INTERNATIONAL DESIGN CONFERENCE – DESIGN 2024
https://doi.org/10.1017/pds.2024.229

Designers' perceptions of a sensor-enabled diary method for enhancing user research

Yuki Taoka ✉, Tomoyuki Tanaka, Momoko Nakatani and Shigeki Saito

Tokyo Institute of Technology, Japan

✉ taoka.y.aa@m.titech.ac.jp

Abstract

This study proposes a diary method enabled with IoT sensors for user research in design. It addresses the limitations of diary methods by incorporating sensor data to trigger user self-reports. The focus is on how sensor data influences self-reports and designers' perceptions. Results show that sensor-enabled diaries offer more diverse content and overview of users' lives and designers perceived the proposed method potentials, suggesting significant potential for IoT in user research.

Keywords: user interview, diary method, design methods, data-driven design, internet of things (IoT)

1. Introduction

In new products and service developments, empathic user research plays a significant role in gaining a deep understanding of users and contexts (Mattelmäki *et al.*, 2014). User research has long relied on qualitative methods to identify latent user needs that users are not aware of (van Boeijen *et al.*, 2014). The qualitative methods are powerful tools to collect information on internal factors of users, such as motivation and feelings through interviews and objective information on users' behavior through observation (Bae *et al.*, 2015). However, the qualitative user research methods also face limitations such as users' recall bias (Norman, 2013) and limited opportunities for observations (van Boeijen *et al.*, 2014). Thus, methods for collecting users' data in users' contexts are needed (Visser *et al.*, 2005). The recent development in information and communication technologies (ICT) enabled us to record large amounts of behavioral data. Behavioral data is defined as *"a collection of specific information, referring to data from sensors, self-logging, telemetry, or social networks which capture people's behaviors and patterns"* (Gomez Ortega *et al.*, 2022). Behavioral data may overcome the limitation of qualitative user research methods because behavioral data is collected in users' environments through the Internet of Things (IoT), wearable devices, and social media. Thus, our objective is to explore the potential of behavioral data for user research.

Behavioral data are widely used for users to objectively reflect on their behavior in the area of personal healthcare management. However, the use of behavioral data in design research is still limited (Bertoni, 2020; Yasuoka, Nakatani, *et al.*, 2023). The majority of existing approaches using data, for example, through Data-Driven Design (DDD) studies, used big data, e.g., social media and online reviews of commercially available products, as data sources (Bertoni, 2020; Yasuoka, Nakatani, *et al.*, 2023). A few examples include biological data (Ishio and Abe, 2017), pictures (Carter and Mankoff, 2005), video recordings (Arvola *et al.*, 2017), use of smart speaker (Gorkovenko *et al.*, 2019), and use of baby bottle (Bogers *et al.*, 2016). The studies suggest that behavioral data can be used to explore users' contexts in user research (Gorkovenko *et al.*, 2023). Camera-based methods face ethical concerns because users cannot know what is logged beforehand. Sensor data may be ethically more acceptable, while data itself has less information than camera-based methods. Behavioral data can objectively acquire and graphically visualize users' specific behavior, while a single sensor can only collect limited users' behavior and lacks holistic visualization of users' behavior (Tanaka *et al.*, 2022). Gorkovenko *et al.* (2023) tested the behavioral data from multiple devices and pointed out that qualitative data from users is also essential to elicit insights on users when making sense of behavioral data. Given that sensor data is time series data, qualitative information should also be obtained in time series. The diary method, a traditional qualitative method, might be most appropriate in the context of behavioral data. Diary methods let participants respond to questions at spontaneous or pre-designated times (Carter and Mankoff, 2005), and the participants' memories are fresh and accurate in the collected data. Traditional diary methods also have limitations. For example, frequently filling in diaries is a considerable burden for the participants. The burden decreases the response rate and the quality of the answers and causes the participant to omit their behavior from the reports (Baxter *et al.*, 2015). Behavior data could mitigate the need for more holistic data in diary methods because it can detect users' behavior regardless of users' intentions.

In sum, the combination of a diary method and behavioral data from multiple sensors has the potential to be a powerful user research method. We developed a diary method that uses sensor data to trigger users' self-reports. The acquired materials from users are used for post-data collection interviews by designers. The proposed method was experimentally validated by comparing it with the traditional voluntary self-reports in terms of designers' perceptions of data and the contents of self-reports. This paper aims to investigate the following two questions on the proposed user research method integrating qualitative diary and behavioral data through sensors embedded in users' houses.

Q1: How do data from sensors embedded in users' living environments influence users' self-reports?

Q2: How do designers perceive the usefulness of sensor data and self-reports for user research?

