creative commons

COMMONS DEED

저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

Ph.D. Dissertation

Designing Personally Adaptable Ambient Information Visualizations for Individualized Everyday Contexts

개인화된 일상 맥락을 위한 맞춤형 주변 정보 시각화의 설계

August 2023

Dept. of Electrical Engineering & Computer Science
College of Engineering
Seoul National University

Guhyun Han

Designing Personally Adaptable Ambient Information Visualizations for Individualized Everyday Contexts

Advisor: Jinwook Seo

Submitting a Ph.D. Dissertation of
Electrical Engineering & Computer Science

May 2023

College of Engineering
Seoul National University

Guhyun Han

Confirming the Ph.D. Dissertation written by
Guhyun Han

June 2023

Chair	Youngki Lee	(Seal)
Vice Chair	Jinwook Seo	(Seal)
Examiner	Bongwon Suh	(Seal)
Examiner	Bohyoung Kim	(Seal)
Examiner	Hyunjoo Song	(Seal)

Abstract

Guhyun Han

Department of Electrical Engineering & Computer Science

College of Engineering | Seoul National University

Ambient information visualization is a category of visual representations that longitudinally convey information in the periphery of user attention in an aesthetically pleasing form. It aims to expand users' informational benefits with manageable distractions in today's information-overloaded world. Ambience, its unique characteristic shaped by information, delivery method, and representation, has been predetermined by visualization designers, resulting in visualizations not aligned with the end-user's design perspectives and attentional demands.

This dissertation presents a conceptual framework, Personally Adaptable Ambient Information Visualization (PAAIV). It offers flexible customization capabilities for ambience, allowing end-users to reflect their design perspectives on resultant visualizations. Moreover, it suggests an extended design space for attention-adaptive ambient information visualization, incorporating users' real-time attention demands using Brain–Computer Interface (BCI) technologies. I argue that PAAIV will provide end-users with visualizations of desired information, noticeability, and engagement in harmony with individualized contexts by addressing the following three research questions.

The first question involves designing flexible and usable customization capabilities for PAAIV for smartphone users. DataHalo, a customizable ambient notification visualization system, was designed and evaluated in a lab environment. With flexible customization support, it enabled smartphone users to prioritize

notifications of interest, define their longitudinal delivery, and represent them as prolonged notification badges on the home screen. A usability study demonstrated its potential to embody personalized ambience, allowing participants to rediscover everyday information they had missed due to notification overload.

The second question investigates the commonalities and differences among end-users' personalization strategies regarding PAAIV in a real-world context. An improved DataHalo was deployed in a three-week field study, revealing that the system met participants' diverse needs and improved satisfaction. Mixed-method analyses of interview and log data identified users' personalization patterns at multiple levels, suggesting the guidelines for future systems that support end-users in building and refining their PAAIV more effectively.

The third question examines the technological constraints and design opportunities of attention-adaptive features applied to PAAIV. Brain waves measured by electroencephalography (EEG) were used to infer the user's real-time attentional information. First, the technological constraints were examined under the constraints of everyday contexts where ambient information visualizations are deployed. Next, a semi-structured interview was conducted to understand the end-user's needs for utilizing real-time attentional information. The results revealed limitations, potentials, and design guidelines for further integration with PAAIV.

This dissertation contributes to the fields of Human–Computer Interaction by proposing a user-centric framework (PAAIV), which includes designing and evaluating a practical system (DataHalo), providing insights from real-world usage, and exploring the potential for attention-adaptive features. They inform design guidelines for personalized information delivery across various displays in personal contexts. The findings also suggest directions to integrate attention-adaptive features into ambient information visualization, alleviating user distraction and realizing calmer technology.

Keywords: Ambient Information Visualization; Personalization; Smartphone Notifications; In Situ Data; Brain–Computer Interface; Electroencephalogram

Student Number: 2013-20899

Contents

Abstract i

CHAPTER 1

Introduction 1

1.1 Background and Motivation 1

1.2 Research Questions and Overview of Approaches 5

1.3 Contribution 7

1.4 Structure of Dissertation 8

CHAPTER 2

Related Work 9

2.1 Personal Ambient Information Visualization and Its Challenges . . 10

2.1.1 Background and Practice in the Personal Context 10

2.1.2 Necessity of Personalization Support 11

2.1.3 Necessity of Tight Accordance with User's Attention 12

2.2 Personalization of Notification Method 13

2.2.1 Notification Delivery Mechanisms to Minimize Distraction . . . 13

2.2.2 Supporting Longitudinal Interactions with Notifications . . . 15

2.3 Integration of BCI into Attention-Adaptive Ambience **17**
2.3.1 Background of EEG and BCI 17
2.3.2 Prediction of Attention State based on EEG data 18

CHAPTER 3
DataHalo: Customizable Smartphone Notification Visualization System to Embody PAAIV **20**

3.1 Design Rationale **21**
3.2 System Design and User Interfaces **24**
3.2.1 System Components 24
3.2.2 Halo Configuration 24
3.2.3 Halo Visualization 27
3.2.4 Use case 28
3.3 Usability Study **31**
3.3.1 Study Procedure 31
3.3.2 Participants 32
3.3.3 Analysis 32
3.3.4 Result 33
3.4 Implications and Design Improvements **35**

CHAPTER 4
Understanding User's Design Strategies toward PAAIV in Daily Context **38**

4.1 Deployment study **39**
4.1.1 Study Design 39
4.1.2 Participants 40

4.1.3 Technical Specifications 41
4.2 Data Analysis **43**
4.2.1 Data Preprocessing 43
4.2.2 App-Level Customization Analysis 44
4.2.3 Virtual Category-Level Design Analysis 45
4.2.4 Halo Notification Analysis 45
4.3 Result **46**
4.3.1 Notification Management and Attendance before Using DataHalo 46
4.3.2 App-level Usage of DataHalo 47
4.3.3 Participants' Design Motivations and Choices 54
4.3.4 Virtual Category-Level Design Strategies 58
4.3.5 Halo Notification Analysis 62
4.3.6 Usability and User Experience 64
4.4 Discussion **65**
4.4.1 Customization Support towards the Personalized Ambience . 65
4.4.2 The Ease of Adaptation to Trace the Moving Sweet Spot of User Experience 67
4.4.3 Generalizability beyond Culture and Population 68
4.4.4 Implications for Better Personalization Support 69

CHAPTER 5
Exploring Design Space of Attention-Adaptive Ambient Information Visualization **71**

5.1 Expected Technological Challenges of A3InfoVis **73**
5.2 EEG Collection & Statistical Feature Analysis **76**
5.2.1 Study Design 76

5.2.2 Participants 78
5.2.3 Data Preprocessing 78
5.2.4 Statistical Analysis 79
5.2.5 Result 80
5.2.6 Analytic Insights 85
5.3 Analysis on Technological Challenges 86
5.3.1 Method 87
5.3.2 Result 90
5.3.3 Challenges Revisited: Possibilities and Limitations 94
5.4 Analysis on Design Opportunities 97
5.4.1 Participants 97
5.4.2 Study Design 98
5.4.3 Analysis 98
5.4.4 Result 99
5.5 Design Implications 102
5.5.1 Visualization of Predicted Attention State with Credibility . . . 103
5.5.2 Adequate Application Scenarios of A3InfoVis 104
5.5.3 End-user Support for Design of Personalized A3InfoVis 106

CHAPTER 6
Discussion 108

6.1 Lessons Learned 108
6.2 Design Implications for PAAIV 110
6.2.1 Personalization According to User's Design Perspectives . . . 112
6.2.2 Personalization According to User's Real-time Attentional Demands 113
6.2.3 Integration of Two Personalization Approaches 114

6.3 Future Research Opportunities for PAAIV **116**
6.3.1 Personalized Visualization Design Empowered by Image Generation Models 116
6.3.2 Extending Expressiveness of Halo Visualization 117
6.3.3 Design and Evaluation of Attention-Adaptive Ambience Visualization for Smartphone Notifications 117
6.3.4 Integration of Other Contextual Information into PAAIV 118
6.4 Limitations **118**
6.4.1 DataHalo 119
6.4.2 A3InfoVis 119

CHAPTER 7
Conclusion **121**
7.1 Summary of Approaches **121**
7.2 Summary of Contributions **123**
7.3 Summary of Future Research Opportunities **123**

APPENDIX A
Halo Visualization **138**
A.1 Connection with Constructive Visualization **138**

APPENDIX B
Usability Study of DataHalo **139**
B.1 Virtual Categories during the Task 2 . . . . . **139**

APPENDIX C
Deployment Study of DataHalo 141
C.1 Post-study questionnaire 141
C.2 Popular Design Strategies 141

Abstract (Korean) 143

Acknowledgments (Korean) 145

List of Figures

Figure 1.1 Three dimensions of Personalized Ambience. Ambience is a unique characteristic shaped by information, longitudinal delivery method, and representation of a given information display. The personalized ambience contributes to a satisfactory user experience in one's personal context, which maximizes the informational benefits and minimizes distraction costs. 2

Figure 3.1 Visual interface of DataHalo: (a) DataHalo integrated with the smartphone home screen. Each app has its own *halo* visualization that extends a conventional app notification badge; (a-b) explanation of the side-lighting visualization. A dot on the left side in (a) provides users with a visual guidance for the unsighted notifications on the left home screen pages. Navigating to the left page, users can discover another *halo* activated; (c) entrance to the *halo* design interface triggered by a long touch; and (d-g) example *halos* with their source apps. Each graphical mark corresponds to a notification. For each designated app, users can customize data, mapping, and encoding rules for distinct marks. 22

Figure 3.2 The main tabs of the halo configuration page (top) and the underlying pipeline of the visualization construction (bottom). 25

Figure 3.3 The halo visualization Han created. 28
Figure 3.4 The change of graphical mark based on visual encoding. Elapsed time and Han's attendance triggered the change. 29
Figure 3.5 Examples of app halos created by participants during the usability study. In the end of each halo design task, we captured the preview from the halo configuration page. 34
Figure 3.6 Customization flow of example-based importance modeling. 36
Figure 3.7 Customization flow of example-based visual encoding. 37

Figure 4.1 Halo creation and edit events by app category. We aggregated the events according to app category provided by Google Play Store. The color encodes the week the event occurred. 47
Figure 4.2 Halo creation and edit events by participant. The color encodes the week the event occurred. 48
Figure 4.3 Personal difference in usage of DataHalo. (a) P02 pursued pragmatism while (b) P06 valued aesthetic experience. Upon P06's request of not disclosing the photo of her favorite cat directly, we masked its face with a sunglasses. 53
Figure 4.4 An app-level design motivation (**M1**) and four virtual category-level design motivations (**M2-5**) with corresponding choices. For each design motivation, the grey-colored design choice implies that user had the minimal motivation to customize their interaction with notifications of a designated virtual category. 55
Figure 4.5 Distribution of design choices for the user-created virtual category and the Remainder category. 59
Figure 4.6 Relationship between design motivations, choices, virtual category-level design strategy, and graphical mark of P02's Loop Habit Tracker. 60

Figure 4.7 The result of halo notification analysis. 62

Figure 4.8 The result of 7-point Likert scale questionnaire about the UX of DataHalo. 66

Figure 5.1 List of commercial EEG devices available for purchase as of 2023. . . 75

Figure 5.2 Correspondence between electrodes of the 10-20 system and the brain region variable. 79

Figure 5.3 Results of the post-hoc analysis on the brain regional effect. The row corresponds to the brain regions: (i) prefrontal, (ii) frontal, (iii) central, (iv) parietal, (v) temporal, and (vi) occipital regions. The column corresponds to the relative power of frequency bands: (a) delta, (b) theta, (c) alpha, (d) beta, and (e) gamma bands. For example, a subfigure at (iii, c) represents all the statistically significant pairs that involve the central region (iii) regarding the relative alpha power (c). In the subfigure, we colored the regions using three colors: light grey, red, and blue. The average relative power of the red colored regions were significantly higher (e.g., occipital $>$ central), and that of the blue colored ones were significantly lower than that of the light grey region (e.g., frontal $<$ central). 81

Figure 5.4 Pipeline of the real-time classification of the user's attention state. The classifier receives the brain wave feature matrix (19 channels * 24 features) every one second and predicts the corresponding attention state. For example, in this figure, the predicted attention state of the user at the current time t is the 'video-watching'. For more information on feature matrices, see the Feature Selection subsection of the text. 87

Figure 5.5 Average performance of personal models trained with channels of the selected regions. For classifiers, XGBoost (solid blue) always outperformed C-SVC (light blue). Among the region-specific models, the prefrontal, frontal, and temporal region models consistently performed better than others. 91

Figure 6.1 PAAIV Framework and its two personalization approaches. 111

Figure A.1 The components of constructive visualization and corresponding design constraints of *halo* visualization. **Basic Token** refers to a discrete graphical mark representing a data unit. **Token Grammar** refers to a mapping of a graphical mark's visual channels to data properties. **Environment** refers to a spatial constraint on the assembling methods of tokens based on Token Grammar. **Assembly Model** refers to a construction model of the visual representation. 138

List of Tables

Table 5.1 Two-way RM ANOVA Results of Relative Power of Frequency Band . . 82

Table 5.2 Two-way RM ANOVA Result of Task Engagement Index 84

Table 5.3 Average Performance Metrics of Personal Models by Selected Features 93

Table 5.4 Comparison of Performance Metric between Personal Models and General Model. 94

Table B.1 Smartphone applications that the usability study participants chose for the second halo design task and virtual categories they created. . 140

Table C.1 Questions for inquiring the user experience during the DataHalo usage. 141

Table C.2 Frequent and widely used design strategies among participants. *#P* stands for the number of participants who shared the strategy and *Freq.* stands for the total frequency of the strategy. 142

Chapter 1

Introduction

1.1 Background and Motivation

Taking a balance between ever-expanding information and our daily lives has been an increasingly important issue in the Human–Computer Interaction (HCI) community [1, 42, 86, 104]. These days, people are surrounded by multiple smart devices which connect them with a variety of information technology (IT) services. Each service aims to constantly produce 'useful' information and deliver it 'timely' to individual users, mainly via notification method. However, this information collectively exceeds the users' cognitive capacity and its delivery often disrupts the users' primary task.

Initiated by the notion of *calm technology* [104] suggested by Weiser and Brown, HCI researchers have been interested in these problems. In the same context, Google, which distributes one of the most influential operating systems of smart devices, also suggested the concept of *digital well-being* [32] and has launched commercial and experimental projects to handle the limitations. Smartphone notifications, often cited as the cause of user distraction in exchange for delivering information relevant to the personal context, are one of the technologies intensively handled in their projects.

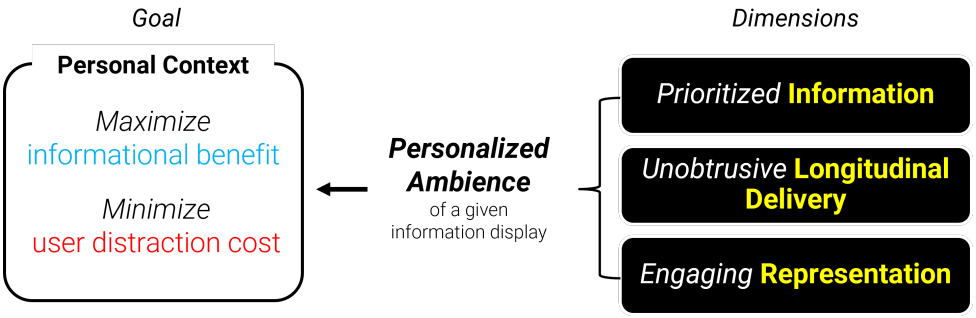


Figure 1.1: Three dimensions of Personalized Ambience. Ambience is a unique characteristic shaped by information, longitudinal delivery method, and representation of a given information display. The personalized ambience contributes to a satisfactory user experience in one's personal context, which maximizes the informational benefits and minimizes distraction costs.

Ambient information visualization [65, 86] is a concept of visualization methodology that tackles the balancing problem between informational benefit and user distraction mentioned above. It visualizes the information of users' interests in an aesthetically pleasing form, enables users to longitudinally access the information in the periphery of their daily context through various personal displays, and persistently provides the change in the information in a less disruptive manner. By doing so, it aims to improve both the amount and quality of information people can reach and provide a satisfactory experience of longitudinal interaction with manageable distractions.

To fulfill the goal, the key challenge is to personalize *ambience* for individual users (See Figure 1.1): It should (1) provide the prioritized information that fits the user's interest and information capacity, (2) convey the information for the opportune time span when users feel it useful without disrupting them, and (3) represent the information using the preferred visual encoding that persistently motivates users to engage with it.

However, previous ambient information visualizations have provided users with limited capability to customize the information, delivery, and repre-

sentation as they wish. They usually had to choose a ready-made visualization as-is, of which information, delivery, and representation are determined by designers. Thus, the ambience provided by the pre-carved visualization could be sub-optimal for users from the beginning or become less satisfactory as their usage context changes in the long term. While previous works [17, 18] have highlighted the importance of providing customization capabilities for users, only a few works [14, 70] have addressed the issue, indicating the limited knowledge about *what kinds of capabilities end-users need* and *how end-users utilize these capabilities to reach personalized ambience.*

Moreover, previous ambient information visualizations have been impossible to finely accord with the user's attentional demands in real-time. Designers of ambient information visualization take heed of the user's attention capability. They have sought a suitable amount of the delivered visual information that fits the user's attention capacity and an acceptable level of notification strength that does not disrupt the user's daily task. However, they had to depend on indirect and post-hoc methods that probed the attention state, such as retrospective user feedback [63] or observation of user behaviors [6, 17, 19, 97]. These methods could hardly adapt to real-time changes. In addition, the subjective responses through self-reflection or observation results might not be consistent with the internal attention state.

This dissertation presents a conceptual framework called *Personally Adaptable Ambient Information Visualization* (*PAAIV*) to overcome the limitations mentioned above. PAAIV aims to provide end-users with personalized ambient information visualizations aligned with end-users' (1) design perspectives and (2) real-time attentional demands. First, it enables end-users to personalize the ambience by supporting a flexible and usable customization interface. This approach can support end-users to build a more satisfactory

ambient visualization than pre-carved ones and improve it according to the temporal change in their preference. Next, it suggests an extended design space in which end-users can alleviate distraction further according to their real-time attentional demands. For example, they might update information faster or add make the visualization more noticeable (e.g., graphical marks get bigger, and the animation starts) only when they feel less disruptive.

The framework was developed based on two works that studied the personalization approaches, respectively. The first work presents a design study in which a customizable ambient visualization system was suggested for smartphone notifications and was deployed to the real usage context of smartphone users. The second work covers an empirical study on electroencephalogram (EEG) data in which the technological feasibility of attention-adaptive features, followed by an interview with end-users experienced in designing ambient information visualization, which explored possible application scenarios.

Thanks to the implications of these two works, PAAIV is expected to provide design guidelines for customizable ambient information visualization systems, which could be generalized to various personal smart devices incorporating visual displays and notification methods, such as personal computers, smartwatches, and extended reality (XR) devices. It can also suggest the future directions of integrating attention-adaptive features into the design of ambient information visualization, which would alleviate user distraction further and realize the *calmer* technology.

1.2 Research Questions and Overview of Approaches

Over the course of the dissertation, I have sought ways to address the following research questions.

- Q1. *How should we design flexible customization capabilities for PAAIV, allowing people to enhance informational benefits while effectively managing distractions across diverse types of personal information?*
- Q2. *What would be commonalities and diversities in people's customization strategies for embodying their personalized ambience by utilizing the PAAIV system in everyday contexts?*
- Q3. *What are the limitations and opportunities for integrating attention-adaptive personalization into PAAIV, considering technical constraints and user needs related to real-time attentional information?*

To address Q1, I designed and developed DataHalo, a customizable notification visualization system that represents smartphone notifications as prolonged ambient visualizations on the smartphone home screen. It is implemented as an Android home screen launcher embedded with fine-grained personalization capabilities of (1) prioritizing app notifications into multiple virtual categories, (2) modeling longitudinal interaction with them, and (3) representing them with extended notification badges specified by flexible visual encoding. The feasibility of DataHalo was assessed through a usability lab study. The result showed that the customization capabilities provided by DataHalo were feasible to embody PAAIV: participants could manage notifications better and rediscover the value of diverse information delivered by smartphone notifications. On the other hand, customization fatigue was

identified as a potential obstacle, leading to improvements based on user feedback.

To address Q2, I deployed the improved version of DataHalo into the real-world context to explore people's personalization strategies for better management of notification information. A three-week field deployment study collected usage logs of DataHalo and partial notification logs. The results indicated that DataHalo helped actualize the participants' diverse needs for better modeling, expressing, and consuming notification information. Mixed-method analyses involving the participants' customization events, partial notification data, and interview data identified the personalization patterns from three perspectives: (1) customization over time, (2) their design strategies to embody the personalized ambience, and (3) notification logs affected by halos. On these bases, the suggestions for the future smartphone notification system were summarized, which could support laypeople to build and refine their PAAIV more easily and expressively.

To address Q3, I investigated both the technological constraints and design opportunities of the Attention-Adaptive Ambient Information Visualization (A3InfoVis), which aims to integrate the Brain–Computer Interface (BCI) method into the design of the PAAIV. First, three technological challenges (i.e., the feasibility of attention state prediction for everyday tasks, the limited spatial resolution of EEG data available for end-user devices, and the burden of data acquisition for building user-specific classifiers) were examined through an empirical study. The results revealed the possibilities and current technological limitations. Next, a semi-structured interview for DataHalo users, experienced in customizing ambient information visualizations, was conducted to understand end-users' needs for utilizing their real-time attentional information to enhance their informational benefits and mini-

mize distraction. On these bases, the design guidelines for A3InfoVis were summarized, which further extends the design space of PAAIV.

1.3 Contribution

The core contributions of this dissertation are as follows:

1. Design and implementation of the DataHalo, a flexible notification visualization system that helps people to realize their PAAIV for the information delivered by smartphone notifications.

2. Empirical findings, acquired through a field deployment study of DataHalo, that clarify the design motivations and relevant choices posed by smartphone users to achieve their personalized ambience through managing notification information.

3. Empirical findings, acquired through an empirical study and a user interview, that provide insights on the technological constraints and design opportunities of integrating attention-adaptive personalization into PAAIV.

Thesis Statement Personally Adaptable Ambient Information Visualization (PAAIV) is a conceptual framework that enables end-users to personalize the longitudinal interaction with visual information in everyday contexts according to their (1) design perspectives and (2) real-time attentional demands, eventually providing them with the visualization of desired information, persistence, and appearance in harmony with attentional demands of the personal context.

1.4 Structure of Dissertation

The rest of this dissertation is organized as follows. Chapter 2 presents prior works that serve as a theoretical basis for PAAIV from three perspectives: (1) ambient information visualization, (2) smartphone notification, and (3) attention state prediction leveraging brain waves.

Chapter 3 illustrates the design of *DataHalo*, a customizable notification visualization system that supports smartphone users to build PAAIV for diverse app notifications. The chapter also covers its evaluation in the lab environment ($N = 17$) to prepare for field deployment.

Chapter 4 reports on the qualitative and quantitative findings regarding how smartphone users construct PAAIV in their smartphone home screen. The chapter describes a three-week field deployment study ($N = 12$) with DataHalo and the results of three mixed-method analyses.

Chapter 5 presents the limitations and potentials of *A3InfoVis*, Attention-Adaptive Ambient Information Visualization, which integrates attention-adaptive features into PAAIV. The chapter covers (1) an empirical study including EEG data collection ($N = 27$) followed by performance analysis of ML-based attention state classifiers and (2) a semi-structured interview ($N = 7$) with end-users familiar with customization of ambient information visualizations.

Chapter 6 discusses design challenges raised by the above studies and further implications for improving PAAIV. The limitations of the studies are also acknowledged.

Finally, Chapter 7 concludes the dissertation by summarizing the findings in this dissertation and the opportunities for future research.

Chapter 2

Related Work

This chapter covers the related work of the dissertation. Section 2.1 covers the background of ambient information visualization and its practice, focusing on usage in the personal context. Barriers and challenges for these areas are also indicated from two perspectives: (1) customization support and (2) tight accordance with the user's attention. Section 2.2 reviews prior works on (smartphone) notification methods, which provide the theoretical background of DataHalo. It highlights the prior efforts on (1) improving notification delivery mechanisms to minimize distraction and (2) supporting longitudinal interactions with notifications. Section 2.3 presents the background of BCI technologies applicable to A3InfoVis. It covers (1) the relationship between brain waves and the user's attention and (2) the prediction of attention information by leveraging EEG data.

2.1 Personal Ambient Information Visualization and Its Challenges

2.1.1 Background and Practice in the Personal Context

Ambient information visualization [85] longitudinally conveys useful but non-critical information with an abstract representation in the casual and personal context. Following the notion of calm computing, which utilizes the periphery of user-attention to convey information [104], researchers have incorporated *ambient* displays to visualize personal data in a longitudinal and less disruptive manner [64, 65, 85].

Unlike traditional information visualizations such as charts, ambient visualizations tend to involve abstract and artistic expressions to deliver the subtle change of data in an aesthetically pleasing manner [64, 85]. Typical information visualization usually assumes that people dedicate their full attention to its visual analytic task during a focused session. On the other hand, ambient information visualization targets audiences who might be less interested in analytic insight but need to reach useful information in diverse contexts of everyday life. To involve these daily contexts, ambient information visualizations have covered different information sources (e.g., bus arrivals [96], physical activities [19]), devices (e.g., public displays [96], smartphone [8, 19], smart wearables [34], dedicated artifacts [14]), and visual representations (e.g., geometric shapes [95, 96], flowers [59], collage [28, 70]). For example, Kandinsky system [28] was inspired by the abstract art of Kandinsky to visualize the text data like email. Skog and colleagues' work [96] was inspired by the abstract art of Mondrian to visualize the bus arrival information. Ubifit Garden visualized various aspects of physical activity using flowers and insects in a virtual garden on the phone screen [17, 18, 19].

Similarly, Info-Lotus visualized email notifications as decorative flowers on a peripheral region of the desktop screen [108].

Moreover, they are often deployed in personal context [42] that involves personal devices (e.g., PC [108], smartphone [8, 19]) or private spaces (e.g., bedroom [53]). Such ambient visualization systems maintain information on a glanceable spot such as screen corners on the desktop (e.g., screen time [52]), phone screen background (e.g., sleep-related activities [8], physical activity [17, 18, 19]), smart wearable (e.g., physical activity [34]), physical environment (e.g., auditory activity data [95]), or dedicated devices other than ones for main tasks (e.g., a clock-like object visualizing time schedules [57, 58], an LCD display in the bedroom [53], a mobile system visualizing the communication via multiple smartphones [59]).

2.1.2 Necessity of Personalization Support

Because individual users might have different perspectives on useful information, attentional demand, and aesthetics, some previous works [14, 70] provided users with customization capabilities to personalize ambient visualization to achieve satisfactory user experience. However, end-users often had to use ready-made ambient visualizations without the ability to adjust them to individual contexts. The long-term deployment study by Consolvo and her colleagues pointed out the potential limitation due to the lack of customizability [17]. Participants suggested supporting their preferred visual themes in addition to the pre-defined theme to prevent boredom and keep them motivated to use the ambient visualization.

Some systems have been designed to provide users with a certain degree of freedom to create custom ambient visualizations. In the InfoCanvas system [70], users could generate a custom ambient visualization from a pre-

defined set of data and visual representations. Although the data and visual resources could be further expanded by editing XML codes, this seems out of reach to lay users. Calm Automaton [14] introduced a programmable, physical ambient display. Users could define a simple visual mapping between the data values and the motion of artifacts and change the visual metaphor by replacing images. DataSelfie [51] provided a more advanced interface for generating custom, personal visualizations. Users could define the visual mapping between random categorical data variables and user-created visual expressions by manipulating a questionnaire-style interface.

However, these researches are limited in further exploring the appropriate level of customization capabilities that should be offered to users. Doing so can contribute to broadening the understanding of users' design perspectives, which has been pointed out by Huang et al. [42] as an important design challenge. This dissertation intended to explore the question by investigating how users make use of the flexible customization interfaces for each design step of ambient visualization from processing a raw notification to specifying a visual presentation.

2.1.3 Necessity of Tight Accordance with User's Attention

User's attentional demand has been a crucial factor in the design of ambient information visualization. It is anchored in the notion of calm technology [104], which aimed to help users to access more valuable information about their surroundings in a less disruptive manner.

