyable enough for every participant, it is even more important to avoid any unnecessary questions.
4. **Evaluate the application continuously using analytics**: It is important to embed analytics to enable continuous design and development. Moreover, this type of data can be collected easily and does not require work from the participants.
5. **Integrate digital elements to increase compliance**: It is important to integrate gamification elements for the content and include more insightful data visualization, which would improve the entertainment value of the app since not all participants enjoyed filling out the diary. Some of them mentioned that notifications increased their motivation to fill out the symptom tracker.

Conclusions

In this paper we show that the pen-and-paper version of the sleep diary has limitations and a digital sleep diary can enable closer monitoring of participant compliance and decrease memory bias. From our findings we derive five design guidelines for the design of digital symptom trackers in general and digital sleep diaries in particular. The five design guidelines are the following: i) use the native environment, ii) utilize established input methods, iii) embed customization to minimize participant workload, iv) evaluate the application continuously using analytics, and v) integrate digital elements to increase compliance. Moreover, we find that the key advantages of moving the sleep diary from a pen-and-paper format to a digital sleep diary stem from the ability to interact with the participants through notifications and custom application content. This is a valuable feature for increasing participant engagement which ultimately leads to more reliable data collection. We also show that the notifications have proven to be a useful and desirable element. Based on that, future work could include further examining and optimizing nudging and its effects on users. For example, testing different phrasing of notifications and how they effect user behaviour could be one way to improve compliance to a greater extent. Also, gamification elements could increase the attractiveness of tasks that are considered to be monotonous now. In addition to making the diary as accessible as possible, it is also important to make it fun for the end-users, to increase compliance.

Acknowledgements

We thank Bergrós Pálmadóttir Morthens, and Birta Líf Baldursdóttir for their contributions. This research is a part of the Sleep Revolution project, with funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 965417. Senior author: Anna Sigridur Islind.

References

Åkerstedt, T., Hume, K., Minors, D., and Waterhouse, J. (1994). "The subjective meaning of good sleep, an intraindividual approach using the Karolinska Sleep Diary," *Perceptual and motor skills* (79:1).

Arnardottir, E. S., Islind, A. S., and Óskarsdóttir, M. (2021). "The Future of Sleep Measurements: A Review and Perspective," *Sleep medicine clinics* (16:3), pp. 447–464.

Bianchi, R., Laurent, E., Schonfeld, I. S., Bietti, L. M., and Mayor, E. (2020). "Memory bias toward emotional information in burnout and depression," *Journal of Health Psychology* (25:10-11), pp. 1567–1575.

Brooke, J. (1996). "Sus: a "quick and dirty'usability," *Usability evaluation in industry* (189:3).

Buysse, D. J., Ancoli-Israel, S., Edinger, J. D., Lichstein, K. L., and Morin, C. M. (2006). "Recommendations for a standard research assessment of insomnia," *Sleep* (29:9), pp. 1155–1173.

Byrne, J. and Humble, Á. M. (2007). "An introduction to mixed method research," *Atlantic research centre for family-work issues* (1), pp. 1–4.

Carney, C. E., Buysse, D. J., Ancoli-Israel, S., Edinger, J. D., Krystal, A. D., Lichstein, K. L., and Morin, C. M. (2012). "The consensus sleep diary: standardizing prospective sleep self-monitoring," *Sleep* (35:2).

Cerna, K., Grisot, M., Islind, A. S., Lindroth, T., Lundin, J., and Steineck, G. (2020). "Changing Categorical Work in Healthcare: the Use of Patient-Generated Health Data in Cancer Rehabilitation," *Computer Supported Cooperative Work (CSCW)* (29:5), pp. 563–586.

Hassenzahl, M., Burmester, M., and Koller, F. (2003). "AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität," in *Mensch & computer 2003,* Springer.

Islind, A. S., Lindroth, T., Lundin, J., and Steineck, G. (2019a). "Co-designing a digital platform with boundary objects: bringing together heterogeneous users in healthcare," *Health and Technology* (9:4), pp. 425–438.

Islind, A. S., Lindroth, T., Lundin, J., and Steineck, G. (2019b). "Shift in translations: Data work with patient-generated health data in clinical practice," *Health informatics journal* (25:3), pp. 577–586.

Knapp, J., Zeratsky, J., and Kowitz, B. (2016). *Sprint: How to solve big problems and test new ideas in just five days,* Simon and Schuster.

Lavallee, D. C., Lee, J. R., Austin, E., Bloch, R., Lawrence, S. O., McCall, D., Munson, S. A., Nery-Hurwit, M. B., and Amtmann, D. (2020). "mHealth and patient generated health data: stakeholder perspectives on opportunities and barriers for transforming healthcare," *Mhealth* (6).

Lindroth, T., Islind, A. S., Steineck, G., and Lundin, J. (2018). "From narratives to numbers: data work and patient-generated health data in consultations," in *Building Continents of Knowledge in Oceans of Data: The Future of Co-Created eHealth,* IOS Press, pp. 491–495.

Lu, D. J., Girgis, M., David, J. M., Chung, E. M., Atkins, K. M., and Kamrava, M. (2021). "Evaluation of mobile health applications to track patient-reported outcomes for oncology patients: a systematic review," *Advances in radiation oncology* (6:1), p. 100576.

Maich, K. H., Lachowski, A. M., and Carney, C. E. (2018). "Psychometric properties of the consensus sleep diary in those with insomnia disorder," *Behavioral sleep medicine* (16:2), pp. 117–134.

Mogg, K., Mathews, A., and Weinman, J. (1987). "Memory bias in clinical anxiety." *Journal of abnormal psychology* (96:2), p. 94.

Óskarsdóttir, M., Islind, A. S., August, E., Arnardóttir, E. S., Patou, F., Maier, A. M., et al. (2022). "Importance of Getting Enough Sleep and Daily Activity Data to Assess Variability: Longitudinal Observational Study," *JMIR Formative Research* (6:2), e31807.

Sateia, M. J. (2014). "International classification of sleep disorders," *Chest* (146:5), pp. 1387–1394.

Sein, M. K., Henfridsson, O., Purao, S., Rossi, M., and Lindgren, R. (2011). "Action design research," *MIS quarterly* (), pp. 37–56.

Tonetti, L., Mingozzi, R., and Natale, V. (2016). "Comparison between paper and electronic sleep diary," *Biological Rhythm Research* (47:5), pp. 743–753.

Văcăreţu, T., Batalas, N., Erten-Uyumaz, B., Van Gilst, M., Overeem, S., and Markopoulos, P. (2019). "Subjective sleep quality monitoring with the hypnos digital sleep diary: Evaluation of usability and user experience," in *12th International Conference on Health Informatics, HEALTHINF 2019,* SCITEPRESS-Science and Technology Publications, Lda., pp. 113–122.

Vallo Hult, H., Islind, A. S., and Norström, L. (2021). "Reconfiguring professionalism in digital work," *Systems, Signs & Actions* (12), pp. 1–17.