2. Method

2.1. Proposed sensor-enabled diary method

We proposed a sensor-enabled diary method that collects users' self-reports triggered by sensors embedded in users' homes. Figure 1-left shows IoT sensors attached to the user's kitchen appliances. The installed devices consist of an iPhone and sensors. The iPhone was installed to collect signals from sensors. The sensors were small IoT devices (MESH, Sony Group Corporation), each 24mm x 48mm x 20mm in size, and a single-board computer, ESP32, with magnetic sensors. We used three types of MESH sensors: brightness, motion, and button. In the proposed method, researchers rather than designers conducted data collection. A pre-defined message asking for self-report, "tell us the use of *[a device triggered the message],"* was semi-automatically sent to users following a predefined algorithm during data collection, as shown in Figure 1-right. The notifications were sent to an iPhone installed by the researcher via Slack, and the tones were played. Then, the users reported a voice message via the Slack app from their own smartphone. The installed iPhone was used to ensure the notification ring, and the user's phone was used for reporting for the user's convenience. The users were instructed to send a

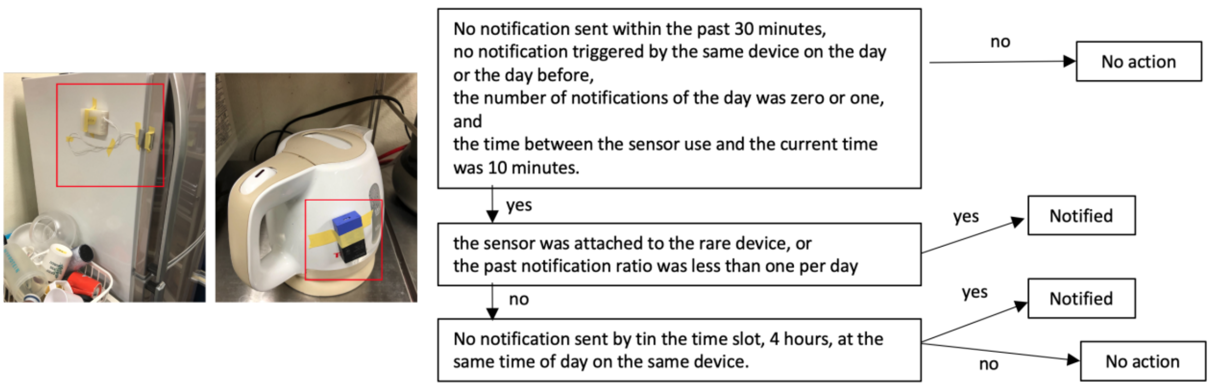

Figure 1. (left) sensors attached to kitchen appliances; (right) algorithm for sending notification

https://doi.org/10.1017/pds.2024.229 Published online by Cambridge University Press

one-minute voice report within ten minutes after each notification. Following-up questions to users' responses were not sent to avoid the influence of the researcher's bias. The algorithm used in this study aims to spread out the devices and time used for notifications. The number of daily reports was targeted to be one or two to avoid heavy burdens for responding. The day was divided into six time*frames*, each lasting 4 hours; in Figure 1-right, rare devices were selected based on the sensor responses of the first four days. Any devices that were not used within the first four days were assigned as unusual devices. The self-report consisted of three aspects: behavior, thought, and emotion. Behavior referred to what the user did or was doing; thought referred to reasons for reported behavior or behavior before and/or after the reported behavior, and emotion referred to positive and/or negative emotions about the behavior. Positive emotion was explicitly included to avoid collecting only complaints.

2.2. Experiment overview

This study compared the sensor-enabled diary method with a traditional diary method based on voluntary responses with ten user and designer participants each. The professional designer performed a user research task, "improving the kitchen life of students living alone", under two experimental conditions: a sensor-enabled diary condition (sensor condition) and a voluntary diary condition (voluntary condition). Each condition consisted of two weeks of collecting diaries and a 2-hours post-data collection interview. Figure 2 shows a session timeline of the data collection and the post-data collection interview. The user participants acted as user research subjects who collected sensor data at their homes and recorded diaries. Before each user's first data collection, the user participant sent a picture of their kitchen and their basic information on their lives, such as transportation to schools, ages, and duration of living in the current house. The designer participants attended the post-data collection interview as interviewers. Designer participants, three males and seven females, had professional experience in user interviews. The user participants were graduate or undergraduate students who lived alone and cooked for themselves at least twice a week, eight males and two females, ages 21-24. We adopted an experimental order that considered counterbalance; half of the participants conducted the sensor condition in their first session, and the other did the voluntary condition first. In addition, two measures were taken to minimize the impact of order effects: the interval between sessions was at least one month, and the same combination of user and designer was avoided in the second session. The participants received 12,000 JPY as a reward for participating in the experiment.