Emphasis on the virtue of manageable distraction has influenced the formulation of design dimensions and evaluation criteria that assess if a visualization delivers an affordable amount of information with acceptable notification strength [63, 65, 85, 93]. Pousman and colleagues [85] investigated 19

ambient information visualization systems according to their design dimensions, which included **information capacity**, the number of discrete information sources the visualization delivers, and **notification level**, the degree to which system interrupts users. Bakker [6] modeled and explained the attentional transition between the center and periphery based on the observation of user behaviors. On these bases, designers evaluated their ambient information visualization by observing the usage in the lab or field and by collecting the retrospective feedback after the usage [6, 17, 19, 97].

However, the previous design and evaluation approaches have been limited to indirect methods in terms of involving the user's internal attention state. For example, the inquiry methods depend on the user's reflection so they hardly capture the real-time transition of attentional states. The observation could characterize the distinct user behaviors during the usage but is ineffective in figuring out what the exact internal state is. Therefore, this dissertation tried to examine if brain waves, real-time physiological data, would be a useful source of information to improve the user experience of ambient visualization.

2.2 Personalization of Notification Method

2.2.1 Notification Delivery Mechanisms to Minimize Distraction

The HCI community put intensive effort into collecting and understanding smartphone notifications (*c.f.*, NotificationLog [102]). A rich body of research has pointed out the negative aspects of notifications, particularly interrupting people's daily activities and causing stress [50, 56, 72, 80, 81, 94]. For example, Kushlev and colleagues found that interruptions caused by smartphone notifications could trigger higher levels of inattention and

hyperactivity of people, which harms their mental well-being [54]. The HCI and UbiComp communities have explored ways to alleviate interruption and distraction primarily by delivering (1) only relevant notifications and (2) at opportune moments. For example, there exists a line of research to understand people's mental modal on the relevance of notifications [5, 61, 91, 103]. Shirazi and colleagues found that smartphone users assess the importance of notifications by the category of the source app, the topic of the content, and the sender that triggered a notification (e.g., system, self, and family members) [91]. As a result of their in-situ study, Weber and colleagues [103] pointed out that a substantial portion of daily notifications was not important (38.91%) and not urgent (51.75%) for smartphone users. Users' assessment of notifications was discrepant with the system-side priority, carved by app providers. Lin and colleagues also utilized people's rating of notifications in four dimensions—importance, urgency, sender attractiveness, and content attractiveness—to analyze the seven types of notification attendance they suggested [61]. The highly-rated notifications tended to be handled faster than others regardless of the system-provided display order of notifications. These works call for a fine-grained and personalized approach for determining the relevance of individual notifications, even from a single app. However, major smartphone operating systems rarely support filtering out unnecessary notifications based on their content because users can suppress notifications as a whole, by an app, or by a channel that an app developer predefined.

Another line of research explored ways to deliver notifications at opportune moments to be interrupted. "Do Not Disturb" is the one primitive and straightforward method to specify when is *not* the opportune moment. Some systems incorporated automated prediction mechanisms for opportune mo-

ments. For example, Social Context-Aware Smartphone Notification System (SCAN) delays notifications depending on the person's social context (e.g., presence of others) [76]. Mehrotra and colleagues proposed a data-driven mechanism for classifying notifications and delivering them at opportune moments based on the content of the notification and other contextual information such as location and surrounding sound [68]. The authors further extended the mechanism into a semi-automated notification control system called PrefMiner [67], which presents the data-driven rules in a human-readable form and allows people to manipulate them. Pielot and colleagues explored a machine learning approach to predict the opportune moment for engaging with a notification with hundreds of features covering the data about the users and their contexts [79]. Cadvar and colleagues analyzed and modeled the association between mobile app usage and the users' workload in an office setting [11] to study their responsiveness to a break-reminder notification. However, one common limitation of these systems is that the content on the deferred notifications is likely outdated because it is not updated since the notification was originally generated.

2.2.2 Supporting Longitudinal Interactions with Notifications

Smartphone notifications are often designed to be instant and ephemeral: Once a person taps on a notification to run the source app, it is discarded from the notification drawer/list. On the other hand, research has shown that people often interact with notifications in an opposed way. For example, Pielot and colleagues analyzed people's attendance to notifications by capturing the transition over the interaction stages (shown, seen, checked, consumed, and removed) over time [82]. They reported that not all seen or checked notifications are immediately consumed (i.e., tapping) or removed,

and participants showed different patterns depending on the source app. Some people deliberately keep specific notifications on a drawer to leverage them (or badges on an app icon) as visual reminders [61]. However, people often lose notifications by unintentionally tapping on the list or the 'dismiss all' button. In this vein, numerous mobile apps that enable people to review past notifications were released in practice, such as "Notification History" [73], which was downloaded over five million times from Google Play Store. Weber and colleagues found that managing the history of dismissed notifications could relieve the user's anxiety of losing important notifications due to careless actions [101].

People also *snooze* a notification when they want to be notified again later. However, the snoozing feature of both Android and iOS allows people to set only a fixed time interval; people often repetitively snooze a notification for a short term until they are ready to consume it. Weber and colleagues explored how people snooze notifications when provided with extended operations (e.g., snooze for time duration or snooze until specific time points) [101] and reported that it was necessary to provide personalized options that fit individuals' contexts and behaviors.

Previous works revealed the smartphone users' desire beyond the one-off interaction and the desire could be diverse based on individual users' smartphone environments (e.g., installed apps and their usages). However, to the best of our knowledge, few studies further explored what the specific scenarios are and how these scenarios could differ among people. Hence, This dissertation aimed to inquire about the day-to-day scenarios of people and understand what kinds of support they need. In addition, a novel method to personalize the duration for persisting informative notifications was contemplated to overcome these limitations.

2.3 Integration of BCI into Attention-Adaptive Ambience

2.3.1 Background of EEG and BCI

Brain waves refer to *rhythmic fluctuations of voltage between parts of the brain resulting in the flow of an electric current* [25]. Electroencephalography (EEG) is a non-invasive method to record brain wave activities by calculating the temporal and spatial sums of the post-synaptic field potential measured at the scalp. Electroencephalogram refers to the data recorded by the method. The electrodes of EEG devices cover different brain lobes of distinct functions so that the EEG channels illustrate different spatial patterns. Researchers apply frequency domain analysis methods to understand the complex patterns underlying EEG data, such as calculating the power spectral density (PSD). There are five popular frequency bands for analysis: delta (1-4Hz), theta (4-8Hz), alpha (8-12Hz), beta (12-30Hz), and gamma (over 30Hz). Previous works have identified the connection between the features of these bands and the subject's attention state from diverse perspectives. Not only the individual bands (e.g., [38] for the delta, [16] for the theta, [29] for the alpha, [26] for the beta, and [40] for the gamma) but also the interaction of multiple bands [10, 78, 100] could explain the different status of cognition and memory.

Researchers have been interested in predicting the user's engagement with the target task using one's EEG data. A heuristic task engagement index ($\beta/(\alpha+\theta)$) [84] suggested by Pope and colleagues has been widely used. It involved the relative power of the theta, alpha, and beta bands as parameters. Researchers have interpreted subjects' attention states during the task of interest (e.g., playing games [66]) based on the measure. HCI researchers also have utilized it to prototype BCI systems adaptive to user's attentional

change (e.g., a storytelling robot that reacts to the drop of the audience's engagement [98], a visualization for the lecturer that monitored the students' engagement during the class [39]). The engagement measure itself might not reveal the target task where the user is currently focused. The recent work [47] presented the possibility of separating the task-related and task-unrelated thoughts by inspecting the P300 (P3) component of event-related potentials (ERP), the brain response of positive amplitude measured as the direct result of stimuli within a delay of 250-500ms [83], and alpha power variability.

2.3.2 Prediction of Attention State based on EEG data

Researchers also have addressed personalized ML approaches to classify distinct attention states. Nuamah and colleagues computed the Pope's index values for every EEG channel and trained ML models with them [74, 75]. They could effectively distinguish five cognitive tasks (i.e., baseline, multiplication, letter composition, geometric figure rotation, and visual counting). Berka and colleagues [9] derived an individualized, four-class model of task engagement (i.e., high, low, relaxed wakefulness, and sleep onset) based on quadratic DFA (Discriminant Function Analysis). They then utilized the prediction probability of the 'high' class to be the task engagement index. They examined the index while the subjects performed various cognition and memory tasks (e.g., forward/backward-digit-span, mental-addition). The result indicated that their task engagement measure could be feasible to capture the allocation of attention within the subject. Lelis-Torres et al. [60] adopted a similar approach to analyze the user's motor skills.

Despite the successful results, there exists a gap between previous works and the goal of applying BCI methods to ambient information visualization.

First, they explored the attention state with very simple and specific cognitive tasks (e.g., mental-addition) in a highly controlled environment (i.e., laboratories). However, everyday tasks that take place in users' casual context, such as watching Youtube videos, consist of complex audiovisual stimuli and are hard to quantify the performance. Second, most of them utilized professional EEG devices with a large number of channels (typically over 64 channels) that densely cover the brain regions. These devices are expensive and hardly usable; users cannot wear these devices for themselves without aid. On the other hand, end-user EEG devices cover the limited brain regions with a few electrodes so the data quality might not be sufficient to reproduce previous results. Finally, individual users should collect EEG signals by themselves to train personalized ML-based classifiers that distinguish the attention states of their interest. The user burden of data acquisition needs to be clarified to support users to utilize personalized attention information.

Chapter 3

DataHalo: Customizable Smartphone Notification Visualization System to Embody PAAIV

This chapter[1] presents DataHalo, a customizable notification visualization system that represents notifications as prolonged ambient visualizations on the smartphone home screen. It covers the first research question: *How should we design flexible customization capabilities for PAAIV, allowing people to enhance informational benefits while effectively managing distractions across diverse types of personal information?* Smartphone notifications cover diverse kinds of timely information generated in the personal context, including notices from app services, communications with others, and self-organized events. On the other hand, their overflow burdens users with two challenges: (1) prioritizing the

[1]The preliminary version of Chapter 3 and Chapter 4 were published as a paper [37] in Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. In these chapters, "we" denote me and my collaborators: Jaehun Jung, Young-Ho Kim, and Jinwook Seo.

informative notifications as they wish and (2) retaining the delivered information as long as they want to utilize it. To provide flexible customizability, DataHalo supports keyword-based filtering and categorization and displays favorable graphical marks based on the time-varying importance model to enable longitudinal interaction with the notifications. The feasibility of DataHalo was assessed via a usability lab study, which revealed that the customization features offered by the system could enhance the management and rediscovery of various information communicated through smartphone notifications. Based on usability feedback, DataHalo was improved further to alleviate the possible customization fatigue.

3.1 Design Rationale

DR1: Support personalized categorization of notifications. Prior literature suggests that people have individualized assessment on the relevance of notifications [5, 61, 91], which is not well supported by existing smartphone operating systems. As such, we aimed to allow people to categorize notifications based on the intrinsic information. In addition, considering the diversity of mobile apps [21], the categorization method should be applicable to any apps. As a result, we enabled people to define categorization rules by an app, based on keywords in the textual information of the notification (e.g., sender and content).

DR2: Support longitudinal interactions with notifications. Although the existing notification drawers (usually) dismiss a notification upon attendance, prior literature suggests that attendance proceeds longitudinally and people often want to retain notifications even after attendance [61, 82, 101]. As such,

Figure 3.1: Visual interface of DataHalo: (a) DataHalo integrated with the smartphone home screen. Each app has its own *halo* visualization that extends a conventional app notification badge; (a-b) explanation of the side-lighting visualization. A dot on the left side in (a) provides users with a visual guidance for the unsighted notifications on the left home screen pages. Navigating to the left page, users can discover another *halo* activated; (c) entrance to the *halo* design interface triggered by a long touch; and (d-g) example *halos* with their source apps. Each graphical mark corresponds to a notification. For each designated app, users can customize data, mapping, and encoding rules for distinct marks.

we allowed people to customize how long the notifications should be accessible and how the relevance of the information should change by elapsed time and user interaction, which we call the 'importance model' as a whole. Because the Android system did not allow much flexibility on the design of notification drawers for developers, we incorporated dedicated 'Notification History' views inside a companion app where people can access past notifications regardless of the status of the notification drawer.

To effectively display notifications according to the importance model, we incorporated ambient visualization on the home screen by representing each notification as graphical marks (See Figure 3.1). With marks, we can visualize the varying importance of a notification through visual properties such as color and size. We chose the home screen as an ambient display for two main reasons: (1) We can display the graphical marks for notifications in close proximity to icon of the source app; and (2) app icons tend to be arranged according to the frequency of the attendance so the notification visualizations for frequently used apps are likely to be exposed more pre-attentively.

DR3: Facilitate lay individuals to customize the visual encoding. Our target audience is the general public who uses smartphones and might not be familiar with information visualization. Therefore, we followed the bottom-up, or constructive design paradigm [43] (Figure A.1), in which (1) an individual data item is assigned to a unit graphical mark without aggregation, (2) the user-defined visual grammar specifies the presentation of marks, and (3) the marks are assembled based on user-defined rules to construct the final visual representation. This bottom-up approach is known to be more accessible to lay individuals [69]. We mapped each notification element with a mark,

assuming that the number of marks around each source app would not be many considering the average number of daily notifications [91]. In addition, we enabled people to choose the marks that can best depict the paired notification category. For example, one can assign a heart icon for notifications regarding messages from their significant other. By allowing personalized visual encoding, we expected people to easily catch the comprehensive trend of the current notifications at a glimpse.

3.2 System Design and User Interfaces

3.2.1 System Components

DataHalo was built upon Rootless Pixel Launcher [107], an open-source launcher on Android that allows us to customize the home screen. The system consists of two main components: (1) A halo configuration page that can be accessed via a context menu (See Figure 3.1c) of the installed app icons either on the home screen or the app drawer and (2) halo visualizations on the home screen. In its back-end, DataHalo includes (1) a new data structure of enhanced notifications and its periodic update logic, (2) an enhanced notification manager based on Android *NotificationListenerService*, and (3) a manager of users' *halo* design specifications.

3.2.2 Halo Configuration

Figure 3.2 describes the pipeline for halo configuration of an app (in this case, Kakaotalk, a messenger app that is widely used in South Korea). Here, people can configure halo visualizations for the app by defining categorization rules, importance model, and visual encoding.

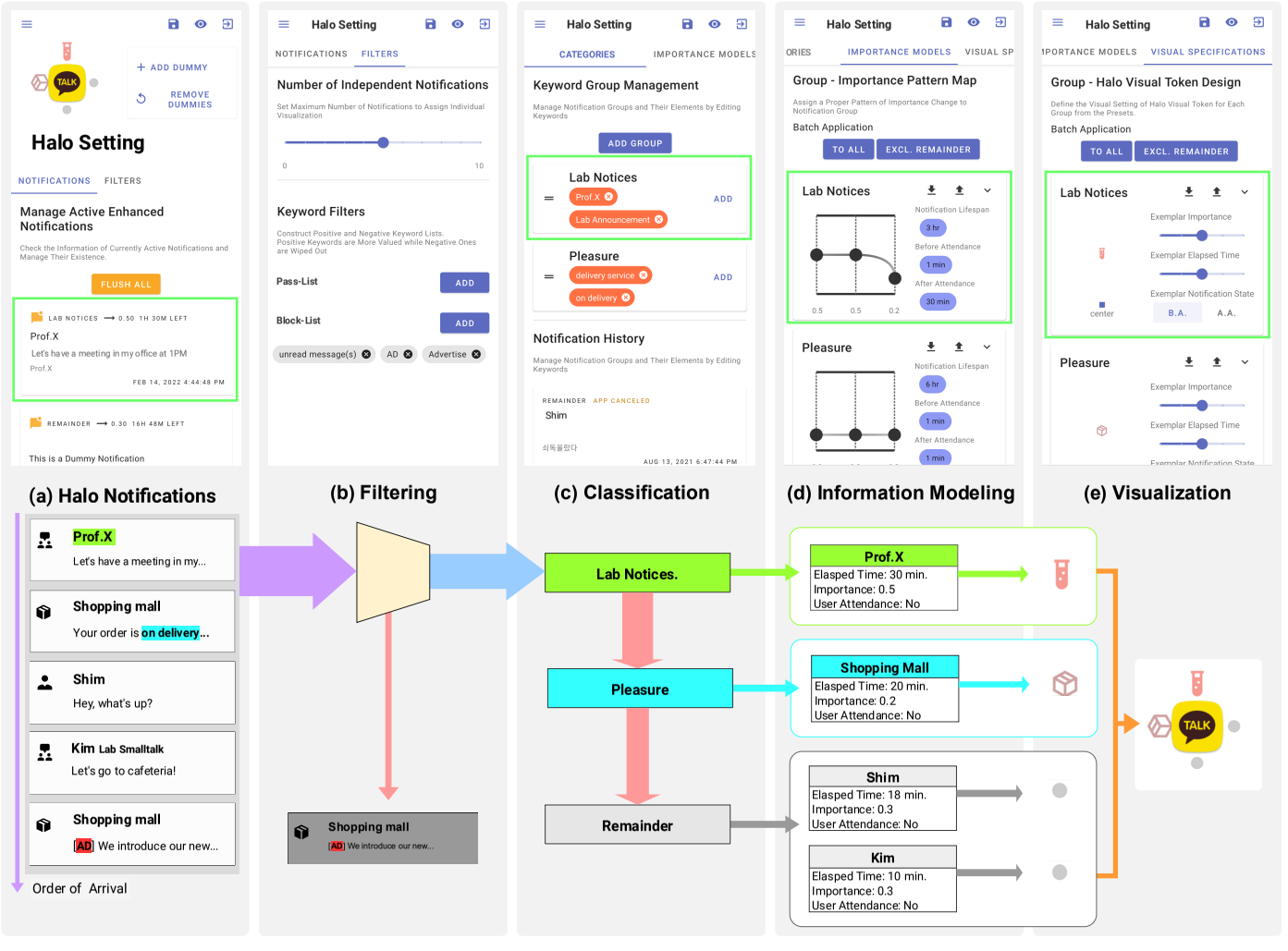


Figure 3.2: The main tabs of the halo configuration page (top) and the underlying pipeline of the visualization construction (bottom).

The **Notifications** tab (Figure 3.2a, top) lists the notifications that are currently being manged by DataHalo, regardless of their status on the notification drawer, in the order of their current importance.

On the **Filters** tab (Figure 3.2b, top), people can define a global filter for notifications that they want to see as halos (DR1). For example, people may filter out notifications for advertisement messages, which are likely to include an "[AD]" prefix in their text, by adding "AD" to the 'Block-List' (See Figure 3.2b); the notifications filtered out do not affect the visualization pipeline.

On the **Notification Categories** tab (Figure 3.2c, top), people can define *virtual* categories with keyword-based rules applied to the textual information of notifications (e.g., sender, title, and content; DR1). According to these categorization rules, DataHalo automatically categorizes the incoming notifications of the app. The notifications unhandled by the categories people created belong to a special virtual category, "Remainder". For each virtual category, people can assign independent importance model and visual encoding.

On the **Importance Models** tab (Figure 3.2d, top), people can define how the importance of the notification should change over time, before and after people attend it (i.e., tap on the notification; DR2). First, people can set the **initial importance** between 0.0 and 1.0 and the **target importance** to pursue until attendance. They can set the duration for the pursuit before attendance to the notification ("Before attendance" in Figure 3.2d, top) and if the notification is not attended after the importance reaches the target, the importance value remains the same. Second, people can set the **saturation importance** and the duration for the pursuit after attendance to the notification ("After attendance" in Figure 3.2d, top). Once people attend the notification, its

importance moves to the saturation value for the duration. Since the initial delivery, a notification persists in the (DataHalo) system for the **lifespan** duration people set. By default, the importance model behaves identical to the ordinary notification; the importance value does not change until the attendance (**initial** importance = **target** importance = 0.3); it disappears just after the attendance (**target** importance = 0.0). For novice users, we included a set of preset importance models, based on the implications from the usability study (Section 3.4).

3.2.3 Halo Visualization

Each notification that passed the global filter (See Figure 3.2b) is represented as a graphical mark around the source app icon on the home screen (DR2, DR3). People can configure how marks will look for each virtual category and encode importance/time variables or constants to visual properties of the mark. The system sorts the notifications by importance and draws them in clockwise order, starting from the top. If people do not customize the visual encoding of the virtual category, its default graphical mark is a static gray circle (See the mark of the Remainder category in Figure 3.2d and e).

On the **Visual Specifications** tab (Figure 3.2e, top), people can configure the visual encoding of each virtual category (DR3). DataHalo supports modification of five visual properties: (1) shape, (2) color, (3) animation, (4) size, and (5) distance from the center. As a shape of the graphical mark, people can choose one from a set of icons, upload a custom image, or use a text for the name of the category. For animation, people can combine four types of looped animations (blink, scale, rotate, jitter) to make the mark more attractive. Beside setting a constant value for the properties, people can encode importance, a ratio of elapsed time since delivery to lifespan, and a binary

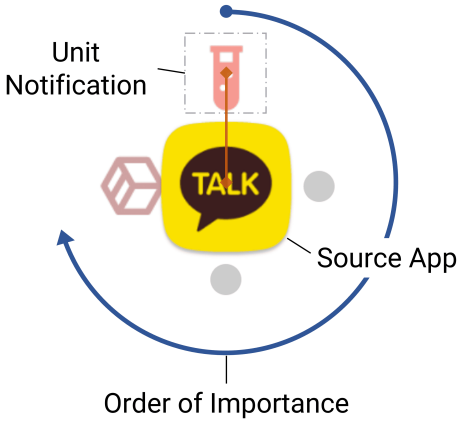


Figure 3.3: The halo visualization Han created.

status of attendance to any of color, animation, and size. For example, one can configure a graphical mark for depicting a visual metaphor of sunrise; its color changes from red to yellow over time; its distance from the center becomes greater as importance increases. For novice users, we included a set of preset graphical marks with the visual encoding configured, based on the implications from the usability study (Section 3.4).

The discoverability of the halo visualization could be sub-optimal as users cannot easily glance at the visualization if it is not located on the first page of the home screen. To overcome this problem, we introduced a side-lighting visualization (Figure 3.1a-b). Each graphical mark provides a navigational cue that one or more activated halos exist if the user moves along the designated direction. Its brightness is proportional to the sum of the existing informative values of the activated halos along the direction.

3.2.4 Use case

Here we describe DataHalo's interactions with a usage scenario of Han, a graduate student who wants to better manage text messages, from a messen-

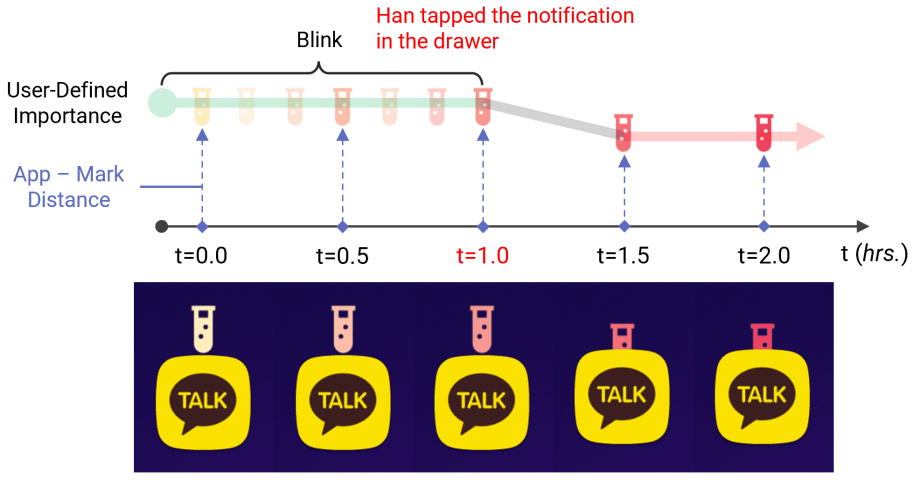


Figure 3.4: The change of graphical mark based on visual encoding. Elapsed time and Han's attendance triggered the change.

ger app called KakaoTalk, regarding important notices from his laboratory and his advisor. Han receives about a hundred of notifications per day from KakaoTalk, about incoming messages from friends and group chat rooms. Not to be distracted, Han has sustained the smartphone notification feedback silent for most of the day and checked the smartphone for about every 2-3 hours. Skimming a pile of notifications, Han often dismissed them in bulk by tapping "dismiss all" in the notification drawer. Meanwhile, Han occasionally dismissed and forgot important or urgent messages about his research and school. Hence, Han runs DataHalo to manage the KakaoTalk notifications by emphasizing ones related to his school. In particular, he wants to easily distinguish the lab notices from other KakaoTalk messages, and maintain them for several hours as a visual reminder even if he dismisses the notifications on the drawer by mistake.

Defining Global Filter and Virtual Category

Han long-presses the KakaoTalk icon on the home screen and opens the Halo Configuration page via a context menu. He first configures a global filter to exclude unnecessary advertisement spams (Figure 3.2b). To separate notification regarding his school, Han decides to capture the direct messages from his advisor and ones from the group chat room with his colleagues. On the **Notification Categories** tab, he adds a new virtual category named "Lab Notices" and appends two keywords, "Prof.X", his advisor's name, and "Lab Announcement", the name of the group chat room (Figure 3.2c).

Importance Modeling

Now Han turns to the importance modeling for the Lab Notices category. He wants to retain the halos as low importance, but for a long period even after he attends the notifications. Han opens the **Importance Models** tab and sets the **lifespan** to be *3 hours*, which is longer than the usual period that he checks his phone (Figure 3.2d). To keep the low importance throughout the lifespan, Han sets both the **initial** and **target** importance to be 0.5, but the **saturation** importance to be 0.2 as he wants to lower the importance after he checks the notification.

Visual Encoding

To configure visual encoding for the category, Han opens the **Visual Specification** tab. He first sets the **shape** of the symbol to be a *test tube* icon, indicating a "laboratory." To make the symbol noticeable, he configures the **color** to gradually changes from *yellow* to *red* over time, and assigns a *blink* **animation** to be played until attendance. Finally, Han encodes the *current importance value* to the **distance from the center of app** so that the symbol

would move closer to the app icon as the importance decreases. The resulting appearance of Han's halos for KakaoTalk and its dynamic change based on the visual encoding is described in Figure 3.3 and Figure 3.4, respectively.

3.3 Usability Study

To assess the feasibility of the personalization features provided by DataHalo and identify usability issues, we conducted a usability study. The study was approved by the institutional review board.

3.3.1 Study Procedure

At the beginning of each session, we conducted a semi-structured interview asking about participants' usual patterns and strategies of (1) configuring notification feedback, (2) dealing with notifications stacked in the status bar, and (3) managing app icons on the home screen. The interview took about five to ten minutes.

After a short introduction of DataHalo, we installed the DataHalo launcher on participants' smartphones and had a ten-minute tutorial. In the tutorial, we provided them with a short material describing the interface and example usages of DataHalo. We then proceeded to two *halo design tasks* (task 1 and task 2). For each task, participants designed an alternative notification method of a selected app by reflecting on their daily inconveniences. For task 1, participants chose their most frequently used messenger because messenger apps are the most significant source of notifications smartphone users receive. For task 2, participants freely chose any other app freely. During the task, we encouraged participants to think aloud; they were also free to ask questions or request technical support. We also recommended that they try all the customization capabilities to collect holistic feedback.

After completing the tasks, we had a ten-minute interview covering (1) participants' strategies to set up satisfying halos and (2) suggestions to improve the usability of DataHalo. The whole process took about an hour. We rewarded participants with 10,000 KRW (equivalent to 8.50 USD). We audio-recorded both the interviews and the design tasks, logged their halo settings during the design task, and collected the System Usability Scale (SUS).

3.3.2 Participants

From a local university, we recruited 17 participants (U01–17; 4 women and 13 men) who have used smartphones for at least a year and are currently using Android smartphones with version 8.0 Oreo or later. We employed opportunity sampling by recruiting participants through an online forum for university students and faculty members. All participants were undergraduate or graduate students in their 20-30s.