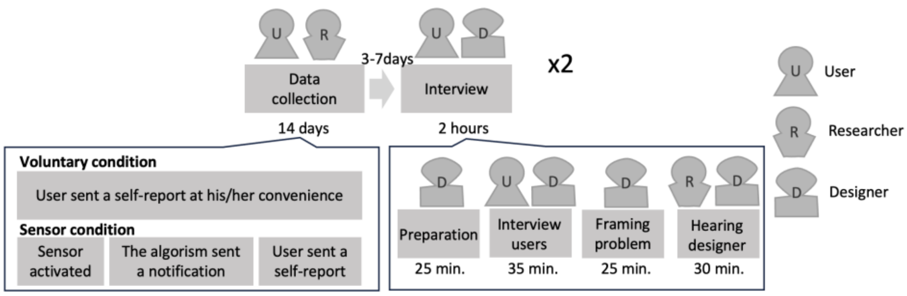

Figure 2. An overview of each session consisting of data collection and interview

In the interview session, the following task was given to designer participants: *"Develop an understanding of the user's life and explore his/her needs and design opportunities in order to come up with ideas to improve people's kitchen lives."* The interview consisted of four steps. At the preparation step, the designer planned the interviews with given materials and wrote down how they used self-reports and sensor data, and the interview plan. The materials included a basic profile of the user, a picture of the user's kitchen, and self-reports in both conditions, as shown in Figure 3. Graphs of sensor data were also included in the sensor condition. All the materials were printed out and sent to designers beforehand. The designers were instructed not to open the envelopes before the session began. No instructions were given on how to utilize the sensor or self-reported data. At the user interview step, Each designer conducted online user interviews based on the data gathered during data collection period. The user did not have access to the materials presented to the designer during the interview. Following the user interview, the designer filled out a worksheet with interesting findings and user needs and the

https://doi.org/10.1017/pds.2024.229 Published online by Cambridge University Press

next steps of user research based on the information obtained from the interviews. Finally, the researcher conducted a questionnaire and hearing with the designer. Post-experiment questionnaire and hearing aimed at understanding designers' perceived usefulness of the user-collected materials, which were diary and sensor data. In the post-experiment hearing of the second session, we also asked designers to compare their experience in both conditions. The questionnaires had two questions about perceived usefulness and interestingness on a seven-point Likert scale.

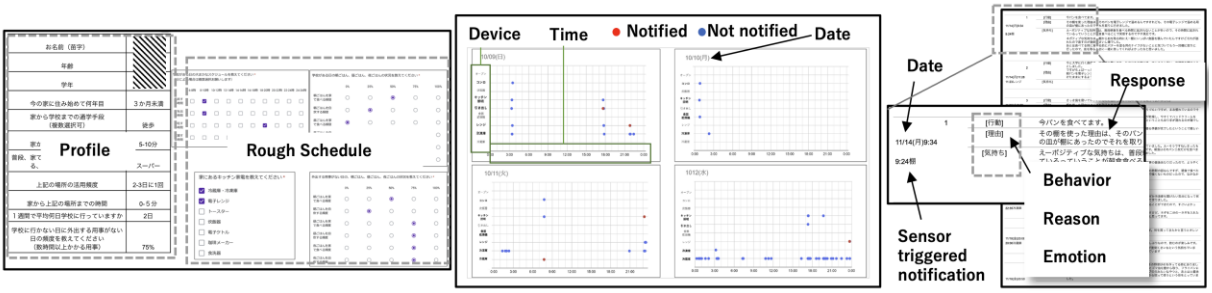

Figure 3. Example of printed materials sent to designers in the sensor condition: (left) user profile and rough schedule of users; (middle) sensor graphs; and (right) a list of self-reports

The self-report contents during data collection had the same structure in sensor and voluntary conditions. The participants should have recorded three aspects separately: behaviors, thoughts, and positive and negative emotions. In both conditions, the participants orally reported using the Slack audio clip function.

The sensor-enabled diary condition: The method proposed in the previous sub-chapter was used. We installed as many sensors as possible at user participants' houses. An average of 7.7 (minimum: 6, maximum: 9) sensors were installed. All homes had sensors installed in the refrigerator (door of freezer and refrigerator), microwave, and on shelve(s) in the kitchen. Table 1 shows other objects with sensors installed.

Table 1. List of kitchen appliances and objects which sensors were attached to

Objects	Lightning	Stove	Rice cooker	Kettle	Toaster	Coffee Machine	Dishwasher
Counts	8	8	6	5	2	1	1

The voluntary diary condition: The user participants were asked to send a voice recording once or twice daily at their convenience. There was no daily request for a report during data collection. The recording contents were instructed to include experiences around the kitchen and kitchen appliances, including before and after meals and other non-cooking times at home. Considering the possibility that user participants may have forgotten to report it, a reminder was sent when more than two days had elapsed since the last report.

2.3. Assessment of the experiment

The data collection was evaluated by quantitatively analyzing the time the self-reports were sent and the amount of the self-reports. The number of self-reports was counted, and the mean response frequency per two weeks of data collection was calculated. The number of Japanese characters in each transcribed message was counted to evaluate the length of the self-reports, and the mean and standard deviation were calculated. The Wilcoxon signed-rank test statistically tested the differences. The self-report contents on behavior were categorized by open coding to evaluate varieties of the contents. We also quantitatively analyzed sensor data by counting how often the sensor responded per day and which sensors triggered what kind of self-report contents to evaluate the sensor condition. The visualization of when the sensor responded in a day, sensor graphs, was also analyzed by counting the number of clusters of sensors' responses. The two sensors were defined as a single cluster if a sensor reacted within an hour after another sensor, as shown in Figure 4.
The hearing with designer participants was transcribed. The contents of the hearing transcripts were qualitatively analyzed to explore how designers perceived and used the sensor data. All qualitative analyses were conducted by one author and confirmed by another author.