3.3.3 Analysis

Two researchers transcribed the audio recordings from the interviews and the halo design tasks. We then categorized participants' responses in terms of the following topics: (1) the usefulness of DataHalo, (2) the assessment on the halo representation, and (3) the feedback about the customization capabilities. Based on the topics, we investigated 32 completed halo settings (Task 1: 17, Task 2: 15; we excluded two because of incomplete logging) that illustrated how participants categorized notifications, modeled the importance, mapped data variables to visual channels, and encoded graphical marks of the halo visualization for a selected app; we utilized the results to triangulate the interview findings. We graded the collected SUS data according to the traditional school grade [7].

3.3.4 Result

App Selection & Virtual Category Creation

In the first halo design task of a messenger app, 15 out of 17 participants chose KakaoTalk, and the rest (U03, U09) chose the default text messaging app. They thought of diverse virtual categories that curated communication with people of interest (e.g., family, one's beloved, advisor, personal fitness trainer) and that subscribed to their topic of interest (e.g., lab meeting, payment & shipping information of product, a favorite sports team). In the second task, they chose apps from 10 distinct categories: secondary messenger, email, social media, financing, video streaming, movie ticketing, language learning, food delivery, game, and paid survey. We elaborated on all the virtual categories authored during the second task with the reason for creation in Appendix B.1.

Visual Design of Graphical Marks

Figure 3.5 presents examples of the *halos* created by participants. Participants were motivated to combine multiple visual properties to create desired graphical marks. Shape and color were the most popular properties they customized. Pictograph was the most preferred option to encode the shape of a graphical mark; the provided icon set was broad and attractive enough to delineate their intention. U16 worried that images (e.g., portraits of her intimates) or text might easily expose the notification content to somebody in reduced-privacy situations such as shoulder surfing. However, the *image* option could provide users with an enjoyable experience during the design process. P9 was satisfied with selecting an attractive food selfie from her photo gallery when designing *halo* for a food delivery app. When choosing

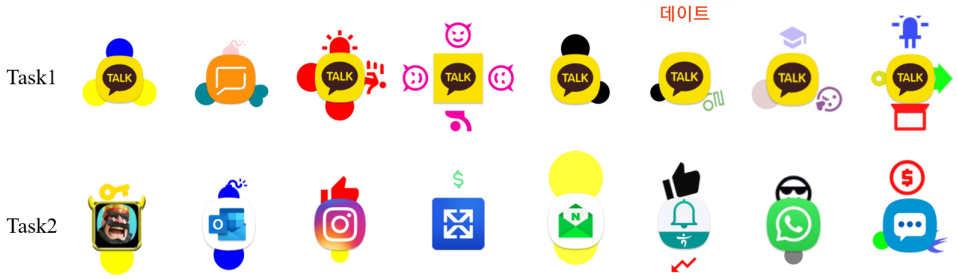


Figure 3.5: Examples of app halos created by participants during the usability study. In the end of each halo design task, we captured the preview from the halo configuration page.

the color, participants considered the followings: (1) to highlight the urgency of information, (2) to tell apart distinct virtual categories quickly, and (3) to reflect their color tastes. They frequently encoded the importance value to the size or distance from the center. Most participants considered animation as an auxiliary option to highlight graphical marks further.

Scalability & Discoverability of Halo Visualization

We asked participants about two possible limitations of halo in our design constraints: the issues of (1) scalability, and (2) discoverability. Firstly, scalability could be problematic if halo must present numerous app notifications. We currently limited the maximum number (currently up to 10) of graphical marks in the halo to prevent visual clutter due to the overlap among graphical marks. All but one (U10) responded that this issue would not matter to them, and they thought conveying the exact number was unnecessary. If it exceeded a certain level, they accessed the notification drawer or the target app. Participants limited the maximum number of graphical marks in a halo to 3.78 on average ($min = 2, max = 10, median = 3, std = 1.92$), which is much lower than our limit. Secondly, participants agreed that the side-

lighting visualization could resolve the discoverability issue adequately. P10 further commented that it would be beneficial when utilizing halo for some apps for which he needs more privacy (e.g., games). He has been reluctant to place these apps on the first page of the home screen due to possible privacy leaks (e.g., shoulder surfing).

Usefulness of DataHalo & Customization Fatigue

Participants' reactions to DataHalo were generally positive. The average SUS of DataHalo was in the range of OK and GOOD. After usage, they all agreed that DataHalo could successfully support their demand for personalized notification management for various apps. In addition, the flexibility of DataHalo also encouraged participants to rethink the value of previously ignored notifications. For example, U09 had deactivated most app notifications except phone calls and text messages since she felt overwhelmed by unnecessary notifications. After completing her design task for the food delivery app, however, she expressed her desire to reactivate more app notifications on DataHalo. On the other hand, they expressed the fatigue of configuration from scratch. Compared to editing virtual categories, they struggled with configuring the importance model and visual encoding, which needed the support of the experimenter.

3.4 Implications and Design Improvements

Despite the promising results from the usability study, we thought it would be crucial to relieve the user's customization fatigue not to demotivate usage of DataHalo in the wild. We improved the customization flow of the importance model (Figure 3.6) and visual setting (Figure 3.7) of a virtual category by modifying predefined or user-created examples. First, we added a set of

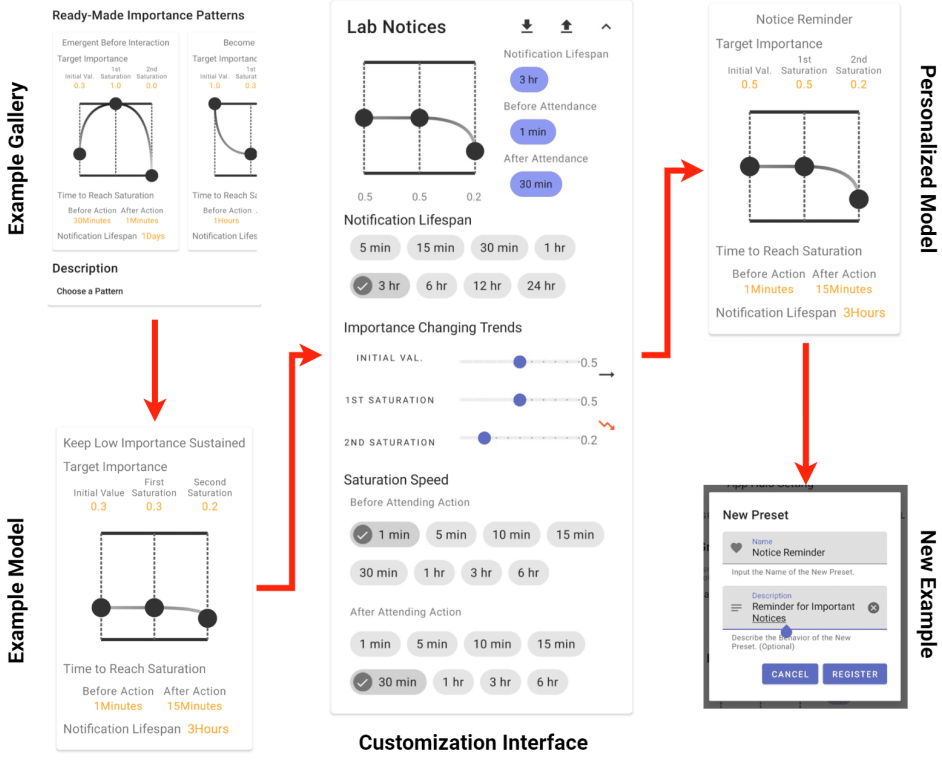


Figure 3.6: Customization flow of example-based importance modeling.

predefined examples for importance models and visual encodings to alleviate the burden of initial customization. People could quickly start customization by choosing an exemplary importance model and visual setting from the gallery. After completing the customization, people could choose to register the results to the gallery. Then, they can easily reuse these new presets when customizing other virtual categories across different apps.

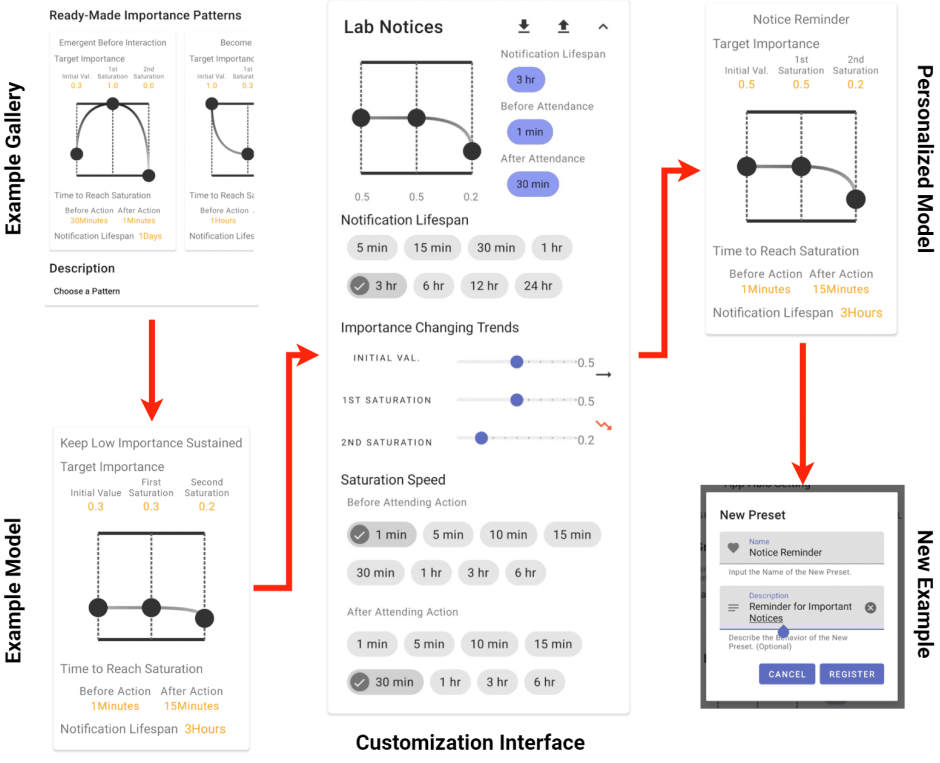


Figure 3.7: Customization flow of example-based visual encoding.

Chapter 4

Understanding User's Design Strategies toward PAAIV in Daily Context

This chapter reports the field deployment study conducted to understand smartphone users' personalization strategies for better longitudinal interaction with notification information, emerging in the real-world context. It covers the second research question: *What would be commonalities and diversities in people's customization strategies for embodying their personalized ambience by utilizing the PAAIV system in everyday contexts?* The improved version of DataHalo was installed into 12 Android smartphone users as an open source launcher. They could apply customization capabilities of DataHalo to any smartphone apps they actualize use in daily context. Participants had interviews before and after deployment. During the three-week deployment, DataHalo usage data and partial, anonymized notification data were collected. As a result of the mixed-method analyses involving usage data, partial notification data, and interview data, participants' customization patterns and underlying design motivations to realize their desirable interaction

with smartphone notifications were revealed. On these bases, the chapter finally covers the suggestions for the future smartphone notification system, which could support end-users to build and refine their PAAIV more easily and expressively.

4.1 Deployment study

With the improved interface, a field deployment study was conducted to explore the commonalities and differences among people's personalization strategies for better notification management, emerging in the real-world context. The study was approved by the institutional review board.

4.1.1 Study Design

Pre-study session

We met each participant on Zoom. We first conducted a semi-structured interview inquiring about daily experiences related to notifications. The questions were identical to those of the pre-study interview of our usability study. Participants then had an exercise of designing a halo for the KakaoTalk app where they (1) created a virtual category to manage the study announcements from the experimenter, (2) set an importance model, and (3) specified a visual setting of a graphical mark. Finally, they responded to the set of apps that they currently planned to manage halo during the usage.

Deployment

Participants immediately started using DataHalo. For the following 3 weeks, we communicated with participants via the 1-to-1 chat room of KakaoTalk. Participants were free to ask questions or request technical support. We sent

announcements such as app updates and weekly summaries of their DataHalo usage. The content of the weekly summary consisted of (1) an initial usage plan summarized from the participants' pre-study interview responses, (2) a list of apps for which they currently manage halo, (3) the most frequently edited halo, and (4) the accumulated number of app notifications influenced by their halo settings; we sent three in total per participant. If some participants did not utilize DataHalo since they received the last week's summary, we added a sentence, *"Do you satisfy with the current settings of DataHalo? If not, how about creating a new halo or adjusting the existing halos for better notification management?"* to the summary of this week.

Debriefing

After three weeks of usage, participants had an exit session for about an hour. We asked participants about (1) the experience of using DataHalo's capabilities of managing notifications, (2) halo design goals and strategies for the apps they customized. We also collected the System Usability Scale (SUS) and a 7-point Likert scale questionnaire (Appendix C.1) inquiring the user experience of notification management before and after using DataHalo. Both the feasibility of core capabilities and their influence on improving the experience of notification management were positively assessed. We rewarded them 50,000 KRW (equivalent to 44 USD) for their complete participation. If participants hoped to continue using DataHalo launcher, they could voluntarily use more.

4.1.2 Participants

We employed opportunity sampling by recruiting participants through the two most active online communities for students and alumni of a local uni-

versity. We advertised our study with an introductory material. The requirements were primarily identical to those of the formative lab study, except for additional conditions: (1) they should be available for a video conference and (2) we log the usage of DataHalo and the part of the information from their daily notifications (specified in Section 4.1.3).

Eighteen people initially applied, and four withdrew during the pre-study session; one dropped because of the compatibility issue with the open source launcher, and the rest resigned because three weeks were too long for them to participate. We also screened out two applicants because they did not utilize DataHalo during the deployment despite our additional reminders. Of the final 12 participants (six males and six females; we call them P01-P12 in the following), 11 participants were in their 20s, and one was in her 40s. Seven participants were undergraduate students, 4 were graduate students, and 1 was a full-time employee. They utilized ten different Android smartphones from two manufacturers (9 Samsung models and a single Xiaomi model); the distribution was consistent with the market share of Android smartphones in South Korea, where Samsung occupied 91% [88]. We sent 2-3 messages per week to each of them, including the announcements and the weekly summaries. After completing participation, four participants (P02, P03, P05, P07) hoped to continue using DataHalo. We monitored their extra usage for about three more weeks.

4.1.3 Technical Specifications

Communication with Participants

Both pre- and post-study sessions were video conferences remotely held through Zoom and recorded for interview analysis. We utilized Android TeamViewer QuickSupport app for screen sharing and remote technical support. During

the deployment, we used the open chat feature of KakaoTalk to communicate with participants; it did not require any personal information and was easily accessible for most South Koreans.

Update Releases

We offered participants an exclusive invitation to install DataHalo via Firebase App Tester for Android. They could receive notifications about app updates and install the latest version from the tool. We monitored the app distribution from the Firebase console and kept participants up to date with DataHalo; we contacted them if they did not install the necessary update. There were five updates to improve (1) app crashes reported by Firebase Crashlytics, (2) device compatibility issues regarding the screen resolution & aspect ratio, and (3) log transfers during the bad data connection (e.g., flight mode). We managed the distribution of the latest version among participants from Firebase console and contacted them via KakaoTalk open chat if they did not install the necessary updates.

Log Data Collection

To understand (1) how participants utilized DataHalo and (2) how the usage affected their interaction with daily notifications, we logged **halo setting logs** and **notification logs**. The halo setting log, the latest configuration of a halo of an app, consisted of (1) virtual categories and their keyword elements, (2) parameter settings of user-defined importance models, and (3) visual encoding results of user-defined visual settings. A notification log was generated when it ended its informative lifespan based on its importance model. Each notification log included information about (1) the identification of the virtual category (e.g., user-id, app name, virtual category label), (2) Android

notification channel (e.g., channel name and system importance), and (3) user interaction (e.g., post time, interaction time, the system flag about the reason of notification removal, and lifespan). We did not log and transmit the sensitive part of notification contents (e.g., title, subtitle, text, and image) outside participants' smartphones. DataHalo sent a stack of halo setting logs and notification logs to Firebase Realtime Database (RTDB) [33] hourly.

4.2 Data Analysis

In the following, analytic approach toward transcriptions of the pre and post-study interviews, halo setting logs, and notification logs was elaborated in detail.

4.2.1 Data Preprocessing

We transcribed the responses of the semi-structured interview. Regarding the interview of the pre-study session, we grouped the transcriptions according to the topics describing their current notification management: (1) filtering, (2) feedback setting, (3) attending, and (4) accessing information in a longitudinal manner. We also categorized the responses about the current home screen management: app shortcuts and their layout.

Regarding the interview of the post-study session, we summarized the transcriptions into a table that listed the participants' considerations during each app halo customization from the following perspectives: (1) their high-level motivation of app choice, (2) notification filtering & categorization, (3) importance modeling, and (4) visual encoding; we named the table as **halo design table** in the following. The table consisted of 91 cases of app halos; it included 205 design scenarios of the graphical mark applied to each virtual category.

Regarding the collected halo setting logs, we reconstructed the temporal events of DataHalo customization; we denoted the data as **halo customization events** in the following. Halo customization events consisted of following properties: (1) *participant_id*, (2) *timestamp*, (3) *target_app*, (4) *app_category* provided by Google Play Store, and (5) the halo configuration at that time.

Regarding the collected notification logs, the following notifications were filtered out to obtain the notifications of interest with which participants actually interacted.

1. Notifications sent by the Android system (e.g., charging state)
2. Notifications not clearable by users (e.g., media control of music/video players accessible in the notification drawer)
3. Notifications controlled by app services (e.g., a summary notification of unread messages in KakaoTalk; the app automatically repeats post and removal whenever the number of unread messages changes.)

Among the filtered notifications of interest, we separated ones from the apps for which participants managed custom halos; we named these notifications as **halo notifications** in the following.

4.2.2 App-Level Customization Analysis

First, we aimed to understand the participants' DataHalo usage in app level: choice of apps and customization over time. We labeled the halo customization events based on (1) customization type (i.e., creation or edit) and (2) weeks (i.e., first week, second week, third week, and extra weeks). We then investigated halo creation and edits over time by app category and participant. Finally, we reviewed the adjacent edit events and the halo design table to understand why participants repeatedly edited app halos over time.

4.2.3 Virtual Category-Level Design Analysis

Next, we aimed to understand the design motivations and choices underlying participants' halo designs. Because the virtual category was the unit of notification management in DataHalo, we drilled down into virtual category-level analysis. The analysis proceeded in two steps. First, we derived the high-level **design motivations**. We analyzed 205 graphical mark design scenarios in the halo design table referring to (1) our design considerations, (2) the major findings from the usability study, (3) user–notification interaction patterns summarized from the interview results of the pre-study session, and (4) design dimensions or evaluation criteria of ambient information visualization [63, 65, 85]. Next, we derived participants' **customization choices** regarding these motivations. We reviewed the last halo customization events of 91 app halos, which included the final configuration of 205 virtual categories. For each virtual category, we compared the corresponding motivation described in halo design table with the relevant part of the halo configuration; we validated if the configuration could output the graphical mark that behaves as participants described. If two data were consistent, we labeled the pair of description and configuration. we iteratively refined and grouped similar labels, and named the final set of labels as choices.

4.2.4 Halo Notification Analysis

We also aimed to understand the similarities and differences between the virtual categories made by participants the notification channels carved by app providers. We first investigated the volume of halo notifications by participant and app category. We then compared (1) the category and (2) the importance of notification carved by the system (i.e., notification channel and

channel importance of the Android System) to those defined by participants (i.e., virtual category and user-defined importance).

4.3 Result

This section starts with a brief summary of the conventional user–notification interaction before using DataHalo. Next, the results of analysis on (1) app-level usage, (2) virtual category-level design, and (3) halo notifications were presented. Finally, usability of DataHalo and its contribution to a better user experience of notification management were reported.

4.3.1 Notification Management and Attendance before Using DataHalo

Most participants (9 out of 12) have struggled to control the volume of app notifications. They decided to disable the advertisements when installing an app or block all the notifications of some apps (e.g., messengers, social apps) later in the usage. To avoid being distracted, most (except P01) preferred less noticeable feedback (i.e., vibration or silence) over the sound as a default feedback method regardless of their spatial or eventual context. It led participants to frequently attend a stacked list of notifications in the drawer when they visited the smartphone home screen. Most (except P08) frequently experienced the need to access some notices for a longer duration when attending notifications. They kept the notification untouched in the drawer or manually duplicated the content using other services (e.g., calendars, image capture). Most participants (except P11) preferred to locate the frequently used app shortcuts on the first page of the home screen, where they could access the apps quickly. Other criteria (e.g., proximity among the apps with similar features, comfortable area for touch, and aesthetics) also influenced

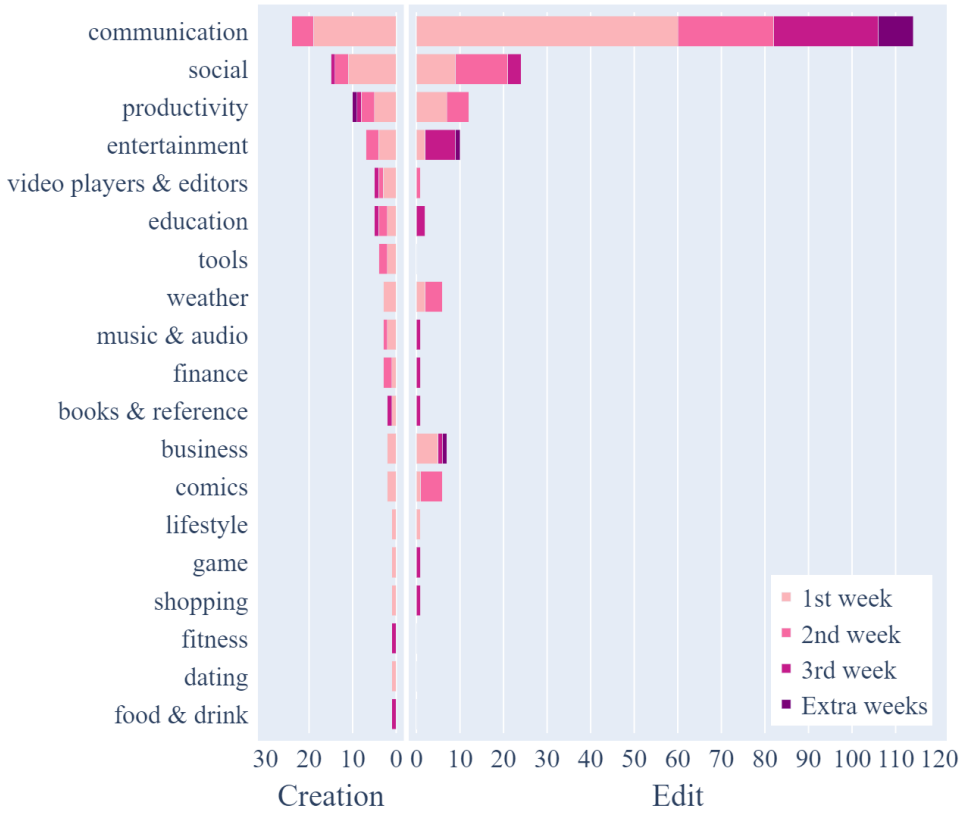


Figure 4.1: Halo creation and edit events by app category. We aggregated the events according to app category provided by Google Play Store. The color encodes the week the event occurred.

their home screen layout. Participants eventually placed some useful apps on the subsequent pages, and they thought the badges of these apps were less effective; they could not glimpse the badges immediately.

4.3.2 App-level Usage of DataHalo

Overview

In total, participants managed 91 halos of 55 different Android apps from 19 app categories provided by Google Play Store (Figure 4.1). On average,

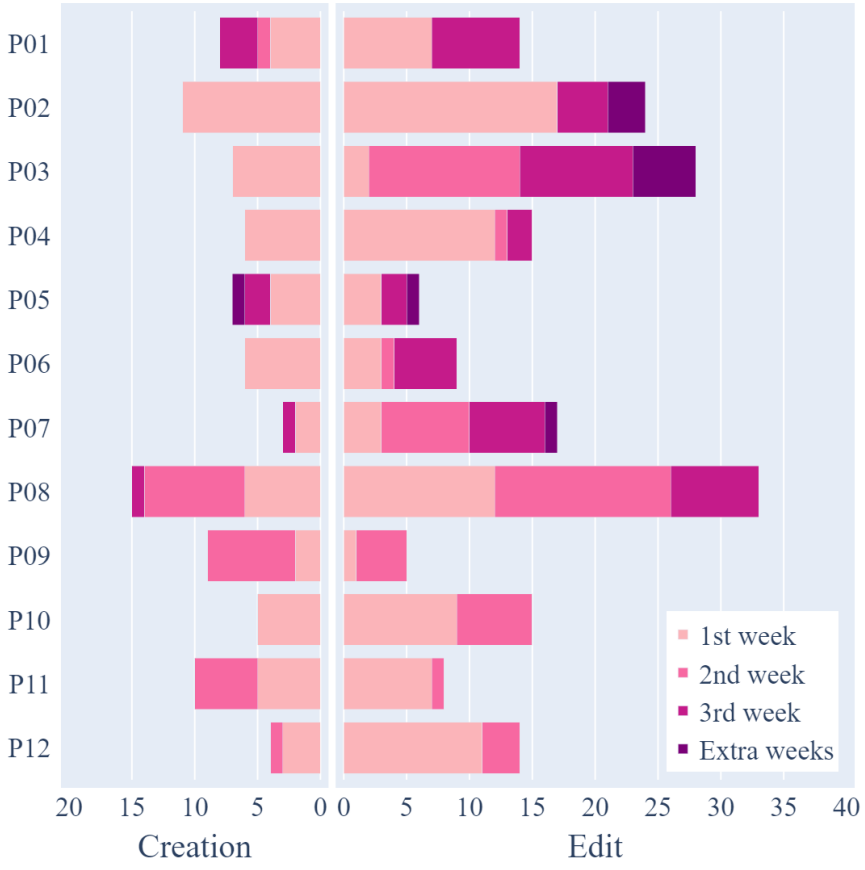


Figure 4.2: Halo creation and edit events by participant. The color encodes the week the event occurred.

participants have created 7.58 app halos $(min = 3, median = 7, max = 15)$. After the creation, participants edited their app halos 188 times in total. They edited the app halos 15.67 times on average $(min = 5, median = 14.5, max = 33)$ during the usage (Figure 4.2). The edit count per app halo was 2.07 times on average $(min = 0, median = 1, max = 18)$. 51 out of 91 app halos have been edited by participants at least one additional time after creation; these apps covered 14 out of 19 categories.

The major categories, where the number of the created app halos was more than five, included the following apps.

1. **Communication** (24): KakaoTalk (12; instant messaging app), Messages (6), etc.

2. **Social** (15): Instagram (5; social media), Between (2; relationship app for couple), EveryTime (2; community app for university students), etc.

3. **Productivity** (10): Calendar (5), Email (3), etc.

4. **Entertainment** (7): KakaoPage (3; web comic & novel service), Netflix (2; video-on-demand service), etc.

The communication category accounted for 61% (114) of total edits, and KakaoTalk (94) mainly contributed to the result. Participants edited KakaoTalk 7.83 times on average ($min = 2, median = 7, max = 18$).

Halo Creations & Edits over Weeks

Overall, halo creation and edits decreased in a different pattern over time. Most participants completed the creation of halos during the first two weeks. Of the 91 app halos, participants created 67% (61) in the first week, 24% (22) in the second week, and 8% (7) in the third week. The rest, 1% (1), occurred during the extra usage of four participants. The number of halo edits also decreased over time, but the speed was slower than that of creation. Of the 188 edit events, 46% (87) occurred in the first week of usage, 26% (49) in the seconds week, and 22% (42) in the third week; the rest 5% (10) occurred during the extra usage. Figure 4.2 supports the difference. Five participants (P02, P03, P04, P06, and P10) completed halo creation in the first week but

continued editing afterward. P02 and P03 had consistently edited app halos of messenger, email, and webtoon services even during the voluntary extra usage to adjust the change. Their primary reason for the edit was to newly create the virtual category or refine the existing ones because new communication partners or webtoons of interest appeared.

Halo Edits for New Notifications of Interest

Participants became constantly motivated to manage notifications of new interests within an app halo; the motivation originated from themselves (e.g., P03's need to curate messages from her intimates) or their surroundings (e.g., P11's need for a subscription to the confirmed COVID-19 cases that occurred on campus). They either created a new virtual category or expanded the keyword-based rules of the existing ones to reflect the new interests; the choice depended on how definite/flexible they conceptualized the virtual category. Participants managed specific virtual categories because they could control notifications finely, especially using various graphical marks. For example, P03 started with a virtual category for curating messenger notifications sent from a friend. Finding the customization helpful and joyful, she created four new virtual categories by person and assigned distinct images to them. On the other hand, participants could also finish editing faster by adding new keywords to the virtual category of broader concepts.