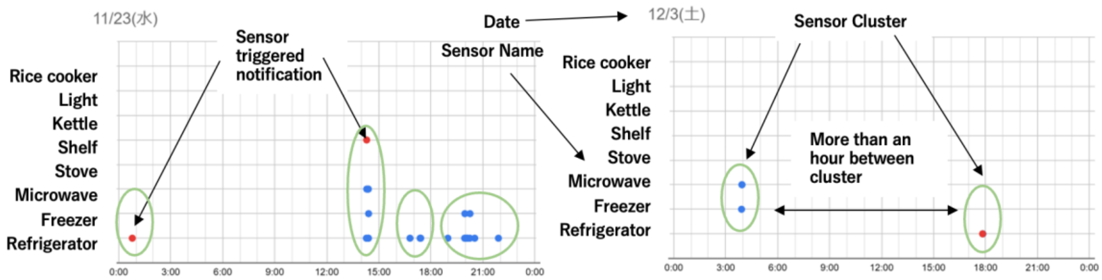

Figure 4. Typical days with lots of or a few sensor data and the concept of sensor cluster

3. Result

3.1. Overview of self-reports and sensor data

Table 2 shows an overview of the user-generated data during the data collection. The total number of reports per person shows how often a participant sent reports in two weeks. The participants reported in sensor conditions more often than in voluntary conditions, which was significantly different, p=0.025, by the Wilcoxon signed-rank test. The higher standard deviation in the voluntary condition shows the variance in the number of reports depending on the participants. The mean characters per report show the number of Japanese characters included in each report, and 200 characters in Japanese is roughly equivalent to 80 words in English, which were not significantly different, p=0.33. It shows that the reported sentences have similar lengths regardless of conditions. In the sensor condition, sensors responded in 11.6 days on average, and the number of times the sensor responded per day was 11.44 on average. Each sensor graph had 2.3 sensor clusters on average. Figure 5 shows when the user participants sent reports and how the sensor reacted. The vertical line shows the total number of reports, and each color shows each participant's value. 60% of reports were sent from 21:00 to 3:00 in the voluntary condition (Figure 5 left), while 50 % of reports were sent in the sensor condition (Figure 5 right). Figure 6-left shows when sensors responded throughout the day, assuming users were out during the daytime. Figure 6-right and Table 3 show the categorized behavior reported in the self-reports. Table 3 shows quotes from the self-reports and the number of codes reported in each experiment condition. In Figure 7, the numbers in bar charts show the number of self-reports categorized into each code. Cooking occupies 66 % of the reported reports in the voluntary but covers about 34 % in the sensor condition. The seven categories covered 95% of the voluntary condition but 88% of the sensor condition.

Table 2. The mean and standard deviation of user-generated self-reports and sensor data

	Total reports in two weeks	Characters/ Report	Dates of sensor responded	Activated Sensor / Day	Sensor cluster/Graph
Sensor	18.1(3.45)	215.5(61.0)	11.6 (3.00)	11.44 (5.17)	2.30 (0.74)
Voluntary	13.2(5.31)	226.7(59.0)	--	--	--

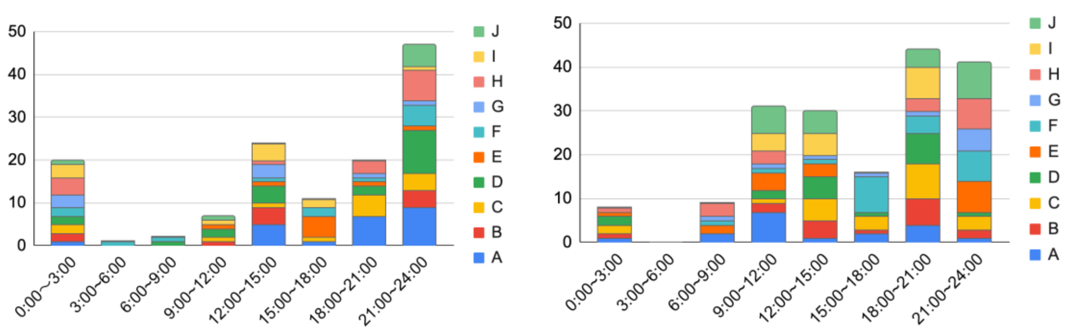

Figure 5. Time of self-reports was sent: (left) in the voluntary condition; (right) in the sensor condition.

https://doi.org/10.1017/pds.2024.229 Published online by Cambridge University Press