Halo Edits for Improvements for Current Category

Participants edited halos when they faced a critical but unseen notification that the current category could not successfully classify. They experienced both the user-side and the app provider-side issues while improving the keyword-based rules of their virtual categories.

When initially building a virtual category, participants skimmed the recent app notifications from the notification drawer or the notification history in our halo configuration page. When the relevant app notifications users could reach were too small, the initial keyword-based rules needed improvements to actualize their goal. For the apps triggering notifications frequently (e.g., messenger app), participants could realize the problem and complement the rules. However, P09, who created a virtual category that subscribed to an irregularly updated podcast, needed a long time until she received the new notification.

P03 explained the "advertisement cop" issue caused by app providers while managing the keyword composition of virtual categories. When she tried to subscribe to notifications about the arrival of new short pants for the fashion app, she found that her keywords failed to hit these notifications. Their notification text, colloquial expression written by a copywriter, changed frequently and utilized synonyms or indirect expressions so that the current keywords became quickly outdated.

We could observe a primitive use case where a participant utilized the app-provided notification control method and DataHalo at the same time. P07 utilized the structured notification features provided by his blog app; he could figure out what notifications the app would trigger easier than looking up the list of the received notifications. He disabled the unnecessary notifications from the app-provided menu and moved to halo customization. Because he thought all the notifications the app triggered were informative, he did not create any virtual category and just edited the graphical mark of the Remainder category. P09 also commented on unifying the app-specific methods of Kakaotalk into the DataHalo interface. She came up with the idea

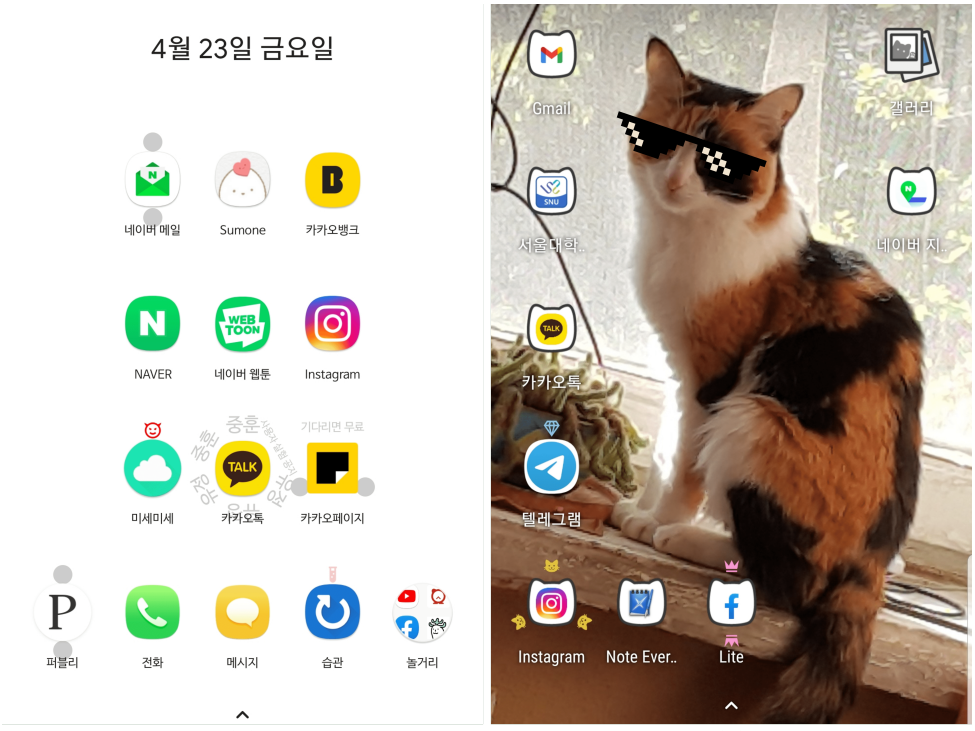
while filtering and categorizing diverse notifications, but did not take action; she felt the app-provided features added another complexity.

Null Halo and Removal from Home Screen

Participants disabled halo visualization for 11 apps in total. They either hid all the graphical marks in the halo configuration setting or removed the halo from the home screen. As a result of the former approach, participants got a "null halo"; they chose it when app notifications were useless, but they wanted to use the app on the home screen. For example, P08 set null halos for the app store, web browser, gallery, and banking apps to maintain his home screen neatly. Participants chose the latter when they decided not to use an app on the home screen. For example, P05 had become indifferent to the matching service provided by the dating app, so he blocked all the notifications and finally dropped the app from the home screen. P01 had removed the halo for the webtoon app, although subscriptions she made were still valuable for her; she worried her halo visualization, which highlighted the update for her favorite webtoons, might disturb her concentration during the midterm period.

Contrasting Use Cases of DataHalo

We present use cases of P02 and P06 that indicate the diversity of DataHalo usage. P02 (Figure 4.3a) was one of the most passionate users of DataHalo according to the number of halos and the total number of virtual categories she created. She managed 42 virtual categories (31 user-created and 11 Remainder categories) in total for 11 halos. The app halo of KakaoTalk even included ten virtual categories. She tried to control the notifications of her interest as sophisticated as possible. Because P02 attached the highest pri-



(a) (b)

Figure 4.3: Personal difference in usage of DataHalo. (a) P02 pursued pragmatism while (b) P06 valued aesthetic experience. Upon P06's request of not disclosing the photo of her favorite cat directly, we masked its face with a sunglasses.

ority on the discernibility of multiple graphical marks, she frequently presented the name of virtual categories directly (i.e., text) or assigned unique pictographs to different virtual categories. Her home screen wallpaper was also solid white, supporting her design goal of extensive and effective notification management.

On the other hand, P06 (Figure 4.3b) created a minimal number of categories. She created only three custom categories in total while managing six halos. Besides two communication apps (KakaoTalk and Telegram), P06 was not motivated to classify notifications. Instead, she made a great effort

to design a home screen that reflected her aesthetic preference. First, she laid out halos where they did not occlude the image of her favorite cat on the wallpaper. Second, she decorated halos by selecting pictographs irrelevant to the app or notification but visually satisfying (e.g., cat for Instagram and crown for Facebook).

4.3.3 Participants' Design Motivations and Choices

As a result of the virtual category-level design analysis, we derived participants' five **design motivations** (**M1-5**) and corresponding **design choices** (see Figure 4.4), consisting of an app-level and four virtual-category-level motivations. We provide an exhaustive list of the virtual category-level scenarios in the supplementary material.

M1. Control the Scalability Issue of App Notification Control

Participants left 54% of the apps (49 out of 91 apps) **as-is** when they had no desire to filter. For 34% (31) of them, participants mixed various features, including Pass-List, Block-List, and the maximum number of visualized marks, to **selectively control**. Participants tried the selective control for 17 of 24 communication apps and 4 of 15 social apps. Lastly, participants **blocked** notifications of 12% (11) of the apps by setting zero to the number of visualized marks.

M2. Classify Notifications based on User-defined Schema

Participants prioritized **self-organized events**, which involved scheduled events, to-dos, or daily goals created and registered by themselves. Their **curated communications** focused on senders of interest who communicate with them via messenger, email, and social apps. **Personal subscriptions** covered di-

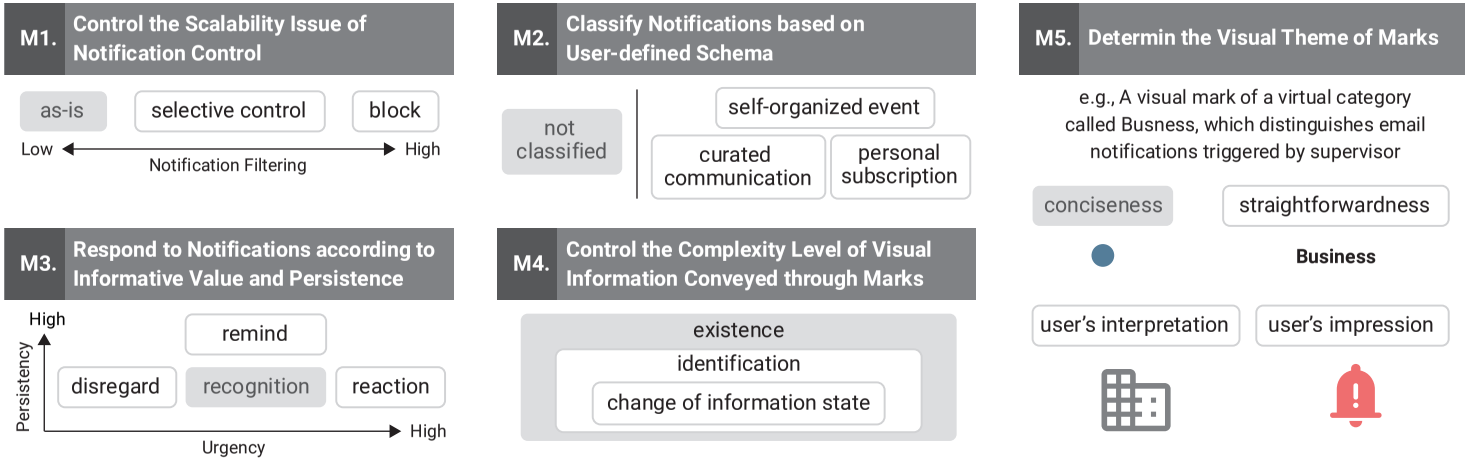


Figure 4.4: An app-level design motivation (**M1**) and four virtual category-level design motivations (**M2-5**) with corresponding choices. For each design motivation, the grey-colored design choice implies that user had the minimal motivation to customize their interaction with notifications of a designated virtual category.

verse topics of interest. The notifications participants **did not want to categorize** belonged to the Remainder category. However, some of them naturally covered the notifications of interest (Figure 4.5a); some apps (e.g., to-do app) triggered only the notifications that the user organized; P07 filtered the blog app notifications using app-provided features so that he could subscribe to notifications of interests from the Remainder category.

M3. Respond to Notifications according to Urgency and Persistence

Participants wanted to **recognize** the useful but non-critical notifications in the same way they used to; they left the default importance model untouched. They wanted to **react** immediately to urgent notifications by imposing high initial and target importance values but a zero saturation importance to the category. They wanted to **remind** themselves of persistently helpful notifications by setting the long lifespan and keeping the saturation importance non-zero. They **disregarded** unnecessary notifications by setting the initial, target, and saturation importance to zero or the lowest value within the app. Figure 4.5b indicates that participants chose more diverse responding approaches for the user-created virtual category, while they thought that the recognition was enough for the notifications of the Remainder category.

M4. Control the Complexity of Visual Information Conveyed through Marks

Participants expected the graphical mark to behave similarly to the dot-shaped notification badge by which they could only guess if any notification **exists**; they left the default visual encoding untouched. Participants made the graphical mark **identifiable** by assigning a distinct combination of visual properties (e.g., shape, color) to the default encoding. Participants addition-

ally wanted to track the **change of information state** by configuring a more complex visual encoding, starting from our predefined examples. In terms of visual variables, the most preferred visual encoding was the following:

- $user_checked \rightarrow shape$ (i.e., the shape changes after the user attends to notification.)
- $elapsed_time \rightarrow color$ (i.e., the color changes over time.)
- $user_checked \rightarrow motion$ (i.e., the motion changes after the user attends to notification.)
- $importance_value \rightarrow size$ (i.e., the size is proportional to the current importance.)
- $importance_value \rightarrow position.$ (i.e., the position is proportional to the current importance.)

Participants tended to configure the more complex visual encoding for the user-created virtual categories (Figure 4.5c). However, P06 managed the category "non-urgent group chats" in her messenger and assigned an unidentifiable graphical mark to disregard the targeted messages.

M5. Determine the Visual Theme of Marks

Participants pursued **conciseness** by utilizing the default graphical mark (i.e., circle) as-is or with color change. They pursued **straightforwardness** by presenting the name of the virtual category directly. Participants wanted to **represented their interpretation** of the category (e.g., a pictograph of the office building). Participants were motivated to express **their impression** to the mark (e.g., a bomb-shaped icon to express the negative and pressing moods). Figure 4.5d indicates that participants applied diverse visual

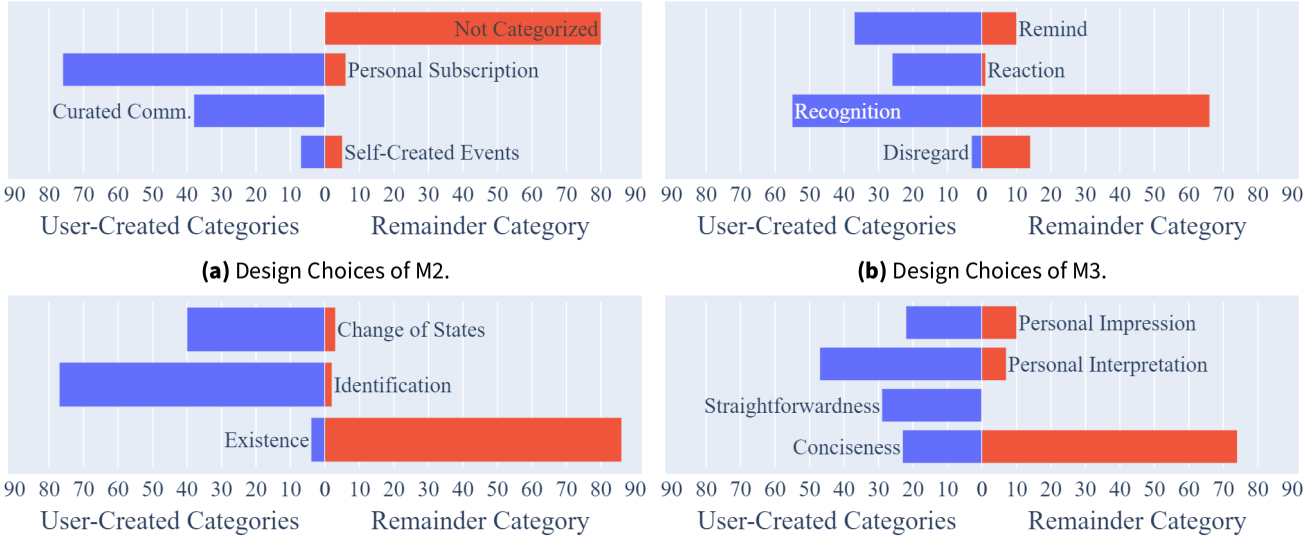
themes to the user-created virtual categories, while they generally favored concise graphical marks for the Remainder category. We also observed the individual difference regarding the preferred theme of the mark. We revisit P02 and P06's usage in Section 4.3.2. 43% (18 out of 42) of the graphical marks created by P02 highlighted straightforwardness, while P06 created 56% (5 out of 9) of the marks emphasizing her impression.

4.3.4 Virtual Category-Level Design Strategies

We denoted a combination of the design choices along the five design motivations participant made for a virtual category as its virtual category-level **design strategy**. In the following, We further describe the commonalities and diversities among the participants' halo designs by comparing their design strategies. We listed the selected design strategies that were (1) frequently adopted and (2) widely shared among participants in Appendix C.2.

Explanation of Terminology by Example

Figure 4.6 describes the relationship between the design motivations, the design choices, the virtual category-level strategy, and the graphical mark with a halo of Loop Habit Tracker created by P02. It was a self-tracking app where she registered her daily goals and monitored their achievement. Although the app could send notifications that remind her of daily goals, they easily disappeared when "dismissing all" from the drawer. So, she decided to manage a virtual category called "Supplements" that reminded her of a daily goal to take nutritional supplements. Our design motivations (**M1–5**) provide the five viewpoints of P02's design scenario. Next, our design choices regarding these motivations give a detailed explanation of her customization behaviors. Finally, the resultant virtual category-level design strategy —M1:



(a) Design Choices of M2.

(b) Design Choices of M3.

(c) Design Choices of M4.

(d) Design Choices of M5.

Figure 4.5: Distribution of design choices for the user-created virtual category and the Remainder category.

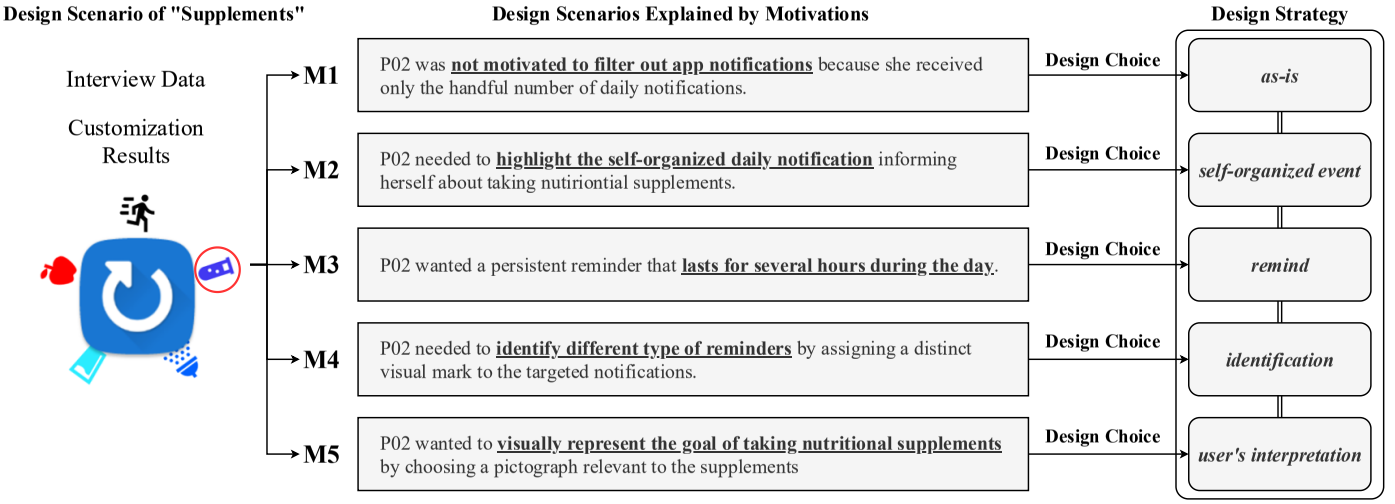


Figure 4.6: Relationship between design motivations, choices, virtual category-level design strategy, and graphical mark of P02's Loop Habit Tracker.

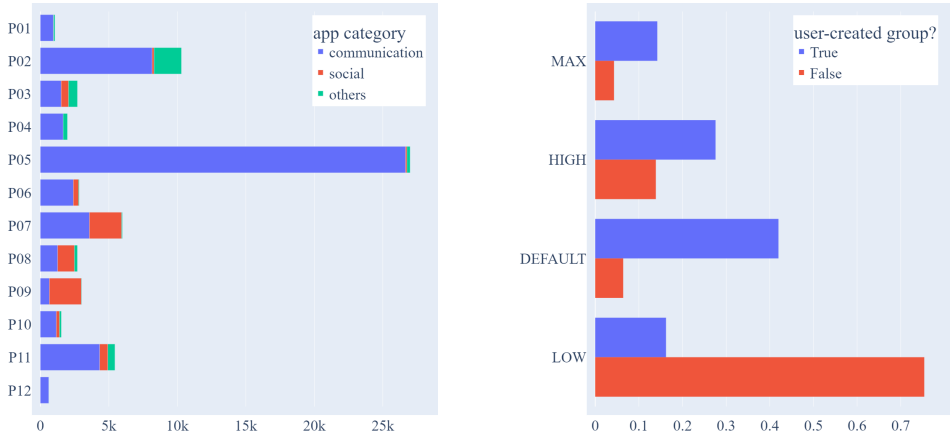
as-is, M2: self-organized event, M3: remind, M4: identification, M5: personal interpretation— comes out; it enables the scenario-to-scenario comparison with other virtual categories made by P02 and by other participants.

No Intersection between Design Strategies of User-Created Category and Remainder Category

We could observe 74 unique design strategies from the 205 virtual category-level design scenarios; 50 of them were for the user-created virtual categories (118 out of 205); 24 of them were for the 87 Remainder categories. There was no intersection between two sets of unique strategies, indicating that participants posed distinct design approaches based on whether they were actively involved in category creation.

Individual Differences of Design Strategies Applied to User-Created Category

Participants tended to reuse the same design strategies multiple times for similar virtual categories in the same app or across apps. 58% (7 out of 12) of participants reused their design strategies. For example, P02 repeated her three design strategies nine, eight, and five times, respectively. Individual design strategies hardly overlapped among participants. 1.61 Participants shared the same design strategy on average ($median = 1, max = 4, std = 1.60$). Only 26% (13 out of 50) of the design strategies were shared by two or more participants. The most widely shared strategy, used by four participants, was to subscribe to the experiment notices (Section 3.2.4) provided as an exercise; all participants initially utilized the same setting, but eight participants changed the strategies by further customization during their usage.



(a) Halo notifications by participant.

(b) Distribution of Android channel importance by type of virtual category.

Figure 4.7: The result of halo notification analysis.

Homogeneity of Design Strategies Applied to Remainder Category

As illustrated in Figure 4.5, design strategies for the Remainder category were more homogeneous than those for the user-created category: The two most frequent strategies took up 55% (48 out of 87) of usage. These two were identical except for the choice of **M1**. They corresponded to the default configuration of the Remainder category; participants did not customize because they focused on managing the categories they created.

4.3.5 Halo Notification Analysis

Overview

Of the 1,088,368 total notification logs collected, we filtered 114,316 notifications of interest. Of the 114,316 notifications of interest, halo notifications occupied 65,200. The communication category was prevailing in halo notifications (52,995 out of 65,200); KakaoTalk dominated this category (51,023 out of 52,995). The second most frequent app category was social (7,910

out of 65,200). Therefore, these two categories, related to communication among users, were the primary source (93%) of halo notifications. On average, participants received 5433.3 halo notifications ($min = 627, median = 2,773, max = 26,992$, Figure 4.7a).

Comparative Analysis on Virtual Categories and Android Notification Channels

At a high level, we found some commonalities between the notification importance imposed by the user's information schema and one provided by the channel structure of the current notification system. We observed a significant difference in the distribution of notification's channel importance (Figure 4.7b) based on whether it belongs to the user-created virtual category or the Remainder category (i.e., not classified by user schema). Similarly, notifications that belonged to user-created categories also had higher channel importance than the Remainder categories. It coincided with the observation from the halo design table; all but one participant imposed a higher or equal importance value to the virtual category they created than the Remainder category automatically generated.

However, the detailed investigation contrasted the schema of virtual categories with Android notification channels. Virtual categories frequently accommodated heterogeneous notification channels. 41% (39 out of 94) of virtual categories included two or more distinct notification channels ($min = 1, max = 9, mean = 1.54$). For example, the "announcement" category created by P01 for KakaoTalk assigned a higher informative value to the designated notifications. However, the halo notifications showed that the notifications from two different Android notification channels, "New Message" and "Notification without Feedback", were classified into her "announcement"

category. These channels assigned contrasting system importance ("New Message" $\rightarrow$ *IMPORTANCE MAX*, "Notification without Feedback" $\rightarrow$ *IMPORTANCE LOW*) to notifications, which shows a discrepancy with user motivation.

Moreover, the categorization based on notification channels was not specific enough to cover the user's schema. 38% (31 out of 82) of Android notification channels involved two or more virtual categories($min = 1, max = 8, mean = 1.77$). For example, P02 managed five virtual categories for MiseMise, a weather app specialized in forecasting air pollution, to subscribe to five-graded fine-dust alarms. On the other hand, the logs indicated that its app provider labeled all the differently-graded alarms with a single Android notification channel called "MiseMise Alarm". In addition, channel importance itself could not effectively differentiate virtual categories. For example, four different virtual categories created by P11 for KakaoTalk involved the identical channel importance, *IMPORTANCE DEFAULT*, although she applied distinct importance models to all of them.

4.3.6 Usability and User Experience

Usability Feedback

Compared with the prior result of the formative study, there was no significant increase in the score of the System Usability Scale (SUS). The score was in the range of OK-GOOD again [7]. However, the subsequent inquiry revealed that the fatigue of customization, the main usability issue of the early prototype, has been successfully alleviated.

User Experience of New Notification Management

Regarding the 7-point UX questionnaire (Appendix C.1), the distribution of responses for every question was left-skewed (Figure 4.8). All participants agreed that these capabilities provided a better experience of notification management: (1) notification categorization, (2) prioritization of information, (3) longitudinal interaction, and (4) delivery in a favorable visual form. Participants who favored managing aesthetic halos (P01, P03, P06, P10) especially loved the capability of personalizing graphical marks for the notifications of interest. For example, P10 responded that she voluntarily had a brainstorming session of decorating halos with her friend for about an hour. She thought DataHalo could appeal to her peers interested in decorating their smartphone home screen.

As a result, all participants responded that their user–notification interaction improved overall during the deployment. They hoped to use the full version of DataHalo from the app store soon. On the other hand, the launcher-related usability issues resulted in less positive responses to Q7, "Continue to Use". Many participants wanted to continue using the DataHalo capabilities in their original launcher (e.g., One-UI). P4 also mentioned that he wanted to use DataHalo with 'NotiStar', a notification management service provided by the One-UI launcher; it supported the global search over the notifications that can harmonize with our app-specific method.

4.4 Discussion

4.4.1 Customization Support towards the Personalized Ambience

Our studies revealed diverse perspectives on the satisfactory user–notification interaction among participants. In the casual context [86], participants used

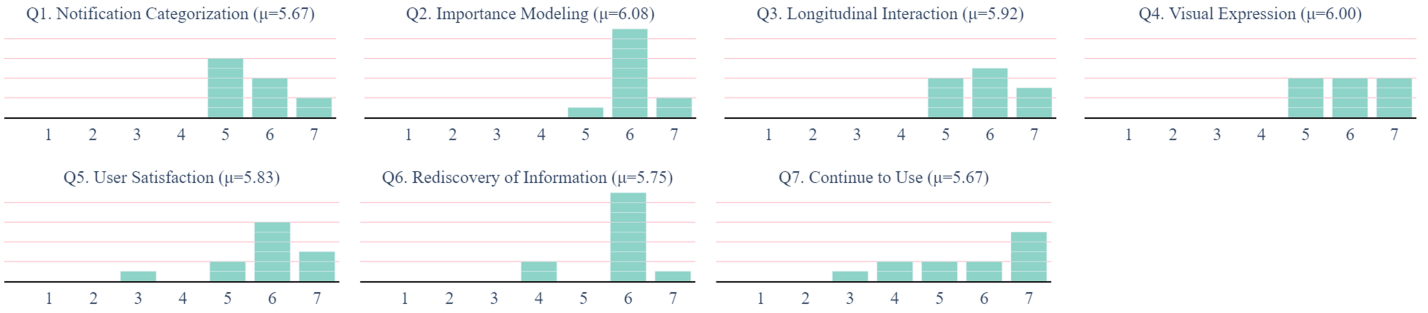


Figure 4.8: The result of 7-point Likert scale questionnaire about the UX of DataHalo.

DataHalo for a broad range of purposes not limited to obtaining analytic insights from notification information; some wanted to decorate their home screen prettier, while others carefully chose graphical marks that implied their emotion. In designing halos, participants reflected their personal preferences (e.g., aesthetic taste) and usage contexts (e.g., app usage, home screen management), producing highly personalized and unique ambient displays. This suggests the potential of DataHalo's customizability in supporting individualized notification management. Although participants generated highly individualized halos, they also tended to apply the same design strategies for multiple virtual categories they created. To streamline such a unified but individualized halo configuration, a future version of DataHalo may support global customization across multiple apps. This could lower the burden of people who want to repeat similar design strategies for virtual categories across multiple apps.

4.4.2 The Ease of Adaptation to Trace the Moving Sweet Spot of User Experience

Another important lesson of our deployment study was the change of personal 'sweet spot' in the user experience over time. Participants added, edited, and even removed halos during usage, similar to personal information trackers [27]. Although participants determined most of the app halos during the first two weeks of usage, they constantly edited halo configurations for extended periods; their sweet spot (i.e., well-personalized setting at a specific time) needed to be updated to adjust the changes originated from themselves and their surroundings. The result indicated that the customization capabilities should be more accessible for people to keep up with the shift of sweet spots. DataHalo successfully exploited the accustomed interaction

methods of the home screen, app icons, and notification badges. Users could intuitively organize an apt information display that adapts to their desire. For instance, they could include or exclude information sources by placing and removing app icons on the home screen. Plus, they could quickly access the halo customization interface of a target app thanks to the context menu triggered by directly manipulating its icon. Likewise, our design study provides an example of integrating novel ideas into widely known interfaces and interaction methods by concinnously extending their underlying semantics.