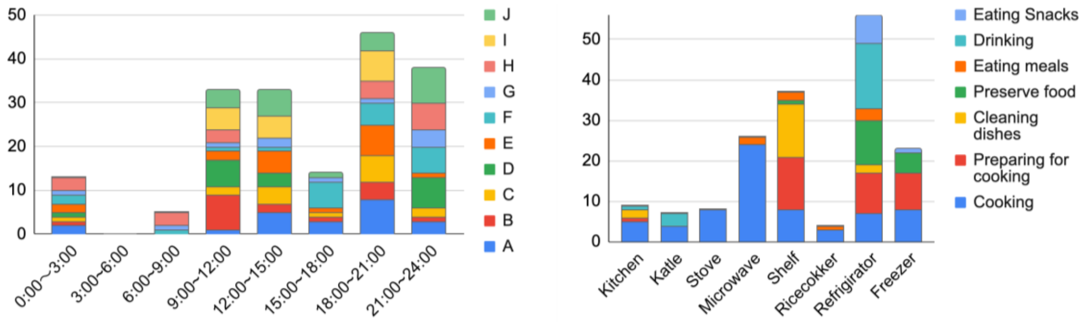

Figure 6. (left) The result of time when the sensors activated; (right) which sensor elicited what kind of self-reports

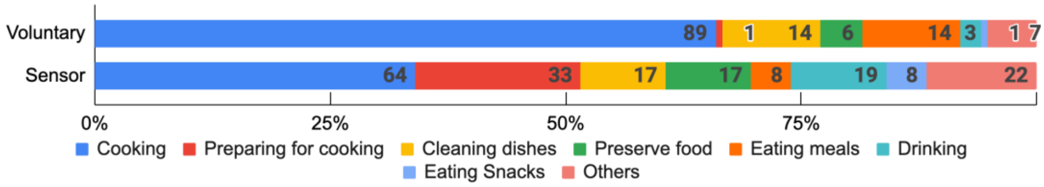

Figure 7. The frequency of the self-reports and sensor-enabled self-reports

Table 3. The categories of behaviors reported

Code	Example
Cooking	Now I'm making sandwiches as a snack because I had a sudden idea.<User I>
Preparing for cooking	I used the top shelf to get a bowl for melting eggs.<User A>
Cleaning dishes	There was a little leftover washing up from yesterday, so I washed it up.<User A>
Preserve food	There was a little prepared food, which I had left out. I put it in the fridge.<User A>
Eating meals	Now, after having yakisoba noodles as supper.<User B>
Drinking	I just took a 2L plastic bottle of tea out of the refrigerator and drank it.<User F>
Eating Snacks	I decided to eat a banana, so I opened the refrigerator.<User H>

3.2. Designers' perception on sensor data

Table 4 shows the result of user-generated sensor data and self-report on the sensor condition and the designer's perception of sensor data. The user-generated data were counted the same way as the result reported in section 3.1. The use of the sensor graphs depends on the designers' self-report during the post-experiment hearing. The average questionnaire scores calculated each participant's scores on two questions about the sensor data's usefulness and interest. The amount of user-generated sensor data varies depending on the user participants. The number of sensor responses per day ranged from 3.14 to 22.29, and the number of self-reports ranged from 11 to 23. The average Japanese characters per report also varied: the shortest was 133.55, and the longest was 324.76. Table 5 shows that two designers did not use sensor graphs during preparation for the interview. Designer B explained the reason was a lack of preparation time, while designer D did not find a good way of using the data. Five designers, including one of those who did not use the sensor data, pointed out the lack of time for preparation and interview. The amount of user-generated data may not have a direct link to designers' use of sensor graphs.

As two designers did not report using the sensor graph, the following focus on those who used sensor graphs. Table 5 and Table 6 show the classified results of the interview. In both tables, the count column shows the number of designers who made statements classified into each code, and the alphabet corresponds to the designers' code in Table 5. Table 6 shows the reported benefits of using the sensor data. Table 5 shows that the designers tried to infer users' behavioral tendencies from the sensor graphs. Four participants reported that they could successfully identify users' unconscious behavior or behavior that did not appear in the self-reports. Three designers found the data useful because it provided basic user information. Three designers found data objectivity useful because voluntary self-reports may not accurately represent users' behavior, but the sensor data were acquired automatically.

https://doi.org/10.1017/pds.2024.229 Published online by Cambridge University Press

Table 4. The relations between user-generated sensor and self-report data and designers' sensor graphs use and questionnaire results

Designers	User-generated sensor data and self-report				Designers perception	
	Sensor responded / Day	Days of Sensor responded	Sensor triggered reports	Mean characters / reports	Use of Sensor Graph	Questionnaire mean scores
A	11.71	12	17	209.94	Yes	3.5
B	8.43	12	16	246.94	No	4
C	6.79	10	22	161.36	Yes	6.5
D	11.64	10	18	175.17	No	2
E	14.29	14	17	324.76	Yes	5
F	22.29	14	22	250.14	Yes	6.5
G	3.14	9	11	133.55	Yes	5
H	9.21	12	19	210.26	Yes	3
I	12.36	11	17	284.47	Yes	5
J	14.5	12	23	158.43	Yes	6