4.4.3 Generalizability beyond Culture and Population

Since we conducted the study in South Korea, some halo design patterns may be subject to regional characteristics of app usage, given that halo creations and edits were active for *communication* and *social* apps. Especially, KaKaoTalk, a communication app most South Koreans use (over 40 million [46]), is more than a messenger. It incorporates various platform services such as shopping, so many commercial chatbots also send messages (e.g., advertisement, invoice, and delivery) to people. Receiving both chat and non-chat messages makes KaKaoTalk a notification-intensive communication app. Most participants needed to curate notifications from this app and subscribe to only useful ones with distinct longitudinal interactions. Because (1) DataHalo successfully covered a case of a notification-intensive app that triggered diverse information and (2) our analysis drilled down into the virtual category level, we believe our findings also apply to notification management of other apps in other regions.

4.4.4 Implications for Better Personalization Support

Integrating User Schema and Android Notification Channels

There was a discrepancy between the categorizations for app notifications that participants implemented using DataHalo and Android notification channels defined by app providers. This gap implies the necessity of the user's participation in modeling the notification information of an app. Furthermore, P03's "advertisement cop" issue revealed the potential incompleteness of inductive methods provided by DataHalo. The exhaustive information for the app notifications from the app provider can supplement the quality of keyword-based rules that relies on people's observable range of notifications at the time. Combining the user's schema and the app-specific control methods could lower the burden of filter and categorizing notifications. However, this approach is not currently viable due to the limited access to Android's notification information. Therefore, future work may consider integrating the notification channel structure of Android with the visualization pipeline of DataHalo.

Semi-automated Tuning of Importance Model

Despite our efforts to alleviate the user's burden of customization by providing more intuitive examples and user interfaces for visual summary and reuse, some participants also reported the learning cost of tuning parameters for predefined models. To tackle this issue, we expect to utilize the set of design motivations and strategies as a source of questions (e.g., "What type of information you classified?", "How do you want to respond to this notification?") and options (e.g., "Personal subscription," "I want to remind the notification for a while"). Based on the similar design choices in the collected

patterns, DataHalo could suggest more sophisticated initial parameters than those of example patterns. In the long run, these parameters could be further adjusted based on user statistics computed by collected notification logs (e.g., the number of notifications and average interaction delay). This semi-automated tuning of the importance model would effectively resolve the remaining usability issues and possibly adapt to the temporal change of the personal sweet spot we mentioned above.

Chapter 5

Exploring Design Space of Attention-Adaptive Ambient Information Visualization

This chapter[1] examines the feasibility and potential of incorporating attention-adaptive technologies in ambient information visualization design. It addresses the third research question: *What are the limitations and opportunities for integrating attention-adaptive personalization into PAAIV, considering technical constraints and user needs related to real-time attentional information?*.

User attention is a crucial factor of ambient information visualization, yet designers have had to rely on indirect methods for assessing user attention, such as behavioral observations or subjective reflections, which may not accurately capture real-time changes. Integrating Brain-Computer Interface (BCI) methods into the design of ambient information visualization could be an effective approach to overcome this limitation. Researchers have actively studied brainwaves measured non-invasively by Electroencephalogra-

[1]This research is in preparation for publication. In this chapter, "we" refer to myself and my collaborators: Dantae An, Hyunjoo Song, Bohyung Kim, and Jinwook Seo.

phy (EEG) to predict human cognitive states. Moreover, advances in mobile technology have allowed people to use their brainwaves in everyday contexts with end-user EEG devices, which are more comfortable to wear and more affordable than professional instruments used in laboratories and hospitals.

On these bases, we proposed Attention Adaptive Ambient Information Visualization (A3InfoVis), a type of ambient information visualization that leverages users' attention states inferred from real-time EEG data. We explored its design space in two aspects: (1) technological constraints of available attention information and (2) users' needs for utilizing and visualizing attention information.

First, we identified three design challenges concerning technological constraints: (1) classifying attention states for everyday tasks rather than highly-controlled cognitive tests in real-time, (2) addressing the limitations of end-user EEG devices (e.g., few channels and narrow brain coverage), and (3) obtaining sufficient data for a personalized classifier targeting attention states of interest. To tackle these challenges, we collected EEG signals from 25 subjects, assessed the applicability of heuristic indices related to user attention, and compared the performance of machine learning (ML)-based attention-state classifiers.

Next, we interviewed seven users of DataHalo [37], a customizable ambient information visualization system for smartphone notifications, who had experience in creating and managing multiple custom ambient information visualizations in their daily lives for at least three weeks. We inquired about (1) how they would enhance the ambient notification visualizations they managed with their real-time attention information and (2) other desirable application scenarios of A3InfoVis in their daily lives. Based on these insights, the chapter concludes with design guidelines for A3InfoVis, cover-

ing (1) personal reflection, (2) revealing attentional status to others, and (3) sharing personal attentional information among others.

5.1 Expected Technological Challenges of A3InfoVis

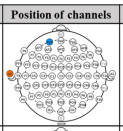
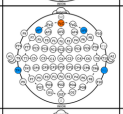
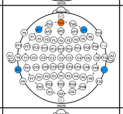
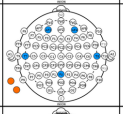
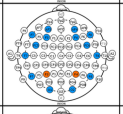
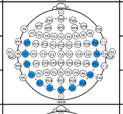
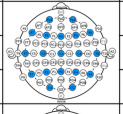
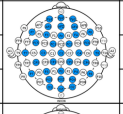
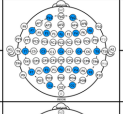
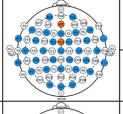
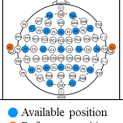
This section summarizes three technological challenges of A3InfoVis based on the literature review and survey on commercially available EEG devices. These correspond to the following questions that need to be addressed to make the users' real-time attention information available in their daily context: (1) Are previous heuristic indices and classifiers for attention state prediction applicable to users' everyday tasks?; (2) What are the minimum requirements of the end-user EEG devices deployed to personal contexts?; and (3) Should the data preparation for individualized attention state classifier be totally subject-specific?

Feasibility of Attention State Prediction for Everyday Tasks. A3InfoVis should be able to distinguish attention states in real-time while users focus on everyday tasks. Previous studies have developed machine learning (ML) classifiers to differentiate between simple cognitive tasks, such as mental addition and forward/backward-digit-span. These tasks assess specific cognitive abilities in highly controlled settings. In contrast, everyday tasks like watching YouTube videos involve more complex (e.g., audiovisual) stimuli that persist for longer durations and vary depending on the video content (e.g., relaxing landscapes and lecture videos). To the best of our knowledge, few studies have examined whether ML-based classifiers can effectively separate these types of tasks and how finely they can differentiate between *similar*

tasks (e.g., debate videos on different topics).

Limited Spatial Resolution of EEG Data Available for End-User Devices. The prediction performance of A3InfoVis should remain reliable even when using end-user EEG devices, which typically have fewer electrodes and cover narrower brain regions compared to professional devices, resulting in lower data quality. As shown in Figure 5.1, commercially available EEG devices exhibit a wide range of prices. Most devices priced under 900 USD have fewer than 8 electrodes and only cover a limited number of electrode spots from the 10-10 system, which is fully covered by medical devices. In addition to the price, usability issues arise when individuals find it challenging to put on and remove an EEG device with a large number of electrodes without professional assistance. Therefore, it is critical to determine if attention state prediction using EEG data with limited spatial coverage remains acceptable.

Burden of Data Acquisition for Building User-Specific Classifiers. The burden of data acquisition for A3InfoVis should be clarified if end-users wish to train a subject-specific attention state classifier that distinguishes between attention states of interest. End-users may find data collection and preparation procedures repetitive and tedious, posing a significant challenge. Furthermore, they are usually uncertain about how much data is sufficient to train a usable model. Given the individual variability in EEG data, it is essential to determine the extent to which models trained on other subjects' EEG data can be reused. If these pre-trained models can be repurposed, it would significantly reduce the burden of initial data acquisition for end-users. Specifically, comparing the performance of a general model, trained on all data,

Device name	Manufacturer	Channels	Position of channels	Price (USD)	Sensor conduction	Battery life	For medical use	For research use	For personal use
MindWave Mobile 2	NeuroSky	1		$109.99	Dry	8 hours	X	O	O
Muse 2 Headband	InteraXon	4		$249.99	Dry	5 hours	X	O	O (meditation and mindfulness training)
Muse S (Gen 2) Headband	InteraXon	4		$399.99	Dry	10 hours	X	O	O (meditation and mindfulness training)
INSIGHT 2.0	EMOTIV	5		$400-900	Semi-dry: polymer	20 hours	X	O	O
EPOC X	EMOTIV	14		$400-900	Semi-dry: saline	9 hours	X	O	O
OpenBCI EEG Headband Kit	OpenBCI	Up to 4 Up to 8 (including reference) Up to 16 (including reference)		$400-900 (including board) $900-1400 (including board) $2000+ (including board)	Dry		X	O	X
EEG Electrode Cap Kit	OpenBCI	Up to 4 Up to 8 (including reference) Up to 16 (including reference)		$900-1400 (including board) $1400-2000 (including board) $2000+ (including board)	Wet: gel		X	O	X
Ultracortex Mark IV EEG Headset	OpenBCI	Up to 4 Up to 8 (including reference) Up to 16 (including reference)		$900-1400 (including board) $1400-2000 (including board) $2000+ (including board)	Dry		X	O	X
Gelfree BCI Electrode Cap Kit	OpenBCI	Up to 8 (including reference) Up to 16 (including reference)		$1400-2000 (including board) $2000+ (including board)	Semi-dry: saline		X	O	X
EPOC FLEX	EMOTIV	32		$1400-2000 $2000+	Semi-dry: saline Wet: gel	9 hours	X	O	X
MINDD SCAN	Ybrain	19		Not for sale	Semi-dry: saline		O	O	X

Available position
Reference position

Accessed:
March 17, 2023

Figure 5.1: List of commercial EEG devices available for purchase as of 2023.

with that of a subject-specific model would provide insights into the challenge regarding data acquisition.

5.2 EEG Collection & Statistical Feature Analysis

This section covers the collection of EEG data and statistical analysis of its underlying patterns. First, we conducted a lab experiment that collected people's EEG data while they were watching video clips. We chose the task for two reasons: (1) it is one of the most typical daily tasks that can easily cover different attention states by varying the video sources; and (2) it does not require frequent keyboard and mouse usages, preventing severe movement artifacts. After collecting the data, we explored the frequency-domain features of the EEG data. We then tested the statistical significance of their variability according to the experimental factors. The analysis result indicated that frequency-domain features would be feasible for attention state classification while their activation pattern was partly inconsistent with an existing index for task engagement.

5.2.1 Study Design

We conducted a lab study at a local university to collect a dataset for analysis and ML model training. The whole procedure was approved by the institutional review board. We set up a dedicated study room in the university that consisted of essential facilities for daily desk work environment (e.g., desks, chairs, desktop PC, and input & output devices). We consistently managed temperature by air conditioning to prevent artifacts due to sweat. Participants sat at the desk and put on MINDD-SCAN [106], an EEG measurement device with 19-channel dry electrodes. We calibrated EEG signal based on its internal SQI (Signal Quality Index). During the measurement, we requested

participants to sustain a comfortable sitting posture as they usually do; we monitored their rapid movements (e.g., shaking heads or legs) that possibly trigger movement artifacts. If the movement induced a noticeable artifact to the EEG stream, we recorded the time interval to exclude the corresponding data from analysis. In case of the minor and brief movements, we reminded them of the request before starting the next measurement.

We measured EEG signals of five tasks from three distinct situations: while (1) taking a rest, (2) watching a landscape video clip, and (3) watching a debate video clip. These situations required "low", "medium", and "high" level of task engagement, respectively; regarding the third situation, we assigned three tasks of different topics to understand its influence to EEG data further. First, we measured EEG while participants took a rest for three minutes (we denote this task as *resting* in the following). They closed their eyes during the measurement to control confounding effects such as eye movement artifacts and unintended attentions to visual stimuli. Second, participants watched a four-minute video clip of a natural landscape (we denote this task as *video-landscape* in the following). We requested them to focus on the scenery with relaxation. Third, participants watched three debate video clips in a counterbalanced order. All the clips were official summary videos of a weekly TV debate show provided by one of the public broadcasting services in South Korea; they covered the three most controversial topics in the country at the time: (1) politics [22] (i.e., abolition of the Ministry of Gender Equality and Family), (2) economics [23] (i.e., overheated investment on the cryptocurrency), and (3) social welfare [24] (i.e., provision of the basic income); the length ranged from 5 to 7 minutes; each clip started with an introduction by the moderator and followed by arguments and rebuttals among two opposing groups of panelists. We named each task as *video-*

debate-p, *video-debate-e*, and *video-debate-s*, respectively. We asked participants about their opinion on the topic before and after watching every debate clip. Specifically, we informed them in advance that they would respond to (1) how interested they were in the topic and (2) which opinion they supported; these were to motivate participants to engage with the video because they might not be interested in some of the topics. The whole process took about an hour. We recorded participants' upper body during the measurement.

5.2.2 Participants

We employed opportunity sampling through online forums for students and faculty members of the local university. We recruited people who had not suffered from brain diseases and cognitive impairments. Twenty-seven people initially applied but we excluded the EEG data of the first two participants due to the unstable data transfer during their sessions. The final count of participants was 25 (13 males and 12 females); their ages ranged from 19 to 31 (μ=24.68, σ=3.51); two of them were left-handed; they were all undergraduate or graduate students. We rewarded them with 13,000 KRW (equivalent to 10.2 USD) for their complete participation.

5.2.3 Data Preprocessing

We preprocessed the collected EEG data using python open source packages, including MNE-python [35], a toolkit specialized in EEG data analysis. We excluded the first four seconds of the raw EEG data to remove abnormal oscillations and applied a high-pass filter ($f > 1.0Hz$), a low-pass filter ($f < 60.0Hz$), and a notch filter ($f = 50.0Hz$) to it. We then applied the ICA (Independent Component Analysis), one of the most famous BSS (Blind Source Separation) methods [45]; we excluded the independent components

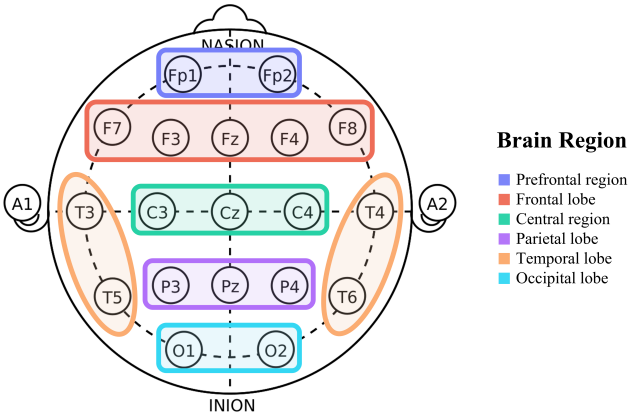


Figure 5.2: Correspondence between electrodes of the 10-20 system and the brain region variable.

suspected of being generated from artifacts (e.g., ocular and muscle movements). Next, we down-sampled the cleaned EEG data from 500Hz to 250Hz and split them into non-overlapping windows for one second. We computed the PSD (Power Spectral Density) for each window by Welch's method [105]. Finally, we got the time-variant PSD data of five frequency bands (i.e., delta, theta, alpha, beta, and gamma). We referred to video records to detect if participants made unexpected, rapid movements or dozed during the measurements.

5.2.4 Statistical Analysis

We first constructed a data table for statistical analysis of frequency-domain features of brain waves. Using the time-variant PSD data, we computed the relative power of the five frequency bands for each measurement. We also included a task engagement index ($\beta/(\alpha+\theta)$) [84], a brain wave index to interpret the user's attention to similar target tasks in HCI research [39, 98]. In addition, we introduced a new categorical variable, brain region (Figure 5.2), to investigate the spatial characteristic of brain waves. We catego-

rized the electrodes into the spatially corresponding brain lobes (i.e., frontal, parietal, temporal, and occipital regions). We then added two more regions (prefrontal and central) according to the international 10-20 system for more sophistication. The final table consisted of the following columns: (1) participant ID, (2) task type, (3) brain region, (4-8) relative power values of the five frequency bands, and (9) task engagement index.

Next, we conducted the two-way repeated measure ANOVA (Analysis Of Variance) to investigate the influence of the factors (task type and brain region) on the participants' relative power and task engagement value. We adjusted p values by the Greenhouse-Geisser correction method [36] to prevent the risk of violating sphericity. We then conducted a post hoc analysis. We applied multiple paired t-tests for every combination of tasks and brain regions; we adjusted p values for multiple comparisons by the Bonferroni method in a 95% confidence level.

5.2.5 Result

Relative Power of Frequency Bands

For all frequency bands, from delta to gamma, both of the factors, task type and brain region, had a statistically significant effect on relative power of frequency band. (Table 5.1). There were also significant interaction effects between task type and brain region for the lower frequency bands (delta, theta, and alpha).

According to the post-hoc analysis of task type, the relative power values of the resting task were significantly different from those of all the video-watching tasks (video-landscape & debate-p, e, and s) for all frequency bands. However, the patterns were contrasting between the alpha and the rest frequency bands. The mean relative alpha power was higher during the resting

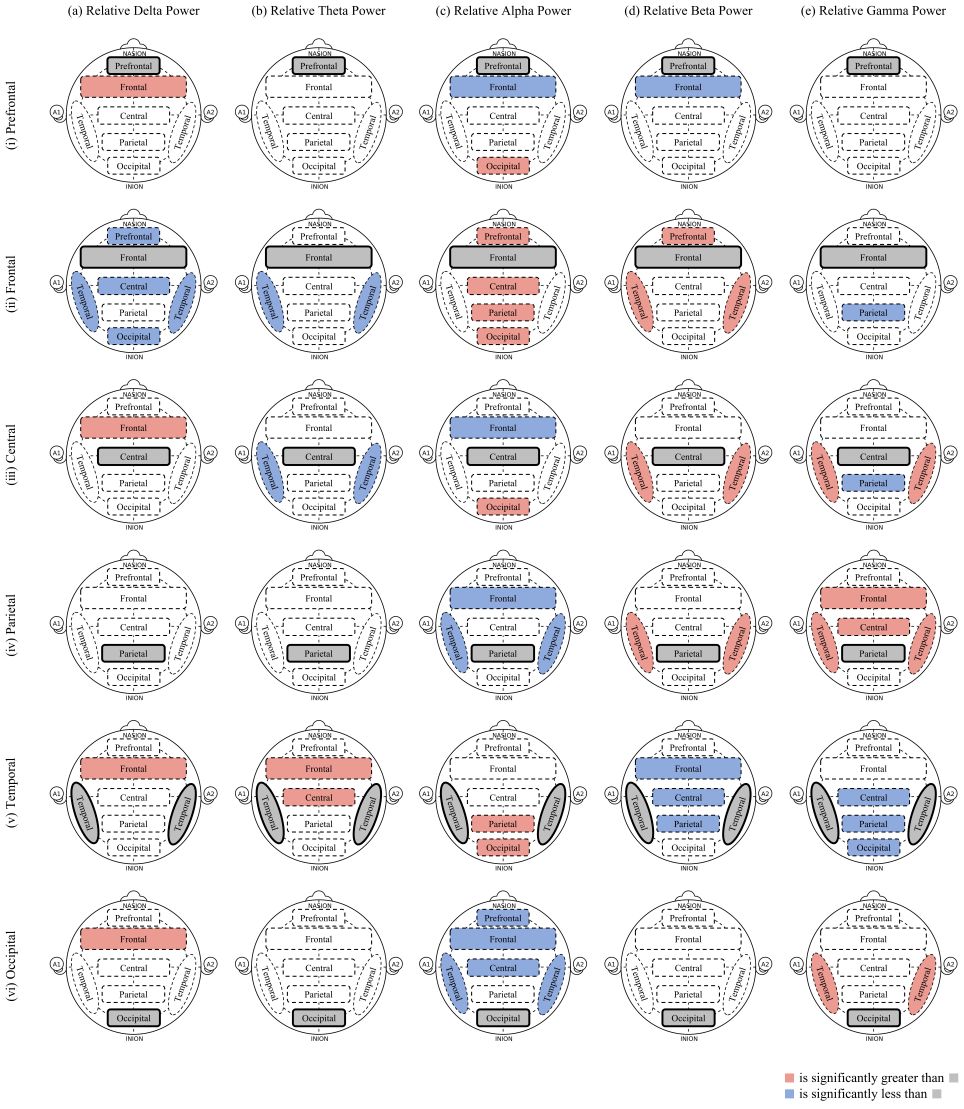


Figure 5.3: Results of the post-hoc analysis on the brain regional effect. The row corresponds to the brain regions: (i) prefrontal, (ii) frontal, (iii) central, (iv) parietal, (v) temporal, and (vi) occipital regions. The column corresponds to the relative power of frequency bands: (a) delta, (b) theta, (c) alpha, (d) beta, and (e) gamma bands. For example, a subfigure at (iii, c) represents all the statistically significant pairs that involve the central region (iii) regarding the relative alpha power (c). In the subfigure, we colored the regions using three colors: light grey, red, and blue. The average relative power of the red colored regions were significantly higher (e.g., occipital $>$ central), and that of the blue colored ones were significantly lower than that of the light grey region (e.g., frontal $<$ central).

band	factor	F-value	p-value	significance
Delta	*task*	$F_{(4,96)} = 39.42$	$p < 0.01$	✓
	brain region	$F_{(5,120)} = 6.4$	$p < 0.01$	✓
	interaction	$F_{(20,48)} = 9.13$	$p < 0.01$	✓
Theta	*task*	$F_{(4,96)} = 24.83$	$p < 0.01$	✓
	brain region	$F_{(5,120)} = 3.32$	$p < 0.05$	✓
	interaction	$F_{(20,48)} = 17.68$	$p < 0.01$	✓
Alpha	*task*	$F_{(4,96)} = 126.28$	$p < 0.01$	✓
	brain region	$F_{(5,120)} = 11.51$	$p < 0.01$	✓
	interaction	$F_{(20,48)} = 22.11$	$p < 0.01$	✓
Beta	*task*	$F_{(4,96)} = 14.22$	$p < 0.01$	✓
	brain region	$F_{(5,120)} = 8.91$	$p < 0.01$	✓
	interaction	$F_{(20,48)} = 2.72$	$p > 0.05$	
Gamma	*task*	$F_{(4,96)} = 15.36$	$p < 0.01$	✓
	brain region	$F_{(5,120)} = 7.73$	$p < 0.01$	✓
	interaction	$F_{(20,48)} = 2.15$	$p > 0.05$	

Table 5.1: Two-way RM ANOVA Results of Relative Power of Frequency Band

task than during the video-watching tasks. On the other hand, those of the four other frequency bands were lower during the resting task than during the video-watching tasks. Moreover, there was an additional significant difference between the video-landscape task and every video-debate task in the theta band. Hence, the mean relative theta power significantly increased in the order of resting, video-landscape, and video-debate tasks.

Figure 5.3 summarized the results of the post-hoc analysis on the brain region among frequency bands. We grouped the statistically significant size pairs for every brain region. We then delineated the reference brain region with solid line and light grey color, and other regions with dashed line. We

colored the region in red if its average relative power is larger than the reference region and colored in blue if the value is smaller.

Relative delta power: the average value of the frontal region was significantly higher than those of four other regions (prefrontal, central, temporal, and occipital; Figure 5.3(a)).

Relative theta power: the average value of the temporal region was significantly lower than those of the frontal and central regions (Figure 5.3(b)).

Relative alpha power: Figure 5.3(c) illustrated the pattern in which the back regions were likely to have higher values than the front regions. The average value of the frontal region was significantly lower than those of the prefrontal, central, parietal, and occipital region; all except the prefrontal region were its back regions (Figure 5.3(ii, c)). The average value of the occipital region, the backmost region, was significantly higher than those of other regions (pre-frontal, frontal, central, and temporal; Figure 5.3(vi, c)).

Relative beta power: Figure 5.3(d) indicated that the prefrontal and temporal regions had higher values. The average value of the prefrontal region was significantly higher than that of the frontal region (Figure 5.3(i,d)). The average value of the temporal region was significantly higher than the frontal, central, and parietal region (Figure 5.3(v,d)).

Relative gamma power: Figure 5.3(e) also delineated the pattern similar to the beta band. The average value of the parietal region was significantly lower than those of the frontal and central regions (Figure 5.3(iv, e)). The average value of the temporal region was significantly higher than those of the central, parietal, and occipital regions (Figure 5.3(v, e)).

index	factor	F-value	p-value	significance
Task Engagement	*task*	$F_{(4,96)} = 18.37$	$p < 0.01$	✓
	brain region	$F_{(5,120)} = 5.24$	$p < 0.05$	✓
	interaction	$F_{(20,48)} = 3.39$	$p < 0.05$	✓

Table 5.2: Two-way RM ANOVA Result of Task Engagement Index

The interaction effect was significant in the relatively low-frequency bands (delta, theta, and alpha). The post-hoc analysis revealed that during the resting task the relative power values of the alpha band tended to increase by moving from front to back regions. In contrast, those of the delta and theta bands tended to decrease in the same direction.

Task Engagement Index

Both of task type and brain region had a statistically significant effect on task engagement (Table 5.2). Also. there was a significant interaction between the effects of task type and brain region. The post-hoc analysis on the task type showed that the task engagement values of the resting task were significantly lower than those of all the video-watching tasks (landscape & debate-p, e, and s). The post-hoc analysis on the brain region showed that the values of the temporal region were significantly higher than those of the central, parietal, and occipital regions. The post-hoc analysis on the interaction revealed that task engagement values of the resting task were significantly lower than those of the video-watching tasks for all but one (occipital) brain regions.

5.2.6 Analytic Insights

Frequency Bands and Task Classification

Along the lines of the literature, our results also indicated that the frequency-domain features of brain waves could contribute to classifying the user's distinct attention states. The post-hoc analysis of the task type especially showed promising results for the real-time binary prediction. All the features extracted from one-second brainwaves could distinguish whether the user was taking a rest or engaging with a video. The theta and alpha bands deserve special emphasis among the five frequency bands. The relative alpha power showed a large difference with a contrasting pattern compared to the other frequency bands, so we thought it would be an essential factor for binary classification. The relative theta power was the only feature that could differentiate three task types (resting, video-landscape, and video-debates).

Necessity of Involving Diverse Brain Regions

Significant interaction effects indicated the necessity of incorporating features from diverse brain regions. They implied that an underlying interdependency between the task type and brain region affected the characterization of distinct patterns within the feature values. Because we need to infer the task type given the feature values of multiple channels from different brain regions, involving as many channels as input might contribute to better prediction.

Inconsistency with the Heuristic Task Engagement Index

Our result was partly inconsistent with the formula of the Pope's task engagement index, which $(\beta/(\alpha + \theta))$ is inversely proportional to the relative

power of the theta band. However, in our findings, the relative theta power increased as our participants performed a task requiring more engagement. Our result was consistent with the literature [10, 100] that the ratio of the theta and alpha is the indicator of cognitive performance. The major reason for this inconsistency might be the different target tasks. Pope and colleagues thought of the task of sustaining a moving symbol to be located within a rectangular boundary by controlling a joystick between two conditions. It was a very focused and highly controlled task that only dealt with visual information of the single mark. However, our video tasks involved more complex visual and auditory information. The difference might affect the contrasting activation pattern of the theta band. The inconsistency implied the necessity of building individualized ML models that can distinguish the person-specific cognitive states, complementing the limitation of the fixed heuristic indices.