Table 5. The usage of sensor data and sensor data triggered self-reports

Code and *Quote*	Count
Infer the user's behavioral tendencies - *"You can tell that he was doing various things in the kitchen at night. Although it's unclear exactly what he was doing, their actions were definitely reflected in the data, which was very helpful."* <designer J>	8
Find users' unconscious behavior and night-time behavior that does not appear in self-reports - *"It is possible to understand behaviors that users do not report. For example, the user opened and closed the refrigerator several times late at night"* <designer I>	4
Reduce interview time by knowing basic information in advance - *"In a normal interview, it takes about an hour to get a daily schedule from morning to night. (This time, we had the data.) The time saving part is also amazing."* <designer A>	3
Sensor data and self-reports can be treated as objective data - *"The fact that the refrigerator and microwave oven were utilized the most was something that was definitely left as sensor data, so I was able to talk with the interviewee with confidence."* <designer G>	3

Table 6. How the designers look at the sensor graph

Code and *Quote*	Count
Find patterns from the data - *"You can understand the pattern of behavior. For instance, whether the person is using the kettle or frequently using the refrigerator. From this, you can guess whether the person is busy or anticipate their behavior patterns."* <designer E>	8
Try to decipher sensor data by comparing it to self-reports -*"I have a habit of looking at odd or regularities. I scraped off the things that I could explore that factor in the report."* <Designer C>	5
Compare with another day's sensor data - *"I use to just look at the data and compare it to other days."* <Designer H>	3
See areas where sensor data is concentrated - *"Much data was gathered at 9:00 to 12:00 p.m., after he ate dinner. I wondered what he was doing."* <designer I>	2

Table 6 shows how the designers looked at the sensor graphs. All designers tried to find behavioral patterns from sensor data. Five designers used the sensor data by linking the sensor graphs to the self-reports. They tried to find self-reports that could describe sensor data that they were interested in. Three designers compared a sensor graph of a day with sensor graphs of other days. Analyzing the sensor graphs and self-reports required significant effort: *"In the case of sensor data, it was indeed necessary to mutually examine the sensor data and self-reports. While this required considerable physical and mental effort, it also provided various insights. However, it also made me realize that a significant level of skill was required." <Designer G>*

https://doi.org/10.1017/pds.2024.229 Published online by Cambridge University Press

4. Discussion

4.1. Sensor data's influence on the self-reports

In response to Q1, we discuss the sensor data and self-reports collected during data collection. Table 4 and Figures 5-7 show that the behavior data through embedded sensors visualized at least part of users' lives. Even a graph without sensor data is evidence that the user did not use the kitchen and the user actually had nothing to report.

More diversified self-reports in the sensor condition: Table 2 and Figure 5 show that sensor notifications increased the number of self-reports and diversities of time reports were sent but kept the same length of self-reports. The contents of self-reports became more diversified in the sensor condition than in the voluntary condition, as shown in Figure 7 and Table 3. These data suggest that the proposed sensor-enabled self-report method created better self-report content than the voluntary condition in exploratory user research. Real-time interactions in the reported situation would further enhance understanding of specific user scenarios (Bae *et al.*, 2015). The result also implies that the users could handle heavier burdens of reporting in sensor conditions because the amount of total reports was increased in sensor conditions. Future studies need to assess participants' impressions of the reporting duties.

Locations for sensors to be attached: Figure 6-right shows refrigerators and freezers sent 48 % of the self-reports, and refrigerators triggered the most varieties of self-reports contents. Interestingly, sensors attached to shelves elicited 37 % of the self-reports having diversified contents, while sensors on microwaves mostly drew self-reports on eating meals. In an experiment where subjects of user research utilized sensor data for recalling memories afterward, the frequently used objects are not good sources for recalling memories (Tanaka *et al.*, 2022). The results suggest that frequently used objects are good sensor data sources in asking for self-reports in the users' lives. The result clarified what type of sensor would elicit a wide variety of comments in the context of user research in kitchen. The results imply that attaching sensors to frequent objects might be good enough to diversify the variety of self-reports. Future research should look at what kind of data increases the diversities of self-reporting in other user research contexts.

4.2. Potentials of sensor-enabled diary method

In response to Q2, designers perceptions were analyzed. The hearing result indicated the potentials.

Objectivity of data: Table 5 shows that designers valued the objectivity of sensor data. The designer who did not use the graph even said that the self-reports became more reliable as the designer was certain that users did something at the time of reports. Table 6 shows that designers used sensor data to understand behavioral tendencies. As in Table 6, grasping users' overall lives was the first thing designers did in the interview. The designers' strategy is aligned with the designers' steps of building empathy, immersing themselves into users' worlds (Kouprie and Visser, 2009).

Capturing unconscious users' behavior and routines: Table 5 shows that designers found the unconscious behavior of users. The sensor graph may have been a beneficial tool because learning the unconscious parts of users' lives that the users themselves are not aware of leads to designers' deeper insights (van Boeijen *et al.*, 2014). Designers participants also reported that they could have found users' behavioral routines through the sensor graphs, which is otherwise challenging to objectively find users' routines because observation and video monitoring usually take too much time (Arvola *et al.*, 2017). User research in design emphasizes capturing users' behavioral patterns and contextual inquiry (Hanington, 2007). The traditional diary method could provide contextual inquiry, but it needs ways to collect what has been missed during data collection. The mix of sensor graphs and the self-reports may become a powerful tool for designers because the result indicated that the sensor graphs provided data for behavioral patterns, and self-reports offered materials for contextual inquiry. This study did not look into the details of the self-reports and the designers' interviews with the users. Future studies should deeply investigate the self-reports and interview transcripts to verify the existence and effect of perceived user behavior in the sensor graphs.

https://doi.org/10.1017/pds.2024.229 Published online by Cambridge University Press

4.3. Challenges of sensor-enabled diary method

This study also revealed the challenges of integrating sensor graphs and self-reports.

Data visualization: In this study, sensor graphs and self-reports were presented on paper rather than in dynamic forms that might have allowed designers to look at both data interactively. As data visualization influences one's perception and understanding of one's behavior (Oh and Lee, 2015), static data representation may limit designers' capability to use data. The data visualization method is an interesting research agenda.

Sensitization of users' behavior data: The results showed that two designer participants did not use the sensor graphs because they could not realize the benefits during interview preparation. Since using data in user research is still an emerging topic, methods of increasing designers' data literacy are also emerging topics in design education (Yi Min Lim *et al.*, 2021). The sensitization session of quantitative data to designers before experiments might have changed the experiment result. How data literacy methods can facilitate designers' use of users' behavioral data is a topic for future research.

Sensor might need more time for interview preparation and data collection: Five designers reported not having enough time to go through collected data in preparation and interviews. It suggests that processing sensor data and self-reports might require more time for preparation. The lack of time may have limited the potential of the proposed method. This study spent two weeks on data collection, which may be short for finding behavior patterns. The duration of data collection may influence the data collection strategy and participants' awareness of data collection, which is another future research topic.

Ethically acceptable method: Another concern is the ethical considerations and users' perception of sensor data collection. This study did not analyze that aspect because it did not analyze interviews with user participants after data collection. As this study took an experimental approach, the users consented to install sensors into their lives beforehand. However, communicating all sensor data to designers may not be widely accepted. Practically, users should have the right to choose what raw data to share with designers as researched and reported in the form of data donation (Yasuoka, Miyata, *et al.*, 2023). Instead of sharing all raw data with users, designers may benefit from users' own reflections based on the collected data, as has been done on video data (Arvola *et al.*, 2017). Long-term reflection by accumulating data is also valuable because it allows users to understand their behavior patterns and trends they have not known (Li *et al.*, 2010). Future research should look at how personal data can benefit user research in an acceptable manner to users.

5. Conclusion

This study proposed and tested a sensor-enabled diary method for user research. The method's effectiveness was evaluated by analyzing designers' perceptions of behavioral data and by comparing the contents of self-reports in two conditions: voluntary and sensor-enabled conditions. The results show that the sensor-enabled diary had more diversified contents and timing in the self-reports, while the user's burdens could have increased. The results suggest that designers perceived that they could identify users' behavioral tendencies and patterns from sensor data and self-reports when eliciting users' latent needs. The challenges for designers to utilize data included data visualization and education on how to use data. Considering recent rapid developments in IoT and wearable devices, this study experimentally demonstrated the potential of integrating qualitative data and quantitative IoT sensor data embedded in user environments for user research.

Acknowledgment

This research was supported by JSPS KAKENHI Grant Number 23K17155 and dLab challenge, Tokyo Institute of Technology. The authors would like to thank a master's student, Kyohei Otani, for his support in the analysis.

References

Arvola, M., Blomkvist, J. and Wahlman, F. (2017), "Lifelogging in User Experience Research: Supporting Recall and Improving Data Richness", *The Design Journal*, Routledge, Vol. 20 No. sup1, pp. S3954–S3965, https://dx.doi.org/10.1080/14606925.2017.1352898.