5.3 Analysis on Technological Challenges

Based on the findings from our EEG analysis, we investigated the feasibility of ML-based attention state classification. In advance, we addressed the three technological challenges by investigating different experimental conditions. First, we analyzed the available sophistication level of attention state classification. We analyzed the performance of three classification models of different target classes: *2-class* (resting and video), *3-class* (resting, video-landscape, and video-debate), and *5-class* (resting and all the video-watching tasks). Next, we investigated the effect due to technical constraints of the EEG devices for end users by limiting the feature selection in terms of the brain region and the frequency band. Finally, we examined the effect of build-

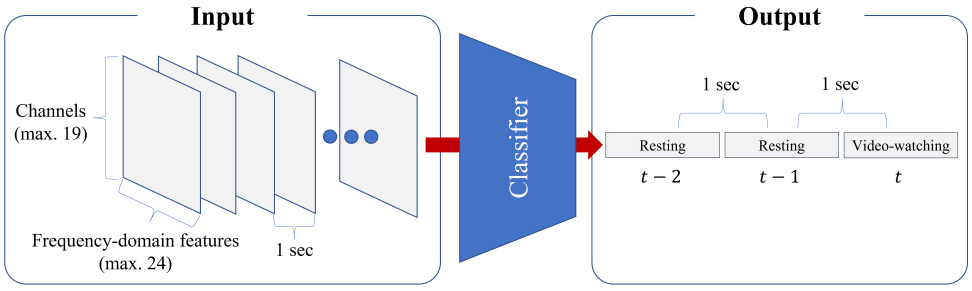


Figure 5.4: Pipeline of the real-time classification of the user's attention state. The classifier receives the brain wave feature matrix (19 channels * 24 features) every one second and predicts the corresponding attention state. For example, in this figure, the predicted attention state of the user at the current time t is the 'video-watching'. For more information on feature matrices, see the Feature Selection subsection of the text.

ing user-specific models trained with individual EEG data compared to the global model trained with the entire participants' dataset.

5.3.1 Method

Figure 5.4 delineates the pipeline of the real-time classification of the user's attention state. The classifier received the brain wave features of one second and predicted the corresponding attention state. We observed the performance in various conditions by changing (1) the number of prediction classes, (2) the inclusion of selective EEG channels or frequency bands, and (3) the scope of train data (i.e., entire dataset versus individual subset).

Specification of Attention States

We wanted to examine how finely the target model could distinguish the user's attention states. We trained the classifiers that predicted the 2-class (binary), 3-class, and 5-class attention states. We varied the number of classes by grouping the similar task types in multiple levels. The 2-class model predicted the resting or the engaged state; the latter included all the video-

watching tasks. The 3-class model predicted the resting or the two distinct video-watching states (i.e., video-landscape and video-debates); each state involved the tasks of different engagement levels. The 5-class model classified all the tasks into distinct states (i.e., resting, video-landscape, video-debate-p, video-debate-e, and video-debate-s).

Model Selection

Considering the size of our dataset, we explored classical ML methods rather than deep learning models. We thought our dataset of 125 measurements from 25 participants would be insufficient to train deep learning models, which would lead to underfitting. Among ML methods, we chose SVM and the ensemble of decision trees. SVM is known to be data-efficient and robust against overfitting while the ensemble of decision trees shows a higher performance despite a higher risk of overfitting. For SVM, we utilized C-SVC (C-Support Vector Classification) model from libSVM [12], one of the most widely used SVM libraries, and used a radial basis kernel function to better capture a non-linear pattern in our data. For the ensemble model, we utilized XGBoost (eXtreme Gradient Boosting) [13]. It is known for one of the most effective classification models for tabular data, even including deep learning methods [4].

Feature Selection

We computed the total 456 features (19 channels * 24 features per channel) for every one-second window of the preprocessed EEG data without overlap. For each EEG channel, we chose the 24 frequency-domain features referring to both the literature [3, 10, 15, 30, 41] and the result of our analysis. Regarding the ratio of two frequency bands, we started with the entire com-

binations and excluded ones that had a negligible impact on performance. We eventually got following four features.

- Absolute power of five frequency bands (N=5)
- Relative power of five frequency bands (N=5)
- Peak power within five frequency bands (N=5)
- Peak frequency within five frequency bands (N=5)
- Ratio between different frequency bands (i.e., delta-to-theta, theta-to-alpha, alpha-to-beta, and beta-to-gamma) (N=4)

We analyzed the effect of the feature selection from two perspectives: spatial and frequency-domain characteristics. First, we compared the models trained with the selective channels that belonged to an identical brain region. It was to understand how the spatial characteristics of brain waves influence the classification of user's attention states. We investigated nine regions: six individual regions (i.e., prefrontal, frontal, central, temporal, parietal, and occipital regions in Figure 5.2), two (left and right) hemispheres, and the entire region.

Next, we compared the models trained with the three different sets of channel features: (1) entire features (n=24), (2) task engagement-related features (n=13), and (3) the theta-specific features (n=8). The second set included the features of the three bands (theta, alpha, and beta) related to the task engagement index; it also involved the four ratio features (delta-to-theta, theta-to-alpha, alpha-to-beta, and beta-to-gamma). The third set included the features of the theta band and the four ratio features. We covered this feature set because the theta might contribute to a more sophisticated classification according to the results of our statistical analysis (Section 5.2.6); it

was the only frequency band whose relative power showed the significantly different statistics among the 3-class attention states.

Training, Validation, and Testing

As a result of feature computation, we obtained about 1,320 labeled data points for each participant (180 points for the resting state, 240 for the video-landscape state, and 300 for the video-debate-p, e, and s states, respectively). We split the data points of each task measurement into train and test data in a ratio of 4:1. To resolve the potential risk of the class imbalance, we applied *the balanced option* to SVM, where the model weights the train data inversely proportional to its class frequency. We posed two different training approaches to figure out the strengths of personal models: first, we built 25 personal models for individual participants' data; second, we trained a single general model using the entire train dataset. During the training phase, we conducted the four-fold cross-validation to monitor if there existed an overfitting problem. After the validation, we compared the average test performances of the personal models and the general model. To better understand the aspects of model errors, we calculated precision, recall, and F_1-score to assess a class-level performance. We then utilized the weighted average of these class metrics to indicate the model performance.

5.3.2 Result

Available Sophistication Level of Classification

Figure 5.5 summarized the average performance of the personal models according to the number of prediction classes, trained with the features selected by the brain region. The first row of Figure 5.5 showed the model performance, trained with the complete feature set. The performance tended to

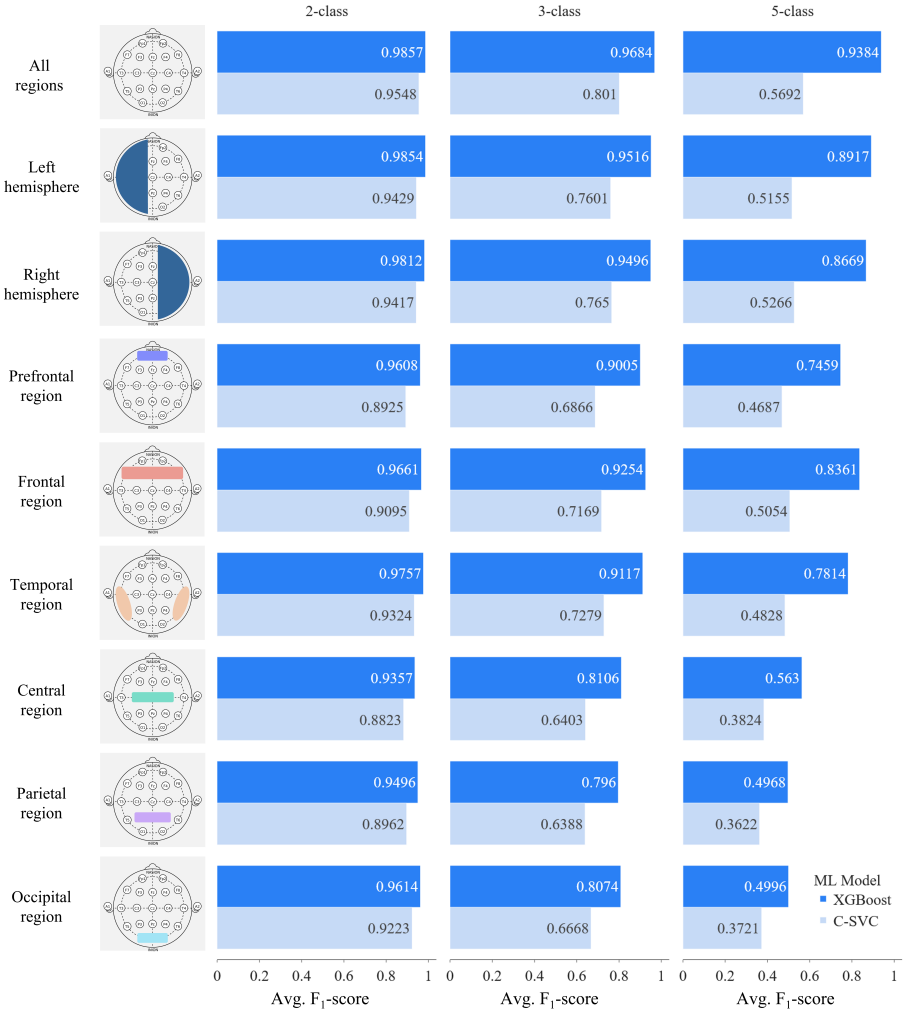


Figure 5.5: Average performance of personal models trained with channels of the selected regions. For classifiers, XGBoost (solid blue) always outperformed C-SVC (light blue). Among the region-specific models, the prefrontal, frontal, and temporal region models consistently performed better than others.

decrease as the number of prediction classes increased. The ensemble tree model, XGBoost, was reliable for all the classification problems, while the performance of C-SVC dropped for multi-class prediction. Detailed analysis indicated that in the 3-class classification, C-SVC was inefficient (F_1=0.51) for separating the user's landscape-engaged state from the other two states. In the 5-class classification, C-SVC also performed unstably except for the resting state prediction (F_1=0.83); the F_1-scores of the other four classes ranged from 0.48 to 0.54.

Brain Regional Effects

Figure 5.5 summarized the average performance of the personal models, trained with the features selected based on the brain region. XGBoost performed better than C-SVC in every condition. The models that covered the features of overall brain regions (the entire region, and the left and right hemispheres) showed better performances than the six other models that only covered region-specific features. The performance difference became greater as the number of prediction classes increased. Among the region-specific models, the prefrontal, frontal, and temporal region models consistently performed better than others. The occipital region model was comparable to the three models in the case of the binary classification.

Frequency Band Effects

Table 5.3 summarized the average performance of the personal models, trained with the features selected based on the frequency band. XGBoost consistently performed better than C-SVC. It achieved the best performance when trained with the entire features (n=456); the performance dropped as the number of features decreased. On the other hand, C-SVC showed a different

ML model	classifier	feature set	precision	recall	F_1
XGBoost	2-class	entire set	0.9860	0.9861	0.9857
		θ, α, β	0.9857	0.9857	0.9853
		θ	0.9764	0.9769	0.9759
	3-class	entire set	0.9692	0.9691	0.9684
		θ, α, β	0.9617	0.9615	0.9605
		θ	0.9341	0.9344	0.9310
	5-class	entire set	0.9393	0.9383	0.9384
		θ, α, β	0.9168	0.9158	0.9157
		θ	0.8395	0.8380	0.8372
C-SVC	2-class	entire set	0.9583	0.9532	0.9548
		θ, α, β	0.9590	0.9541	0.9558
		θ	0.9598	0.9531	0.9552
	3-class	entire set	0.8281	0.7884	0.8010
		θ, α, β	0.8334	0.7953	0.8075
		θ	0.8339	0.7912	0.8045
	5-class	entire set	0.5781	0.5696	0.5692
		θ, α, β	0.5910	0.5825	0.5809
		θ	0.6100	0.5984	0.5962

Table 5.3: Average Performance Metrics of Personal Models by Selected Features

pattern; its performance did not monotonically decrease as the size of the feature set did. Both the task engagement-related features (θ, α, β, n=247) and the theta-specific features (θ, n=152) sustained the model performance comparable to the entire feature set even with a fewer number of features. Moreover, the theta-specific features showed the best performance among the 5-class C-SVC models while the overall performance of the 5-class C-SVC was not good in general.

ML model	classifier	$F_1(general)$	$F_1(p_{avg})$	$F_1(p_{max})$	$F_1(p_{min})$
XGBoost	*2-class*	0.9744	0.9857	1.0000	0.9690
	3-class	0.9192	0.9684	0.9966	0.9294
	5-class	0.8186	0.9384	0.9834	0.8188
C-SVC	*2-class*	0.9338	0.9548	0.9902	0.8624
	3-class	0.6903	0.8010	0.9436	0.6793
	5-class	0.3990	0.5692	0.6880	0.4103

Table 5.4: Comparison of Performance Metric between Personal Models and General Model.

General versus Subject-specific Models

Table 5.4 summarized the performance of the subject-specific models and the general model, trained with the complete feature set, by the number of prediction classes. The performance of the general model was lower than the average performance of the subject-specific models. The gap of F_1-score (2-class=0.02, 3-class=0.05, 5-class=0.12) broadened as the classification became more complex. Specifically, most personal XGBoost models (23 out of 25 2-class models and all the 3-class and 5-class models) showed a higher F_1-score than the general model. The performance of the general XGBoost was good for every classification model, especially as high as 0.9744 for the 2-class classification. In case of C-SVC, F_1-scores of sixteen 2-class, thirteen 3-class, and eleven 5-class models exceeded that of the general model. Except for the 2-class classification, the general model showed low performance, being unsuitable to use for classification.

5.3.3 Challenges Revisited: Possibilities and Limitations

Based on the two analysis results, we revisited the three challenges regarding the technical feasibility of A3InfoVis and summarized the implications.

Feasibility of Real-time Attention State Classification for Casual Tasks

We found it feasible to predict the user's real-time attention states for everyday tasks, provided with the 19-channel EEG device. Both ML models effectively predicted the binary attention states, whether the user is resting or attentive to the task. Predicting a more sophisticated attention state (3-class or 5-class) was also reliable with the ensemble tree model; our 3-class model could successfully differentiate three levels of task engagement; our 5-class model showed the possibility of separating the attention states further while people performed similar casual tasks (i.e., watching debate videos of different topics).

Trade-off between Performance and Usability of EEG technology for End-user's Context

We point out that BCI researchers and practitioners should consider trade-offs between the usability and performance of end-user EEG devices. We investigated the model performance regarding the coverage of the brain region and observed its decrease among the models trained with the limited EEG signals from the single brain region. In this respect, researchers or practitioners should deploy an EEG device that covers as diverse brain regions as possible. However, most multi-channel ($n \geq 19$) EEG devices available in hospitals or laboratories are not usable for end users. They allow sophisticated EEG measurements by covering the whole brain regions with tens of or even hundreds of electrodes but require the help of experts even when putting on and off these devices; the cost of purchase and maintenance is also hardly affordable. On the other hand, casual BCI devices utilize a small number of electrodes for the restricted brain regions. Although they are more

usable and accessible for ordinary people, they would output a lower performance in predicting attention states.

Our result provides more detailed insights about deploying an adequate level of EEG technology into the end-user's context regarding the trade-offs. The binary prediction about the user's attention remained reliable despite the partial coverage of the brain regions. BCI researchers and practitioners would use any of the six (i.e., prefrontal, frontal, temporal, central, parietal, and occipital) brain regions for simple attention state classification. The more sophisticated models (3-class and 5-class models) required more electrodes covering broader brain regions to sustain reliable performance, but we also found three crucial brain regions (i.e., prefrontal, frontal, and temporal regions) that would contribute to better prediction performance. We recommend that researchers and practitioners prepare an EEG device (i.e., commercial product or self-produced prototype) that prioritizes these regions.

Possibility of Managing End-user's Burden for Initial Data Acquisition

Our result implies that end users can construct individual attention state classifiers using an EEG device with a low burden of data preparation and initial model training. We achieved a reliable classification performance with the few-minute EEG signals of the attention states, which ranged from three to six minutes. Our model enabled people to compose multiple target attention states from the daily tasks of interest (e.g., watching of video lecture with engagement or in a daze) and collect the signals for a short duration. As we found in General versus Subject-specific Models, they can also start the usage with the binary classifier (i.e., engaged or not) pre-trained with the other users' data and improve it further by fine-tuning or retraining the model with the new data collected during the usage.

5.4 Analysis on Design Opportunities

To understand the end-users' interests and needs for utilizing their real-time attention information in the design of ambient information visualization, we conducted a semi-structured interview. The interview analysis results indicated that end-users were willing to leverage their real-time attention in two ways: (1) to fine-tune the distraction cost induced by the given ambient visualization and (2) to utilize the information as an improved tool for self-reflection and communication with others.

5.4.1 Participants

To gather responses from end-users experienced in authoring and managing ambient information visualizations in their personal context, we interviewed people who had participated in the deployment study of DataHalo (Chapter 4). DataHalo [37] is a customizable ambient information visualization system targeted at smartphone notifications, which offers smartphone users flexible customization options to design personalized extended notification badges in terms of information, delivery, and visual representation. DataHalo was deployed on their Android smartphones as an open-source launcher for three weeks. Among twelve participants who completed the deployment study, seven of them (P02, P03, P05, P06, P10, P11, P12; two males and five females) accepted the interview. They were senior undergraduates and graduate students in their 20s and 30s. They had at least three weeks of experience using DataHalo and had authored more than five custom ambient notification visualizations during their usage; P02 and P03 used the system for over six weeks.

5.4.2 Study Design

We met each participant on Zoom. At the beginning of the interview, an experimenter briefly reminded the participants of the ambient notification visualizations they had authored and managed during the deployment. Next, the experimenter provided a short introduction to BCI technologies and application examples for promoting better ideation. This included the definition of EEG signals, form factors of end-user EEG devices, and example BCI scenarios of practitioners and researchers in which the real-time user's attention information was visualized (e.g., EngageMeter [39]) or used as a factor to control the system (e.g., Ena [2]). After the introduction, participants were asked two sets of questions. In the first set, the experimenter inquired about how participants would further customize their ambient notification visualizations authored for DataHalo using their real-time attention information. Follow-up questions addressed how their attentional information would enhance the personalization of the custom notification badges they had managed. The second set covered broader application scenarios not limited to smartphone notifications. The experimenter asked participants to think of specific contexts where they would utilize attention-adaptive ambient visualization. The interview took 40 minutes in total, and participants who completed the interview received 15,000 KRW ($\tilde{1}$1.5 USD) as a reward.

5.4.3 Analysis

We transcribed the participants' responses from the semi-structured interview. We first grouped the transcriptions into two situations where participants were interested in utilizing real-time attention information as (1) a factor to further minimize the distraction caused by ambient information visualization, or (2) a new source of information to be visualized in their

personal context. For the first group, we further categorized the transcriptions into attention-adaptive personalization methods concerning (1) information (e.g., filtering), (2) delivery (e.g., duration and noticeability of visual feedback), and (3) visual representation (e.g., visual encoding) of ambient information visualization. For the second group, we separated the usage scenarios into three situations: (1) where attention information is used only for themselves, (2) where their attentional status is exposed to an observer, and (3) where a group of people shares attention information among others.

5.4.4 Result

Possibilities of Fine-tuning User Distraction Induced by Ambient InfoVis

All participants agreed that their real-time attention could be an essential factor to improve ambient information visualization in terms of managing user distraction more delicately. Utilizing the customization features of DataHalo, participants could create ambient notification visualizations that conveyed personally relevant information at a manageable level of distraction. However, they also noted that this "acceptable level of distraction" fluctuated according to shifts in the personal context, which affects their current attentional capacity.

The participants expressed a diverse desire to leverage their real-time attention to fine-tune the ambient visualizations they had created. Especially, they had different opinions on which notifications should be shown and which ones could be omitted. All of them agreed that these standards should be personalized, as they created individualized notification categories within apps to reflect their information of interest while using DataHalo. P02 hoped to alleviate the burden of building the schema from scratch, so she thought of providing data-driven suggestions, which leverage EEG signals and usage

logs, to users. Specifically, she expected that attending notifications valuable for her might induce distinct patterns in brain waves (e.g., high engagement) or in usage behaviors (e.g., faster attendance). By collecting these patterns, the personalization interface can recommend the pass list of notifications during the working context.

All participants identified *working* as the most interesting context, where they had to focus on the task of the highest priority such as studying for an exam, working an internship from home, or doing academic research. For example, P06 suggested an attention-adaptive *Do Not Disturb* approach during the *working* context, which suppresses most notifications if he is highly engaged in the primary task during the focused research time. Likewise, most participants (except P02 and P03) thought more notifications should be filtered out during the working context to alleviate distraction. P02 and P03 worried that they might miss some important notifications as a result of filtering; they suggested other methods such as ordering.

Users' real-time attentional information was valuable but not a dominant factor in determining participants' personalization strategies. Participants considered multiple factors (i.e., delivered information, their current context, and attentional capabilities) to specify their personalization scenarios. They repeatedly mentioned that if any of these factors changed, their personalization strategy would also change. Personalization scenarios among participants were very different, even if these three factors were the same. Nevertheless, there were some commonalities among participants. For example, they usually split their notifications into "attention-invariant" and "attention-dependent" information. Attention-invariant information always took priority over participants' current attentional information.

Participants tried to process real-time attention in different ways. P02 specifically suggested measuring *residual attentional capacity*, similar to peripheral awareness ([2]). In her mind, having time to check notifications is a little different from the low engagement level, which indicates that she has been not fully engaged in her primary task. People frequently explained the state of *distracted*, but the meaning was different. Some wanted to explain the mental state of being overwhelmed while others described the situation they were dizzy.

Opportunities of Ambient Attention Visualization for Self-Reflection and Communication

All participants, consisting of senior undergraduate and graduate students, shared a common interest in accessing and utilizing real-time attention data to self-reflect on their study productivity. While preparing for their qualification exams, P10 and P11 regularly engaged in extensive study sessions at libraries or in their personal spaces. They speculated that their productivity during these sessions might be inconsistent, but no objective means of verification were available. They wanted to undertake simpler tasks or take breaks when provided with visual feedback on their fluctuating engagement levels. In the long run, they hoped to monitor their engagement levels to determine periods of optimal productivity and adjust their study plans accordingly. Similarly, P05, a graduate student, posited that real-time observation of engagement levels could bolster motivation. Given an ambient information visualization of a candle flame that burns brightly when highly engaged and flickers when distracted, P06 believed that facing the flickering candle would encourage him to renew engagement with his research.

Participants expressed varied opinions regarding the continuous observation of their attentional information by others. Some considered positive applications within controlled and specific contexts, such as during business hours or in the library. P10 believed that having someone monitor her engagement level would increase her focus during study sessions compared to self-reflection alone. However, the majority of participants found the presence of an observer highly burdensome due to the perception of surveillance, especially when a hierarchical relationship exists. They believed that their real-time attentional information should remain inaccessible to observers, and individuals should maintain control over the granularity of the information and the frequency of its updates.

On the other hand, participants responded positively to the idea of exchanging real-time attention information with peers in scenarios where they gathered for shared objectives, like study groups. They found real-time sharing particularly helpful in this condition, as it was crucial to the purpose of the gathering. Nevertheless, P03 and P10 were wary of privacy intrusions when sharing study spaces through webcams, and thus mentioned that displaying real-time focus information using ambient visualization would be a preferable solution.

5.5 Design Implications

Based on the two analysis results, we revisited the three challenges regarding the technical feasibility of A3InfoVis and discussed what would be essential in designing ambient visual feedback that leverages the users' attention information.

5.5.1 Visualization of Predicted Attention State with Credibility

We recommend that the designers deliver the predicted attention state and its credibility for A3InfoVis. Although we achieved a competitive level of performance for a single prediction, even intermittent mispredictions would accumulate to a negligible extent due to the short interval (e.g., one second) of real-time prediction. However, presenting only the prediction results cannot give end-users accurate attention information. For example, an alternation between the resting and engaged state in the sequence of the 2-class predictions might reflect the user's unstable task engagement but could be a false alarm simultaneously. If the latter, the corresponding change in A3InfoVis leads to a misunderstanding of one's attentional state and would eventually discredit the whole prediction results. Conveying the prediction result with its credibility information would be an effective solution to differentiate these situations.

We exemplify two information sources that can deliver meaningful insights about credibility. First, the prediction probability, provided by the ML classification model, would be the possible candidate for the uncertainty information. Because many ML classification models output one with the highest prediction probability among entire classes, it could represent the model-side confidence. For example, the prediction probability similar to random guess uncovers that the current model might not differentiate distinct attention states. The designers could also suggest other heuristic indices by monitoring the sequence of prediction results. For instance, an engaged state amid multiple resting states is less likely to be a credible prediction; it might be due to an unexpected occurrence of artifacts. Likewise, we could define a simple index called *consistency*, the ratio between the count of the designated predicted class and that of total predictions for a given time window. It would

be low when the model initially outputs a new class and gradually increase if the same prediction results repeat. These two numeric variables cover the different aspects of credibility; the prediction probability addresses the uncertainty of a single prediction, while the consistency metric is relevant to the time. Hence, delivering them together would give users a deeper insight into the credibility of prediction.

Finally, we suggest a way to manage the increased complexity of visual information due to the additional credibility information. It has been an essential design criterion [85] of ambient information visualization to determine an adequate amount of information delivery that fits the user's capacity. Because involving the multi-faceted credibility would significantly increase the amount of information, users might feel overwhelmed or visually noisy about the visualizations, which harms the virtue of the ambient information system. Designers could refer to the methods suggested by the InfoVis community that aimed at better utilizing a limited budget of human perception and cognition. For example, they could utilize Value–Suppressing Uncertainty Palette (VSUP) [20], a tree-structured bivariate color map, to encode the prediction probability to hue and the consistency to saturation. It restricts the range of expressible hue when the consistency value is low so that an abrupt change in the predicted result (as well as the prediction probability) does not affect the hue immediately. As the new predicted result persists, its hue becomes more distinct, and the color becomes saturated simultaneously, making the credibility information easier to perceive.

5.5.2 Adequate Application Scenarios of A3InfoVis

One of the essential characteristics of ambient information visualization is its longitudinal interaction with people in a casual context. We summarized

two crucial constraints to designing a satisfying interaction scenario. First, we recommend that the designers present a focused usage context where the users naturally perform a single target task. In a more complex usage context of many simultaneous stimuli, EEG data alone is not enough to explain the user's attention states. Users arrange their attentional resources among multiple tasks during usage, but the binary prediction cannot specify to which tasks they pay attention. Additional contextual information (e.g., gaze, computer usage logs, or facial expressions) is required to correctly trace the temporal transitions of the user's target task in the context.

Moreover, the scenario should naturally limit the user's rapid movements. The tasks involving dynamic physical actions (e.g., playing sports) lead to intense artifacts; they would degrade the overall prediction during the usage. Many desk work scenarios could satisfy the constraint. Except for some habitual actions (e.g., shaking one's leg) or accidents (e.g., going to the toilet), people would not continuously make strenuous movements during these scenarios. We also found that during our EEG measurement, casual arm movements related to desk work (e.g., typing keyboards or pointing the mouse) did not cause significant muscle artifacts.

Based on the above constraints, we exemplify a longitudinal interaction scenario of an online group study where A3InfoVis could help people use their attention information. The online group study constitutes a longitudinal, focused, and stationary deskwork scenario. It has become more popular since the outbreak of COVID-19. People use video conference apps, streaming services, and specialized online platforms to form groups and study together. Participants are highly motivated to achieve their study goals. However, the online group study environment is limited to sharing the contextual information (e.g., the presence of other study mates and the academic atmo-

sphere) that further motivates them. Visualizing the attention states makes the user's internal state observable in the environment. Because their attention information directly relates to the attainment of the goal, individual users can monitor the visualization for self-motivation during the session. In addition, sharing the visualization among participants can contribute to the mood creation for the study; they can be further motivated by the visualization of other passionate members' attention information.

5.5.3 End-user Support for Design of Personalized A3InfoVis

We finally summarized the requirements of the interactive system that could support end-users to design A3InfoVis adequate for individual usage contexts. First, supporting the EEG data collection and labeling process would be essential for building a personalized A3InfoVis. The designers should suggest the interface and interaction techniques that help laypeople complete the data preparation. For example, labeling many data points might be tedious work that demotivates users. The designers can consider the semi-supervised learning method. After the user labels a part of the data, the system could propagate the label information to the rest based on the similarity. Integrated with the visual inspection of label prediction results, users can further lessen the cost of validating the predicted labels.

Second, the system should allow users to customize the models based on personalized attention states. The target attention states that the users want to classify would differ according to their goal and context (e.g., target task, the number of classes). To support these needs, end users need more control over the model building beyond data preparation. For example, they might need to label the data with new classes. They also need to evaluate the performance of their model and improve it by inspecting the bottleneck. Recently,

MLOps (Machine-Learning Operations) [99] services that tackle the similar needs of people in the industry have come to fame. However, the designers should also tackle the complexity due to the high customizability. It would be desirable to hide the model details (e.g., hyperparameters) and automate the settings to handle customization fatigue. In addition, the system should support a more comprehensive explanation and suggestions for model improvement. Applying XAI (eXplainable AI) techniques such as LIME [89] and SHAP [62] might further contribute to the more intuitive explanation of the model.

Third, end users should have more control over the visual information delivered by A3InfoVis. As the attention information is innately personal, its embodiment in the environment would trigger negative side effects; laypeople might be afraid of the situation where the visualization might serve as a surveillance method. They must control the information conveyance to resolve the risk. Users can achieve an adequate level of privacy by customizing the abstraction level of information. For example, they could control the temporal resolution to be broader (e.g., from seconds to minutes) or make the information visible for the designated interval. Users also could edit the visibility of the predicted attention states by visualizing only the selected ones. The most effective control would vary according to the longitudinal changes in the user's context (e.g., time, location, events). Therefore, the interface and interaction methods to customize the abstraction strategies should be easily accessible and continuously maneuverable.

Chapter 6

Discussion

This chapter discusses lessons learned from the process of seeking answers to the research questions, the design challenges encountered in this research, and future opportunities for improving the proposed systems. The limitations of the studies are also discussed.

6.1 Lessons Learned

This dissertation endeavored to answer three research questions. The first research question focused on **designing the flexible customization capabilities that expand the informational benefits with manageable distractions for diverse personal information**. As an answer to the question, DataHalo was designed and implemented. To support end-users who have struggled with smartphone notifications, DataHalo integrated a customizable ambient information visualization system into an existing Android notification system. Beyond blocking or suspending all the notifications, DataHalo provided end-users with flexible capabilities to create app notification visualizations of personalized ambience (i.e., information, delivery method, and visual representation) called halos. A remarkable feature of DataHalo was that

it could be applied to general smartphone applications. Thus, every smartphone user could create halo visualizations for the interested apps and utilize their smartphone home screen as a personalized ambient display. The usability study results indicated that customization capabilities provided by DataHalo were useful for better notification management and had the potential to rediscover the benefit of notification information. At the same time, the study pointed out the necessity of alleviating the possible fatigue of customization from scratch.

The second research question focused on **exploring the patterns among the end-users' design strategies to personalize their ambience in their daily context**. To answer the question, an improved version of DataHalo was deployed to Android smartphone users' daily context. Participants agreed on the improved satisfaction of daily notification management and expected the DataHalo features to be accessible in their smartphone environment soon. The field deployment study results revealed a number of interesting findings. As expected, participants managed halo visualizations for diverse apps and their customization strategies to construct halo visualization were highly individualized. Moreover, participants persistently edited their halo visualizations or sometimes removed them, which indicated that personalized ambience is not an immutable state but changes according to various reasons relevant to users themselves or their surroundings. Furthermore, the study revealed the gap between the system-provided notification schema (i.e. Android notification channels) and the user-created ones (i.e., virtual categories). It implies that future notification systems need to empower end-users to restructure the way to categorize notifications as they wish flexibly.

The third research question focused on **examining the limitations and opportunities of attention-adaptive personalization of ambience** . First, an

empirical study was conducted to investigate three technological challenges related to real-time attention state classification in end-users' individualized, everyday context. The study, which involved constructing machine learning classifiers for casual video-watching tasks, revealed that end-users would be viable to build a binary or ternary attention state classifier using few-channel EEG devices, without a heavy burden of initial data collection. Next, semi-structured interviews with end-users highlighted their desire of utilizing real-time attentional information to fine-tune their distraction cost further or to create new ambient attention visualization for self-reflection and non-obtrusive communication. On these bases, design implications of attention-adaptive ambient information visualization covered issues of credible visual representation, possible usage scenarios, and personalization supports.

6.2 Design Implications for PAAIV

Synthesizing the valuable lessons learned from the two works, DataHalo and A3InfoVis, this section presents the structure of the conceptual framework, PAAIV. The goal of suggesting the conceptual framework was to design a better ambient information visualization system that provides end-users with a personalized ambience, which (1) provides the prioritized information that users are interested in with suitable complexity, (2) delivers the information for the opportune time span without user distraction, and (3) visualizes the information in the way that persistently motivates users to engage with it during the time span. PAAIV aimed to cover personalization according to user's design perspectives and real-time attentional demands. In the following, the section presents the related design implications of each personalization approach and then discusses their connection.

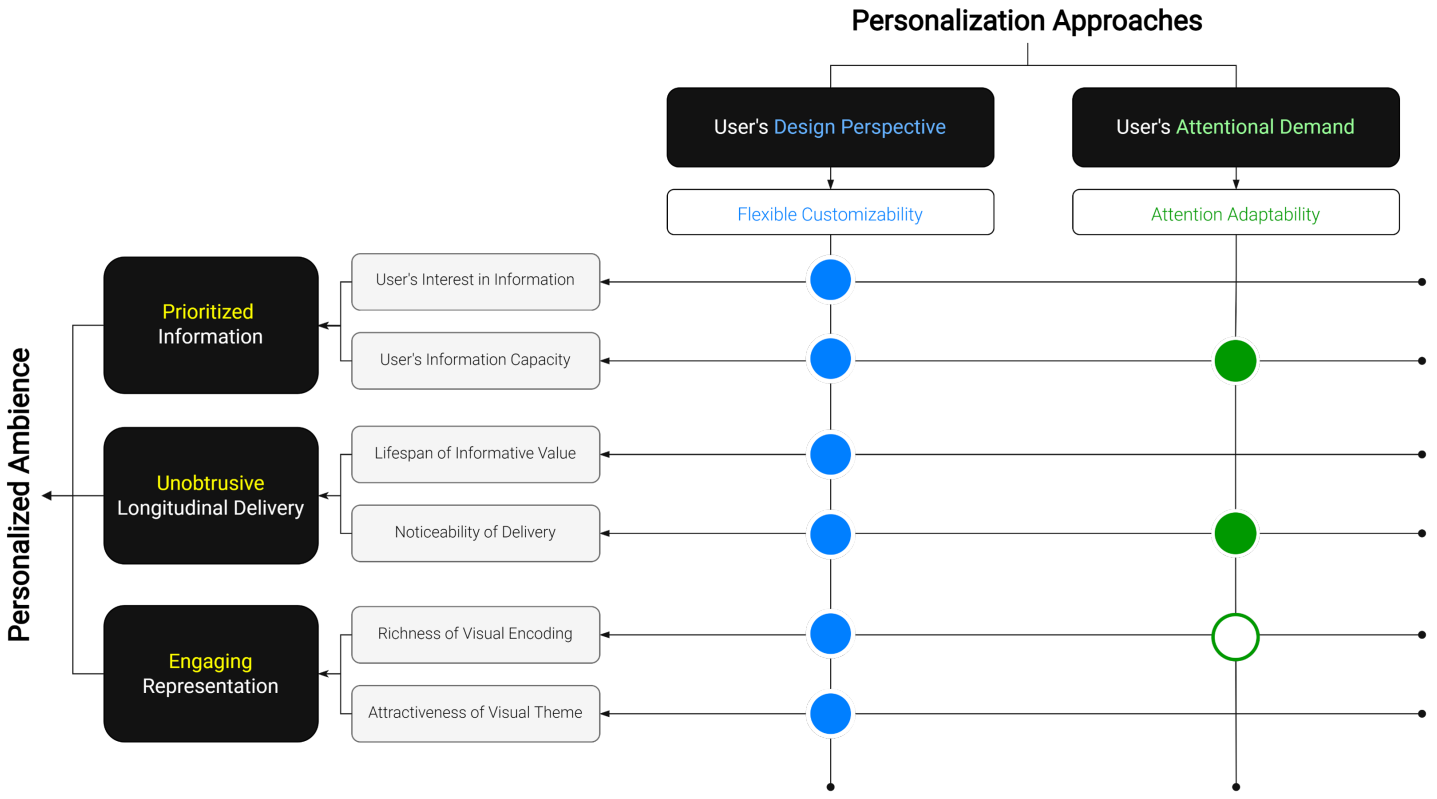


Figure 6.1: PAAIV Framework and its two personalization approaches.

6.2.1 Personalization According to User's Design Perspectives

DataHalo provided end-users with personalization capabilities covering the entire process of design of ambient information visualization for smartphone notifications, from data to resultant visualization (Figure 6.1; blue colored). Specifically, it supported customizable components for information prioritization by keyword-based filtering and categorization, longitudinal delivery based on the urgency and persistency of the information, and visual encoding and theming of the graphical mark.

Designers should notice that users' personalization needs are diverse in multiple aspects, as individual users are interested in different personalization capabilities and have varying expectations for each component's customizability. Some users did not utilize some components for most of their usage. Thus, providing reasonable default settings for untouched components is an essential usability issue. It should not be a vacant setting but one that covers the most familiar usage. For example, DataHalo provided initial settings resembling conventional user–notification interaction with notification badges, rather than rendering a "null halo" of which graphical marks are all hidden.

Moreover, new customization capabilities that have not been granted to end-users might be unfamiliar to them. As DataHalo improved based on user feedback, providing diverse presets can help users learn their personalization capabilities easily and reduce customization fatigue. Additionally, the interface for customizable components should be easily accessible, as DataHalo users frequently adjusted their settings in response to self-induced or environmental changes.

Related Design Implications

- Prepare Default Settings that Resemble the Most Familiar Usage
- Provide Diverse Example Usages for Better Understanding and Alleviation of Customization Fatigue
- Make Component Interfaces Easily Accessible to Adjust the Settings over Time

6.2.2 Personalization According to User's Real-time Attentional Demands

As the empirical study on A3InfoVis indicated, customizable components can be further extended to incorporate attention-adaptive features. The results demonstrated the feasibility of predicting binary attention states in real-time using consumer EEG devices of a limited number of channels. Users' predicted attention state can be introduced into the visual encoding component as a new information property. For example, users can map resting and engaged states into one or more visual channels such as color, size, and animation of the graphical mark. Combined with the user-imposed importance of information, ambient visualization can deliver information more elegantly, aligning with users' attention capability and informational needs.

In addition to the predicted attention state, attention-adaptive ambient visualizations should also consider the prediction credibility for visual encoding. Two properties, *prediction probability* and *consistency*, were suggested to enhance the credibility of a single prediction as well as a sequence of predictions.

Although end-user EEG devices with few channels (e.g., single-channel devices) provide much better usability than medical devices suitable for daily usage, wearing an additional head-mounted device remains a potential

limitation of the BCI method. Casual contexts where people wear a helmet or a head-mounted device, such as bicycle riding [2], VR/AR experiences, and deskwork with noise-canceling earphones and headphones, would be more suitable for proactive utilization of attention-adaptive features.

Finally, data preparation becomes a crucial usability issue when users want to apply the classification of multiple attention states to their personal context. Enhanced interactive support of data collection and labeling would be necessary to build individualized attention state classifiers for end-users.

Related Design Implications

- Enhance the Credibility of Attention State Prediction through Visualization
- Select Usage Scenarios Where Wearing a Head-Mounted EEG Device is Natural
- Improve Interactive Support for End-user Data Collection and Labeling

6.2.3 Integration of Two Personalization Approaches

The semi-structured interview on design opportunities for A3InfoVis with experienced DataHalo users provided a comprehensive view of end-users' needs for personalized ambience. The interview results indicated potential limitations of personalization solely depending on users' design perspectives. With attention-adaptive features, participants considered more diverse personalization strategies, which significantly improved their previous design approaches. The main reason could be the different time scales between the two personalization strategies. Customization using DataHalo occurred

in the long term, primarily through reflection on day-to-day experiences, while real-time computation of attention information yielded predictions every second.

However, due to technological constraints, users' attentional information may not be a universal predictor of their current context. Most participants tended to consider three types of information, current attentional information, the value of delivered information, and simple contextual information like time and location, when thinking of personalization scenarios. As they are easily accessible via smartphone systems. Researchers should consider integrating these factors into personalization systems.

Researchers also need to keep in mind that the attentional states which end-users think of are not always consistent in terms of EEG signals. For example, the residual attention capability which P02 named might be similar to peripheral awareness or not. Moreover, the same word distracted might indicate totally different attentional states. Hence, providers of the end-user customization system which utilizes a certain attentional state as a factor for personalization should consider how to align this gap.

Related Design Implications

- Integrating Attention-adaptive Features can Diversify and Improve the End-user's Personalization Approach
- Additional Contextual Information is Needed to Handle Realistic and Complex Attention Management Scenarios
- Bridge the Gap between EEG Microstates and End-users' Awareness of their Attention State

6.3 Future Research Opportunities for PAAIV

This section suggests future research opportunities to improve the personalization approaches of PAAIV further.

6.3.1 Personalized Visualization Design Empowered by Image Generation Models

Image generation models [31, 48, 49, 71, 77, 87, 90], which computationally generate novel images, have experienced substantial advancements owing to the development of deep learning techniques. Among these, recent text-to-image models such as *Dall-E*[90] and *Stable Diffusion* have demonstrated the ability to produce high-quality images corresponding to natural language prompts. Furthermore, the quality of the generated images can be enhanced by incorporating additional, task-specific conditions (e.g., canny edge, human pose, semantic segmentation results) into diffusion models using *ControlNet*[109]. Employing cutting-edge image generation models can enable end-users to produce a broad range of desirable visual representations with a lower customization burden. In both the usability study and deployment study, many participants enjoyed creating their preferred graphical marks from the available preset pictographs. Although they had other options to access a wider variety of images through the photo gallery, image search, or digital drawing, participants did not choose these alternatives frequently. This can be interpreted as another example of mild customization fatigue, or *cost-effectiveness*: they could effortlessly find an "acceptable" pictograph among the presets, even if it was not ideal. By integrating text-to-image generation capabilities into the customization of visual encoding, end-users' burden of preparing custom images will be significantly reduced.

It could motivate end-users further to try generating more diverse images that express their informational intention or aestheticism.

6.3.2 Extending Expressiveness of Halo Visualization

DataHalo could be further expanded to accommodate more diverse visual expressions. In addition to graphical marks visualizing unit notifications, In future work, DataHalo can additionally provide a new type of aggregated visual expression that represents notifications that belong to the same virtual category, application, or folder. These visual expressions would give a higher level of abstraction over the set of notifications. Blending visual expressions for unit notifications (i.e., graphical marks) and those for aggregated notifications will raise the diversity of visual design. For example, users can represent relatively important notifications as individual graphical marks, while grouping less critical notifications together as an aggregated visual expression.

6.3.3 Design and Evaluation of Attention-Adaptive Ambience Visualization for Smartphone Notifications

This dissertation identified the design challenges of A3InfoVis, verified its technical feasibility, and derived design implications that should be considered to design A3InfoVis. Drawing upon these findings, future research could undertake a design study of implementing an A3InfoVis system for a specific smart display and evaluating the usefulness and effectiveness of attention-adaptive ambience in a real-world context. Extending DataHalo system to incorporate the smartphone users' real-time attention state will be an interesting research topic. It would enable users to create a more flexibly encoded halo visualization. Users could customize graphical marks of the virtual cat-

egories that possibly deliver distracting information (e.g., small chat, the latest update of their favorite webtoons) to be invisible while they are in a highly engaged attention state. Conversely, users could customize the graphical marks of the virtual categories delivering urgent information (e.g., business contacts) to be more noticeable while they are in a resting attention state for a long time. Leveraging attention state classifiers that can distinguish attention states more sophisticatedly would extend the expressiveness of halo visualization even further.

6.3.4 Integration of Other Contextual Information into PAAIV

In addition to the user's attention state, there exist diverse kinds of contextual information that can be used to identify changes in the end-users context. Integrating the information into ambient information visualization systems would allow a more sophisticated personalization of ambience. For example, physiological sensors in smart wearable devices can be used to infer the current stress level of users. Investigating the changes in the suitable level of ambience while users are in a very stressful condition would be an interesting research topic. It would provide valuable insight into the edge case where ambient information visualization might not be effective.

6.4 Limitations

This section addresses the limitations of the conducted studies. As every research approach possesses inherent limitations concerning unforeseen issues and threats to internal or external validity [44], elucidating potential constraints within this dissertation is of significant value.

6.4.1 DataHalo

The present method primarily focuses on extending the notification badge to visualize enhanced notification information for smartphone users. More advancements in user–notification interaction could be realized by allowing the enhanced information to influence other visual components of notification, such as the status bar and notification drawer, and supporting seamless interplay with halo visualization. For example, the notification drawer could facilitate alternative attendance methods, like pinning halo notifications based on their importance models. This approach would grant users intuitive access to individual notifications across app halos. Furthermore, applying halo visualization to other smartphone displays (e.g., Always On Display) or wearable devices could increase accessibility to enhanced notification information, necessitating solutions for battery consumption and privacy leak concerns.

Providing a weekly report of DataHalo usage might have encouraged participants to augment their usage, particularly for halo customization events, despite receiving identical rewards. Future iterations of DataHalo should incorporate features that consistently motivate users to better personalize their user-notification interaction. The limited generalizability of user study results, such as the small sample size in the field study and unbalanced samples concerning age and gender, also constitutes a limitation.

6.4.2 A3InfoVis

The video-watching task was selected to represent casual tasks in a desk work environment. Although the prediction model could accommodate a variety of deskwork scenarios, it might not directly apply to highly interactive tasks, such as online FPS games, due to the necessity for immediate

user responses and ongoing communication with others. Collecting datasets encompassing various contextual data, including behavioral information or other physiological sensors, would be essential to address these tasks.

Given the limited dataset size, the study focused on classical machine learning methods to predict users' attention states. While a feasible performance level was attained, examining the performance of recent deep learning models would be more advantageous. For instance, TabNet [4] demonstrated comparable performance to XGBoost, the selected model, and functioned effectively for self-supervised learning scenarios with a substantial portion of unlabeled data. TabNet's strengths could help alleviate users' data labeling overhead. Additionally, deep learning methods specialized in EEG data (e.g., EEGNet [55], BD-Deep4 [92]) could minimize potential biases during feature selection by internally generating temporal and spatial features.

Chapter 7

Conclusion

In this dissertation, a conceptual framework called *Personally Adaptable Ambient Information Visualization* (*PAAIV*) was proposed to address the limitations of previous ambient information visualizations. PAAIV aimed to offer end-users personalized ambient information visualizations that align with their design perspectives and real-time attentional demands. It expanded upon previous ambient information visualizations through two personalization approaches. First, it allowed end-users to customize the ambience through a flexible and usable interface, supporting them in creating a more satisfying ambient visualization and improving it over time. Second, it proposed an expanded design space where end-users can further improve ambience by reducing distractions based on their real-time attentional needs. The following research questions were addressed in the course of the dissertation.

7.1 Summary of Approaches

To address the first research question (How should we design flexible customization capabilities for PAAIV, allowing people to enhance informational

benefits while effectively managing distractions across diverse types of personal information?), DataHalo, a customizable notification visualization system, was designed and developed. It represents smartphone notifications as prolonged ambient visualizations on the home screen and is implemented as an Android home screen launcher with personalization capabilities. DataHalo's feasibility was assessed through a usability lab study, which demonstrated its potential to embody PAAIV, enabling participants to manage notifications better and rediscover the value of diverse information. However, customization fatigue emerged as a potential obstacle, prompting improvements based on user feedback.

To address the second research question (What would be commonalities and diversities in people's customization strategies for embodying their personalized ambience by utilizing the PAAIV system in everyday contexts?), an improved version of DataHalo was deployed in a real-world context to explore personalization strategies for better notification management. A three-week field deployment study collected usage logs and partial notification logs, revealing that DataHalo met participants' diverse needs and improved satisfaction. Mixed-method analyses identified personalization patterns and provided suggestions for future smartphone notification systems to support users in building and refining their PAAIV more easily and expressively.

To address the third research question (What are the limitations and opportunities for integrating attention-adaptive personalization into PAAIV, considering technical constraints and user needs related to real-time attentional information?), the technological constraints and design opportunities of Attention-Adaptive Ambient Information Visualization (A3InfoVis), were explored. First, three technological challenges were assessed through an empirical study, revealing possibilities and limitations of predicting user's real-

time attentional information. Next, a semi-structured interview with end-users experienced in managing custom ambient information visualizations was conducted to understand their needs for using real-time attentional information to enhance benefits and minimize distractions. Based on these findings, design guidelines for A3InfoVis were established, further expanding the design space of PAAIV.

7.2 Summary of Contributions

This research contributes mainly to the fields of human-computer interaction. The contributions are summarized as follows: (1) the design and implementation of the DataHalo, a flexible notification visualization system that helped people to realize their PAAIV for the information delivered by smartphone notifications; (2) empirical findings, acquired through a field deployment study of DataHalo, that clarified the design motivations and relevant choices posed by smartphone users to achieve their personalized ambience through managing notification information; (3) empirical findings, acquired through an empirical study and a user interview, that provide insights on the technological constraints and design opportunities of integrating attention-adaptive personalization into PAAIV.

7.3 Summary of Future Research Opportunities

This research inspired a number of future opportunities to improve PAAIV.

Expanding PAAIV Approach to Various Devices. DataHalo primarily concentrated on improving notification badges, and future advancements could involve other visual components of smartphone notification systems. The PAAIV approach could also be extended to a broader range of smart displays

and wearable devices. When extending the approach, battery and privacy concerns should be taken into consideration.

Exploring Deep Learning Techniques for Attention State Prediction. With a limited dataset, the study used classical machine learning methods for predicting attention states. Exploring deep learning models could further improve performance and minimize potential biases of manual feature selection.

Investigating Deep Learning Approaches for Enhanced Attention State Prediction. With a limited dataset, A3InfoVis employed classical machine learning techniques for predicting attention states. Exploring deep learning models could further improve performance and minimize potential biases associated with manual feature selection.

Harnessing Image Generation Models to Reduce Customization Burden. Recent advancements in image generation models could enable end-users to create diverse visual representations with a lower customization burden. Integrating text-to-image generation capabilities could reduce end-users' efforts in preparing custom images.

Expanding DataHalo to Accommodate Diverse Visual Expressions. DataHalo could be expanded to include aggregated visual expressions representing sets of notifications, providing a higher level of abstraction and raising the diversity of visual design.

Integrating A3InfoVis with DataHalo. Future research could implement and evaluate A3InfoVis systems on specific smart displays in real-world contexts. Integrating real-time attention states into the DataHalo system could enable more flexible halo visualization and customization based on users' attention levels, ultimately enhancing the expressiveness of the visualization.

Integrating Diverse Contextual Information into PAAIV. Incorporating a variety of contextual information into PAAIV could enable more advanced personalization of the ambience. Examining shifts in optimal ambience under varying user conditions, such as highly stressful situations, may offer significant insights.

Bibliography

[1] Christoph Anderson, Isabel Hübener, Ann-Kathrin Seipp, Sandra Ohly, Klaus David, and Veljko Pejovic. A survey of attention management systems in ubiquitous computing environments. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 2(2), July 2018.

[2] Josh Andres, m.c. schraefel m.c., Nathan Semertzidis, Brahmi Dwivedi, Yutika C. Kulwe, Juerg von Kaenel, and Florian Floyd Mueller. *Introducing peripheral awareness as a neurological state for human-computer integration*. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 2020, pages 1–13.

[3] Efthymios Angelakis, Joel F Lubar, Stamatina Stathopoulou, and John Kounios. Peak alpha frequency: an electroencephalographic measure of cognitive preparedness. *Clinical Neurophysiology*, 115(4):887–897, 2004.

[4] Sercan Ö. Arik and Tomas Pfister. Tabnet: attentive interpretable tabular learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(8):6679–6687, May 2021.

[5] Jonas Auda, Dominik Weber, Alexandra Voit, and Stefan Schneegass. Understanding user preferences towards rule-based notification deferral. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI EA '18, LBW584:1–LBW584:6, Montreal QC, Canada. ACM, 2018.

[6] Saskia Bakker, Elise Hoven, and Berry Eggen. Peripheral interaction: characteristics and considerations. *Personal Ubiquitous Comput.*, 19(1):239–254, January 2015.

[7] Aaron Bangor, Philip Kortum, and James Miller. Determining what individual sus scores mean: adding an adjective rating scale. *J. Usability Studies*, 4(3):114–123, May 2009.

[8] Jared S. Bauer, Sunny Consolvo, Benjamin Greenstein, Jonathan Schooler, Eric Wu, Nathaniel F. Watson, and Julie Kientz. ShutEye: encouraging awareness of healthy sleep recommendations with a mobile, peripheral display. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 1401–1410, Austin, Texas, USA. ACM, 2012.

[9] Chris Berka, Daniel Levendowski, Michelle Lumicao, Alan Yau, Gene Davis, Tristan Zivkovic, Richard Olmstead, Patrice Tremoulet, and Patrick Craven. Eeg correlates of task engagement and mental workload in vigilance, learning, and memory tasks. *Aviation, space, and environmental medicine*, 78:B231–44, June 2007.

[10] James F. Cavanagh and Michael J. Frank. Frontal theta as a mechanism for cognitive control. *Trends in Cognitive Sciences*, 18(8):414–421, 2014.

[11] Seyma Kucukozer Cavdar, Tugba Taskaya-Temizel, Mirco Musolesi, and Peter Tino. A multi-perspective analysis of social context and personal factors in office settings for the design of an effective mobile notification system. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 4(1), March 2020.

[12] Chih-Chung Chang and Chih-Jen Lin. LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2:27:1–27:27, 3, 2011. Software available at `http://www.csie.ntu.edu.tw/~cjlin/libsvm`.

[13] Tianqi Chen and Carlos Guestrin. XGBoost: a scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, pages 785–794, San Francisco, California, USA. ACM, 2016.

[14] Minjoo Cho and Daniel Saakes. Calm automaton: a DIY toolkit for ambient displays. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA '17, pages 393–396, Denver, Colorado, USA. ACM, 2017.

[15] Sommer Christie, Selenia di Fronso, Maurizio Bertollo, and Penny Werthner. Individual alpha peak frequency in ice hockey shooting performance. *Frontiers in Psychology*, 8, 2017.

[16] Laura Lee Colgin. Mechanisms and functions of theta rhythms. *Annual Review of Neuroscience*, 36(1):295–312, 2013. eprint: `https://doi.org/10.1146/annurev-neuro-062012-170330`. PMID: 23724998.

[17] Sunny Consolvo, Predrag Klasnja, David W. McDonald, Daniel Avrahami, Jon Froehlich, Louis LeGrand, Ryan Libby, Keith Mosher, and James A. Landay. Flowers or a robot army?: encouraging awareness & activity with personal, mobile displays. In *Proceedings of the 10th International Conference on Ubiquitous Computing*, UbiComp '08, pages 54–63, Seoul, Korea. ACM, 2008.

[18] Sunny Consolvo, Predrag Klasnja, David W. McDonald, and James A. Landay. Designing for healthy lifestyles: design considerations for mobile technologies to encourage consumer health and wellness. *Found. Trends Hum.-Comput. Interact.*, 6(3–4):167–315, April 2014.

[19] Sunny Consolvo, David W. McDonald, Tammy Toscos, Mike Y. Chen, Jon Froehlich, Beverly Harrison, Predrag Klasnja, Anthony LaMarca, Louis LeGrand, Ryan Libby, Ian Smith, and James A. Landay. Activity sensing in the wild: a field trial of Ubifit garden. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, pages 1797–1806, Florence, Italy. ACM, 2008.

[20] Michael Correll, Dominik Moritz, and Jeffrey Heer. Value-suppressing uncertainty palettes. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, pages 1–11, Montreal QC, Canada. Association for Computing Machinery, 2018.

[21] data.ai. 2017 retrospective: a monumental year for the app economy. `https://www.data.ai/en/insights/market-data/app-annie-2017-retrospective/`, 2019. Retrieved June 28, 2023.

[22] MBC (Moonhwa Broadcasting Cooperation). Summary of the debate on the policies of ministry of gender equality and family. `https://youtu.be/Tr3XP0DvE4s`, 2022. Retrieved June 28, 2023.