Bae, J., Cho, K., Kim, C. and Others. (2015), "Developing a framework of new mixed method, social networking services group diary and its application in practice", *DS 80-11 Proceedings of the 20th International*

Conference on Engineering Design (ICED 15) Vol 11: Human Behaviour in Design, Design Education; Milan, Italy, 27-30.07. 15, designsociety.org, pp. 051–060.
Baxter, K., Courage, C. and Caine, K. (2015), *Understanding Your Users: A Practical Guide to User Research Methods*, Morgan Kaufmann.
Bertoni, A. (2020), “DATA-DRIVEN DESIGN IN CONCEPT DEVELOPMENT: SYSTEMATIC REVIEW AND MISSED OPPORTUNITIES”, *Proceedings of the Design Society: DESIGN Conference*, Cambridge University Press, Vol. 1, pp. 101–110, https://dx.doi.org/10.1017/dsd.2020.4.
van Boeijen, A., Daalhuizen, J., Schoor, R. van der and Zijlstra, J.S. (2014), *Delft Design Guide: Design Strategies and Methods*.
Bogers, S., Frens, J., van Kollenburg, J., Deckers, E. and Hummels, C. (2016), “Connected Baby Bottle: A Design Case Study Towards a Framework for Data-Enabled Design”, presented at the Proceedings of the 2016 ACM Conference on Designing Interactive Systems, Association for Computing Machinery, pp. 301–311, https://dx.doi.org/10.1145/2901790.2901855.
Carter, S. and Mankoff, J. (2005), “When participants do the capturing: the role of media in diary studies”, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, New York, NY, USA, pp. 899–908, https://dx.doi.org/10.1145/1054972.1055098.
Gomez Ortega, A., van Kollenburg, J., Shen, Y., Murray-Rust, D., Nedić, D., Jimenez, J.C., Meijer, W., *et al.* (2022), “SIG on Data as Human-Centered Design Material”, *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, New York, NY, USA, pp. 1–4, https://dx.doi.org/10.1145/3491101.3516403.
Gorkovenko, K., Burnett, D., Murray-rust, D., Thorp, J. and Richards, D. (2019), "Supporting real-time contextual inquiry through sensor data", *Conference Proceedings. Ethnographic Praxis in Industry Conference*, Wiley, Vol. 2019 No. 1, pp. 554-581, https://dx.doi.org/10.1111/1559-8918.2019.01307
Gorkovenko, K., Jenkins, A., Vaniea, K. and Murray-Rust, D. (2023), “Data-Enhanced design: Engaging designers in exploratory sensemaking with multimodal data”, International Journal of Design, https://dx.doi.org/10.57698/V17I3.01.
Hanington, B.M. (2007), “Generative research in design education”, *International Association of Societies of Design Research2007: Emerging Trends in Design Research*.
Ishio, J. and Abe, N. (2017), “Measuring Affective Well-Being by the Combination of the Day Reconstruction Method and a Wearable Device: Case Study of an Aging and Depopulating Community in Japan”, *Augmented Human Research*, Vol. 2 No. 1, p. 2, https://dx.doi.org/10.1007/s41133-017-0006-2.
Kouprie, M. and Visser, F.S. (2009), “A framework for empathy in design: stepping into and out of the user’s life”, *Journal of Engineering Design*, Vol. 20 No. 5, pp. 437–448, https://dx.doi.org/10.1080/09544820902875033.
Li, I., Dey, A. and Forlizzi, J. (2010), “A stage-based model of personal informatics systems”, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, presented at the CHI ’10: CHI Conference on Human Factors in Computing Systems, ACM, New York, NY, USA, https://dx.doi.org/10.1145/1753326.1753409.
Mattelmäki, T., Vaajakallio, K. and Koskinen, I. (2014), “What Happened to Empathic Design?”, *Design Issues*, MIT Press, Vol. 30 No. 1, pp. 67–77, https://dx.doi.org/10.1162/DESI_a_00249.
Norman, D. (2013), *The Design of Everyday Things: Revised and Expanded Edition*, Basic Books.
Tanaka, T., Taoka, Y. and Saito, S. (2022), “Enhancing the Quality of User Research Using Embedded IoT Sensors for Collecting Life Information”, *Proceedings of the Design Society*, Cambridge University Press, Vol. 2, pp. 613–622, https://dx.doi.org/10.1017/pds.2022.63.
Visser, F.S., Stappers, P.J., van der Lugt, R. and Sanders, E.B.N. (2005), “Contextmapping: experiences from practice”, *CoDesign*, Taylor & Francis, Vol. 1 No. 2, pp. 119–149, https://dx.doi.org/10.1080/15710880500135987.
Yasuoka, M., Miyata, T., Nakatani, M. and Taoka, Y. (2023), “Untapped Potential of Participatory Design - Citizen Centered Social Data Utilization for Smart Cities”, *Distributed, Ambient and Pervasive Interactions*, Springer Nature Switzerland, pp. 156–165, https://dx.doi.org/10.1007/978-3-031-34668-2_11.
Yasuoka, M., Nakatani, M., Taoka, Y., Hamaguchi, N. and Saito, S. (2023), “A systematic literature review on utilisation of behavioural data in service design: Unexplored potentials on data utilisation on co-design”, *Service Design and Innovation Conference: Entanglements and Flows Service Encounters and Meanings*.
Yi Min Lim, D., Yap, C.E.L. and Lee, J.-J. (2021), “Datastorming: Crafting data into design materials for design students’ creative data literacy”, *Creativity and Cognition*, presented at the C&C ’21: Creativity and Cognition, ACM, New York, NY, USA, https://dx.doi.org/10.1145/3450741.3465246.