[23] MBC (Moonhwa Broadcasting Cooperation). Summary of the debate on the overheated cryptocurrency market. `https://youtu.be/GKQtcmHyf54`, 2022. Retrieved June 28, 2023.

[24] MBC (Moonhwa Broadcasting Cooperation). Summary of the debate on the basic income guarantee. `https://youtu.be/8NzrPgvi2uY`, 2022. Retrieved June 28, 2023.

[25] Merriam-Webster. Definition of brain wave. `https://www.merriam-webster.com/dictionary/brain%20wave`, 2022. Retrieved June 28, 2023.

[26] Andreas K Engel and Pascal Fries. Beta-band oscillations—signalling the status quo? *Current Opinion in Neurobiology*, 20(2):156–165, 2010. Cognitive neuroscience.

[27] Daniel A. Epstein, An Ping, James Fogarty, and Sean A. Munson. A lived informatics model of personal informatics. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, UbiComp '15, pages 731–742, Osaka, Japan. Association for Computing Machinery, 2015.

[28] James Fogarty, Jodi Forlizzi, and Scott E. Hudson. Aesthetic information collages: generating decorative displays that contain information. In *Pro-*

ceedings of the 14th Annual ACM Symposium on User Interface Software and Technology, UIST '01, pages 141–150, Orlando, Florida. ACM, 2001.

[29] John Foxe and Adam Snyder. The role of alpha-band brain oscillations as a sensory suppression mechanism during selective attention. *Frontiers in Psychology*, 2, 2011.

[30] Lukas A.W. Gemein, Robin T. Schirrmeister, Patryk Chrabąszcz, Daniel Wilson, Joschka Boedecker, Andreas Schulze-Bonhage, Frank Hutter, and Tonio Ball. Machine-learning-based diagnostics of eeg pathology. *NeuroImage*, 220:117021, 2020.

[31] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 27. Curran Associates, Inc., 2014.

[32] Google. Digital wellbeing experiments. `https://experiments.withgoogle.com/collection/digitalwellbeing`, 2019. Retrieved June 28, 2023.

[33] Google. Firebase realtime database. `https://firebase.google.com/docs/database`, 2021. Retrieved June 28, 2023.

[34] Rúben Gouveia, Fábio Pereira, Evangelos Karapanos, Sean A. Munson, and Marc Hassenzahl. Exploring the design space of glanceable feedback for physical activity trackers. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, UbiComp '16, pages 144–155, Heidelberg, Germany. Association for Computing Machinery, 2016.

[35] Alexandre Gramfort, Martin Luessi, Eric Larson, Denis A. Engemann, Daniel Strohmeier, Christian Brodbeck, Roman Goj, Mainak Jas, Teon Brooks, Lauri Parkkonen, and Matti S. Hämäläinen. MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*, 7(267):1–13, 2013.

[36] Samuel W Greenhouse and Seymour Geisser. On methods in the analysis of profile data. *Psychometrika*, 24(2):95–112, 1959.

[37] Guhyun Han, Jaehun Jung, Young-Ho Kim, and Jinwook Seo. Datahalo: a customizable notification visualization system for personalized and longitudinal interactions. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23, Hamburg, Germany. Association for Computing Machinery, 2023.

[38] Thalia Harmony. The functional significance of delta oscillations in cognitive processing. *Frontiers in Integrative Neuroscience*, 7, 2013.

[39] Mariam Hassib, Stefan Schneegass, Philipp Eiglsperger, Niels Henze, Albrecht Schmidt, and Florian Alt. Engagemeter: a system for implicit audience engagement sensing using electroencephalography. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pages 5114–5119, Denver, Colorado, USA. Association for Computing Machinery, 2017.

[40] Christoph S. Herrmann, Matthias H.J. Munk, and Andreas K. Engel. Cognitive functions of gamma-band activity: memory match and utilization. *Trends in Cognitive Sciences*, 8(8):347–355, 2004.

[41] Behshad Hosseinifard, Mohammad Hassan Moradi, and Reza Rostami. Classifying depression patients and normal subjects using machine learning techniques and nonlinear features from eeg signal. *Computer Methods and Programs in Biomedicine*, 109(3):339–345, 2013.

[42] Dandan Huang, Melanie Tory, Bon Adriel Aseniero, Lyn Bartram, Scott Bateman, Sheelagh Carpendale, Anthony Tang, and Robert Woodbury. Personal visualization and personal visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, 21(3):420–433, March 2015.

[43] Samuel Huron, Sheelagh Carpendale, Alice Thudt, Anthony Tang, and Michael Mauerer. Constructive visualization. In *Proceedings of the 2014 Conference on Designing Interactive Systems*, DIS '14, pages 433–442, Vancouver, BC, Canada. Association for Computing Machinery, 2014.

[44] MPH James H. Price Editor PhD and Judy Murnan Assistant. Research limitations and the necessity of reporting them. *American Journal of Health Education*, 35(2):66–67, 2004. eprint: `https://doi.org/10.1080/19325037.2004.10603611`.

[45] Xiao Jiang, Gui-Bin Bian, and Zean Tian. Removal of artifacts from eeg signals: a review. *Sensors*, 19(5), 2019.

[46] Pulse (by Maeil Business News Korea. Kakaotalk in its 10th year has nearly all s. korean population connected, 2020. Retrieved June 28, 2023.

[47] Julia W. Y. Kam, Zachary C. Irving, Caitlin Mills, Shawn Patel, Alison Gopnik, and Robert T. Knight. Distinct electrophysiological signatures of task-unrelated and dynamic thoughts. *Proceedings of the National Academy of Sciences*, 118(4):e2011796118, 2021. eprint: `https://www.pnas.org/doi/pdf/10.1073/pnas.2011796118`.

[48] Tero Karras, Samuli Laine, and Timo Aila. A style-based generator architecture for generative adversarial networks, 2019. arXiv: `1812.04948 [cs.NE]`.

[49] Tero Karras, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. Analyzing and improving the image quality of StyleGAN. In *Proc. CVPR*, 2020.

[50] Minhyung Kim, Inyeop Kim, and Uichin Lee. Beneficial neglect: instant message notification handling behaviors and academic performance. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 5(1), March 2021.

[51] Nam Wook Kim, Hyejin Im, Nathalie Henry Riche, Alicia Wang, Krzysztof Gajos, and Hanspeter Pfister. DataSelfie: empowering people to design personalized visuals to represent their data. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, 79:1–79:12, Glasgow, Scotland Uk. ACM, 2019.

[52] Young-Ho Kim, Jae Ho Jeon, Eun Kyoung Choe, Bongshin Lee, KwonHyun Kim, and Jinwook Seo. TimeAware: Leveraging Framing Effects to Enhance Personal Productivity. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 272–283, Santa Clara, California, USA. ACM, 2016.

[53] Jan Kučera, James Scott, Nicholas Chen, Patrick Olivier, and Steve Hodges. Towards calm displays: matching ambient illumination in bedrooms. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 1(2), June 2017.

[54] Kostadin Kushlev, Jason Proulx, and Elizabeth W. Dunn. "silence your phones": smartphone notifications increase inattention and hyperactivity symptoms. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 1011–1020, San Jose, California, USA. Association for Computing Machinery, 2016.

[55] Vernon J. Lawhern, Amelia J. Solon, Nicholas R. Waytowich, Stephen M. Gordon, Chou P. Hung, and Brent J. Lance. Eegnet: A compact convolutional network for eeg-based brain-computer interfaces. *CoRR*, abs/1611.08024, 2016. arXiv: `1611.08024`.

[56] Yu-Kang Lee, Chun-Tuan Chang, You Lin, and Zhao-Hong Cheng. The dark side of smartphone usage: psychological traits, compulsive behavior and technostress. *Computers in Human Behavior*, 31:373–383, 2014.

[57] Kyung-Ryong Lee, Geon-il Goh, and Young-Woo Park. Quietto: an interactive timepiece molded in concrete and milled wood. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pages 2988–2992, Denver, Colorado, USA. Association for Computing Machinery, 2017.

[58] Kyung-Ryong Lee, Somi Ju, Temirlan Dzhoroev, Geonil Goh, Moon-Hwan Lee, and Young-Woo Park. *Dayclo: an everyday table clock providing interaction with personal schedule data for self-reflection*. In *Proceedings of the 2020 ACM De-*

signing Interactive Systems Conference. Association for Computing Machinery, New York, NY, USA, 2020, pages 1793–1806.

[59] Moon-Hwan Lee, Yea-Kyung Row, Oosung Son, Uichin Lee, Jaejeung Kim, Jungi Jeong, Seungryoul Maeng, and Tek-Jin Nam. Flower-pop: facilitating casual group conversations with multiple mobile devices. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 1(4), January 2018.

[60] Natália Lelis-Torres, Herbert Ugrinowitsch, Tércio Apolinário-Souza, Rodolfo N Benda, and Guilherme M Lage. Task engagement and mental workload involved in variation and repetition of a motor skill. *Scientific Reports*, 7(1):1–10, 2017.

[61] Tzu-Chieh Lin, Yu-Shao Su, Emily Helen Yang, Yun Han Chen, Hao-Ping Lee, and Yung-Ju Chang. "put it on the top, i'll read it later": investigating users' desired display order for smartphone notifications. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, New York, NY, USA. Association for Computing Machinery, 2021.

[62] Scott M Lundberg and Su-In Lee. A unified approach to interpreting model predictions. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems 30*, pages 4765–4774. Curran Associates, Inc., 2017.

[63] Jennifer Mankoff, Anind K. Dey, Gary Hsieh, Julie Kientz, Scott Lederer, and Morgan Ames. Heuristic evaluation of ambient displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '03, pages 169–176, Ft. Lauderdale, Florida, USA. ACM, 2003.

[64] Tara Matthews. Designing and evaluating glanceable peripheral displays. In *Proceedings of the 6th Conference on Designing Interactive Systems*, DIS '06, pages 343–345, University Park, PA, USA. Association for Computing Machinery, 2006.

[65] Tara Matthews, Tye Rattenbury, and Scott Carter. Defining, designing, and evaluating peripheral displays: an analysis using activity theory. *Hum.-Comput. Interact.*, 22(1):221–261, May 2007.

[66] Timothy McMahan, Ian Parberry, and Thomas D. Parsons. Evaluating player task engagement and arousal using electroencephalography. *Procedia Manufacturing*, 3:2303–2310, 2015. 6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015) and the Affiliated Conferences, AHFE 2015.

[67] Abhinav Mehrotra, Robert Hendley, and Mirco Musolesi. Prefminer: mining user's preferences for intelligent mobile notification management. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and*

Ubiquitous Computing, UbiComp '16, pages 1223–1234, Heidelberg, Germany. Association for Computing Machinery, 2016.

[68] Abhinav Mehrotra, Mirco Musolesi, Robert Hendley, and Veljko Pejovic. Designing content-driven intelligent notification mechanisms for mobile applications. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, UbiComp '15, pages 813–824, Osaka, Japan. Association for Computing Machinery, 2015.

[69] Gonzalo Gabriel Méndez, Uta Hinrichs, and Miguel A. Nacenta. Bottom-up vs. top-down: trade-offs in efficiency, understanding, freedom and creativity with infovis tools. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pages 841–852, Denver, Colorado, USA. Association for Computing Machinery, 2017.

[70] Todd Miller and John Stasko. Artistically conveying peripheral information with the infocanvas. In *Proceedings of the Working Conference on Advanced Visual Interfaces*, AVI '02, pages 43–50, Trento, Italy. ACM, 2002.

[71] Mehdi Mirza and Simon Osindero. Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784*, 2014.

[72] Carol Moser, Sarita Y. Schoenebeck, and Katharina Reinecke. Technology at the table: attitudes about mobile phone use at mealtimes. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 1881–1892, San Jose, California, USA. ACM, 2016.

[73] mydloch's dev-team. Notification history. `https://play.google.com/store/apps/details?id=com.myd.android.nhistory2`, 2021. Retrieved June 28, 2023.

[74] J. K. Nuamah and Younho Seong. Support vector machine (svm) classification of cognitive tasks based on electroencephalography (eeg) engagement index. *Brain-Computer Interfaces*, 5(1):1–12, 2018. eprint: `https://doi.org/10.1080/2326263X.2017.1338012`.

[75] Joseph K. Nuamah, Younho Seong, and Sun Yi. Electroencephalography (eeg) classification of cognitive tasks based on task engagement index. In *2017 IEEE Conference on Cognitive and Computational Aspects of Situation Management (CogSIMA)*, pages 1–6, March 2017.

[76] Chunjong Park, Junsung Lim, Juho Kim, Sung-Ju Lee, and Dongman Lee. Don't bother me. i'm socializing!: a breakpoint-based smartphone notification system. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, CSCW '17, pages 541–554, Portland, Oregon, USA. ACM, 2017.

[77] Taesung Park, Ming-Yu Liu, Ting-Chun Wang, and Jun-Yan Zhu. Gaugan: semantic image synthesis with spatially adaptive normalization. In *ACM SIGGRAPH 2019 Real-Time Live!*, SIGGRAPH '19, Los Angeles, California. Association for Computing Machinery, 2019.

[78] Christie Picken, Adam R. Clarke, Robert J. Barry, Rory McCarthy, and Mark Selikowitz. The theta/beta ratio as an index of cognitive processing in adults with the combined type of attention deficit hyperactivity disorder. *Clinical EEG and Neuroscience*, 51(3):167–173, 2020. eprint: `https://doi.org/10.1177/1550059419895142`. PMID: 31875684.

[79] Martin Pielot, Bruno Cardoso, Kleomenis Katevas, Joan Serrà, Aleksandar Matic, and Nuria Oliver. Beyond interruptibility: predicting opportune moments to engage mobile phone users. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 1(3), September 2017.

[80] Martin Pielot, Karen Church, and Rodrigo de Oliveira. An in-situ study of mobile phone notifications. In *Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices & Services*, MobileHCI '14, pages 233–242, Toronto, ON, Canada. ACM, 2014.

[81] Martin Pielot and Luz Rello. Productive, anxious, lonely: 24 hours without push notifications. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '17, 11:1–11:11, Vienna, Austria. ACM, 2017.

[82] Martin Pielot, Amalia Vradi, and Souneil Park. Dismissed!: a detailed exploration of how mobile phone users handle push notifications. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '18, 3:1–3:11, Barcelona, Spain. ACM, 2018.

[83] John Polich. Updating p300: an integrative theory of p3a and p3b. *Clinical neurophysiology*, 118(10):2128–2148, 2007.

[84] Alan T Pope, Edward H Bogart, and Debbie S Bartolome. Biocybernetic system evaluates indices of operator engagement in automated task. *Biological Psychology*, 40(1):187–195, 1995. EEG in Basic and Applied Settings.

[85] Zachary Pousman and John Stasko. A taxonomy of ambient information systems: four patterns of design. In *Proceedings of the Working Conference on Advanced Visual Interfaces*, AVI '06, pages 67–74, Venezia, Italy. ACM, 2006.

[86] Zachary Pousman, John Stasko, and Michael Mateas. Casual information visualization: depictions of data in everyday life. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1145–1152, November 2007.

[87] Aditya Ramesh, Mikhail Pavlov, Gabriel Goh, Scott Gray, Chelsea Voss, Alec Radford, Mark Chen, and Ilya Sutskever. Zero-shot text-to-image gen-

eration. In Marina Meila and Tong Zhang, editors, *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pages 8821–8831. PMLR, 18–24 Jul 2021.

[88] Counterpoint Research. Smartphone market share of south korea in 2021 (in korean). https://korea.counterpointresearch.com/2022ëĚĎ-êμŋëĆť-ìŁďëğĹíŁÿíŘř-ìŃœìđě-1680ëğŒëŒĂ-ìŸĹìČĄ-ìăĎëĚĎëŒĂëźĎ-ìĘŒíŘŋ-ìČĄ/, 2022. Retrieved June 28, 2023.

[89] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. "why should I trust you?": explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, August 13-17, 2016*, pages 1135–1144, 2016.

[90] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 10684–10695, 2022.

[91] Alireza Sahami Shirazi, Niels Henze, Tilman Dingler, Martin Pielot, Dominik Weber, and Albrecht Schmidt. Large-scale assessment of mobile notifications. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, pages 3055–3064, Toronto, Ontario, Canada. ACM, 2014.

[92] Robin Tibor Schirrmeister, Jost Tobias Springenberg, Lukas Dominique Josef Fiederer, Martin Glasstetter, Katharina Eggensperger, Michael Tangermann, Frank Hutter, Wolfram Burgard, and Tonio Ball. Deep learning with convolutional neural networks for eeg decoding and visualization. *Human brain mapping*, 38(11):5391–5420, 2017.

[93] N. Sadat Shami, Gilly Leshed, and David Klein. Context of use evaluation of peripheral displays (cuepd). In *Proceedings of the 2005 IFIP TC13 International Conference on Human-Computer Interaction*, INTERACT'05, pages 579–587, Rome, Italy. Springer-Verlag, 2005.

[94] In-geon Shin, Jin-min Seok, and Youn-kyung Lim. Too close and crowded: understanding stress on mobile instant messengers based on proxemics. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, 615:1–615:12, Montreal QC, Canada. ACM, 2018.

[95] Tobias Skog. Activity wallpaper: ambient visualization of activity information. In *Proceedings of the 5th Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques*, DIS '04, pages 325–328, Cambridge, MA, USA. Association for Computing Machinery, 2004.

[96] Tobias Skog, Sara Ljungblad, and Lars Erik Holmquist. Between aesthetics and utility: designing ambient information visualizations. In *Proceedings of*

the Ninth Annual IEEE Conference on Information Visualization, INFOVIS'03, pages 233–240, Seattle, Washington. IEEE Computer Society, 2003.

[97] John Stasko, Todd Miller, Zachary Pousman, Christopher Plaue, and Osman Ullah. Personalized peripheral information awareness through information art. In *International conference on ubiquitous computing*, pages 18–35. Springer, 2004.

[98] Daniel Szafir and Bilge Mutlu. Pay attention! designing adaptive agents that monitor and improve user engagement. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 11–20, Austin, Texas, USA. Association for Computing Machinery, 2012.

[99] Damian A. Tamburri. Sustainable mlops: trends and challenges. In *2020 22nd International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC)*, pages 17–23, 2020.

[100] Janet P. Trammell, Priscilla G. MacRae, Greta Davis, Dylan Bergstedt, and Ariana E. Anderson. The relationship of cognitive performance and the theta-alpha power ratio is age-dependent: an eeg study of short term memory and reasoning during task and resting-state in healthy young and old adults. *Frontiers in Aging Neuroscience*, 9, 2017.

[101] Dominik Weber, Alexandra Voit, Jonas Auda, Stefan Schneegass, and Niels Henze. Snooze!: investigating the user-defined deferral of mobile notifications. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '18, 2:1–2:13, Barcelona, Spain. ACM, 2018.

[102] Dominik Weber, Alexandra Voit, and Niels Henze. Notification log: an open-source framework for notification research on mobile devices. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*, UbiComp '18, pages 1271–1278, Singapore, Singapore. Association for Computing Machinery, 2018.

[103] Dominik Weber, Alexandra Voit, Gisela Kollotzek, and Niels Henze. Annotif: a system for annotating mobile notifcations in user studies. In *Proceedings of the 18th International Conference on Mobile and Ubiquitous Multimedia*, MUM '19, Pisa, Italy. Association for Computing Machinery, 2019.

[104] Mark Weiser and John Seely Brown. Designing calm technology. *POWERGRID JOURNAL*, 1, 1996.

[105] Peter Welch. The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. *IEEE Transactions on audio and electroacoustics*, 15(2):70–73, 1967.

[106] Ybrain. Mindd-scan: wireless eeg system. https://www.ybrain.com/en/product-professional/, 2022. Retrieved June 28, 2023.

[107] Amir Zaidi. Rootless pixel launcher. https://github.com/amirzaidi/Launcher3, 2019.

[108] Leizhong Zhang, Nan Tu, and Dave Vronay. Info-lotus: a peripheral visualization for email notification. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '05, pages 1901–1904, Portland, OR, USA. ACM, 2005.

[109] Lvmin Zhang and Maneesh Agrawala. Adding conditional control to text-to-image diffusion models. *arXiv preprint arXiv:2302.05543*, 2023.

APPENDIX A

Halo Visualization

A.1 Connection with Constructive Visualization

Components	Corresponding Design Constraints of *Halo*
Basic Token	An enhanced notification from the target app
Token Grammar	User-defined visual encoding for position, size, shape, color, and motion of graphical marks
Environment	Layer of home screen over that of wallpaper and under that of app shortcuts
Assembly Model	A clockwise arrangement around the target app shortcut starting from the top according to user-defined importance

Figure A.1: The components of constructive visualization and corresponding design constraints of *halo* visualization. **Basic Token** refers to a discrete graphical mark representing a data unit. **Token Grammar** refers to a mapping of a graphical mark's visual channels to data properties. **Environment** refers to a spatial constraint on the assembling methods of tokens based on Token Grammar. **Assembly Model** refers to a construction model of the visual representation.

APPENDIX B

Usability Study of DataHalo

B.1 Virtual Categories during the Task 2

	App	Description	Virtual Category / Reason of Creation
U01	Facebook	Social media	**Notice** - to subscribe to announcements from several Facebook pages **Social** - to curate the notice about other user's reactions (e.g., Like, Comment)
U02	Messages	Messenger	**Church** - to curate messages related to religious activity **Private Tutoring** - to curate messages sent by tutee
U03	Youtube	Video streaming	**[Youtuber_Name]** - to subscribe to the updates of the Youtube channel
U04	Kakao Story	Social media	**Story** - to subscribe to the story updates of social feeds
U05	Bank Salad	Financing	**Coffee** - to subscribe to the purchase history of coffee **Chicken** - to subscribe to the purchase history of fried chicken
U06	Megabox	Movie ticketing	**Event** - to subscribe to the notice of discount events
U07	Gmail	Email	**Shopping** - to subscribe to information letters from shopping mall **Major** - to subscribe to mails sent by the university faculty **Payment** - to subscribe to mails notifying purchase cancellations
U08	Cake	Language learning	**Untitled** - to subscribe to the notice of encouraging user to study English
U09	BaeMin	Food delivery	**Coupon** - to subscribe to the notice of discount events
U10	Clash Royale	Game	**Chest** - to subscribe to the arrival of free game items
U11	Outlook	Email	**Babymind** - to subscribe to the email informing the update of github storage **Lab** - to subscribe the email sent by lab people
U12	Instagram	Social media	**Notice** - to curate the notice about user's reaction
U13	PanelPower	Paid survey	**Survey** - to subscribe to the notice about new paid survey
U14	Naver Mail	Email	**University** - to subscribe to mails sent by faculty
U15	Hana Bank	Financing	**Deposit** - to subscribe to the notice informing the event of account deposit
U16	WhatsApp	Messenger	**[User_id]** - to curate the messages sent by the user
U17	Messages	Messenger	**Financial Service** - to subscribe to the messages sent by bank

Table B.1: Smartphone applications that the usability study participants chose for the second halo design task and virtual categories they created.

APPENDIX C

Deployment Study of DataHalo

C.1 Post-study questionnaire

Types	Questions
Q1. Notification Categorization	I could better categorize the information delivered by notifications of interest.
Q2. Importance Modeling	I could better prioritize notifications based on my information assessment.
Q3. Longitudinal Interaction	I could keep in touch with the longitudinal information as long as I wanted.
Q4. Visual Expression	I could access the useful information with my preferable visual representations
Q5. User Satisfaction	DataHalo provided more satisfactory experience of notification management to me.
Q6. Rediscovery of Information	I became more interested in the informative values of my daily notifications.
Q7. Continue to Use	If possible, I hope to continue using the personalization capabilities of DataHalo.

Table C.1: Questions for inquiring the user experience during the DataHalo usage.

C.2 Popular Design Strategies

Notification Control	Information Category	User Response	Information Complexity	Emphasis on Presentation	#P (Freq.)	Example Usage
User-defined virtual category						
selective	personal subscription	remind	change of states	user's impression	4 (4)	To subscribe DataHalo study announcement via instant messenger
as-is	personal subscription	recognition	identification	user's interpretation	3 (6)	To subscribe the withdrawl and deposit events of one's account via bank app
selective	curated communication	reaction	change of states	user's impression	3 (3)	To prioritize the chat with one's beloved via instant messenger
selective	curated communication	recognition	identification	raw information	2 (10)	To distinguish a large number (e.g., 9) of notification groups in the instant messenger with group labels
as-is	self-organized events	remind	identification	user's interpretation	1 (5)	To remind daily goals (e.g., taking a multivitamin tablet) registered in self tracking app
"Remainder" category						
selective	unclassified	recognition	existence	no emphasis	12 (26)	To simply visualize the remaining notifications unhandled by user created groups
block	unclassified	disregard	existence	no emphasis	4 (10)	To deactivate app halo for entire notifications (no user created groups)

Table C.2: Frequent and widely used design strategies among participants. *#P* stands for the number of participants who shared the strategy and *Freq.* stands for the total frequency of the strategy.

국문 초록

앰비언트 정보 시각화(Ambient Information Visualization)는 정보의 범람 속에서 사용자가 겪는 주의 분산을 줄이되, 정보로부터 얻을 수 있는 편익은 증대하기 위해 제안된 시각화 방법론이다. 개별 사용자의 정보 디스플레이에 다양한 일상 정보를 몰입감 높은 시각 표상으로 표현하고, 지속적으로 상호작용 가능하도록 하여 만족스러운 사용자 경험을 제공하는 것을 목표로 한다. 그러나 기존 방법론에는 두 가지 한계가 있다. 사용자 맞춤형 설계가 어렵고 주의 분산 여부도 실시간 반영할 수 없다.

이 논문에서는 개인 맞춤형 및 적응형 앰비언트 정보 시각화(Personally Adaptable Ambient Information Visualization; PAAIV)라는 새로운 개념적 프레임워크를 제시한다. 사용자는 PAAIV를 통해 맞춤형으로 정보를 선택, 전달, 표현하는 시각화를 생성할 수 있으며, 실제 주의 상태의 변화에 적합한 방식으로 상호작용할 수 있다. 이를 달성하기 위해 아래의 세 개의 구체화된 문제를 탐구하였다.

첫째, 다양한 일상 정보에 적용 가능한 유연한 사용자 맞춤 기능을 설계하는 방법을 탐구하였다. 가장 주요하되 부작용 또한 크다고 알려진 일상 정보원인 스마트폰 알림을 대상으로 개인 맞춤형 앰비언트 알림 시각화 시스템인 DataHalo를 설계 및 구현하였다. DataHalo 사용자는 관심 알림을 원하는 기간, 원하는 형태로 스마트폰 홈스크린에 표현할 수 있었다. 사용성 평가(usability study) 결과, DataHalo는 개별 사용자에게 충분한 맞춤 기능을 제공하였다. 또한, 수많은 알림 속에서 놓치기 쉬웠던 일상 정보의 가치를 재발견할 수 있도록 도왔다.

둘째, 개별 사용자들이 일상에서 맞춤형 앰비언트 시각화를 설계하는 전략에 어떤 공통점과 차이점이 있는지 탐구하였다. 사용성이 개선된 DataHalo를 3주간 현장 배포하여 연구(field deployment study)하였다. DataHalo를 활용하면 실생활에서 알림 정보와의 상호작용이 더욱 만족스러워짐을 확인하였다. 배포 중

수집한 로그 및 인터뷰 데이터를 통합연구 및 분석(mixed method analysis)한 결과, 참가자들의 다양한 개인화 방법들로부터 다층적 경향성을 확인하였다. 이를 통해, PAAIV를 만들고 수정하는 과정을 더 효과적으로 돕기 위해 미래의 정보 시스템이 갖춰야 할 점들을 제안하였다.

셋째, 주의적응형(attention-adaptive) 앰비언트 시각화 설계를 위한 기술적 제한사항 및 새로운 디자인 가능성을 탐구하였다. 먼저, 일상 맥락에서 뇌전도(Electroencephalography; EEG) 신호 패턴으로부터 실시간 주의 상태를 추론하는 기술을 활용하기 위한 도전과제를 제시했으며, 제한사항은 무엇인지 실증연구(empirical study) 하였다. 다음으로, 반구조화 인터뷰(semi-structured interview)를 통해 일반 사용자가 일상에서 주의집중 정보를 어떻게 활용하고자 하는지 분석하였다. 이를 통해, 기술과 사용자 측면의 요구사항을 포괄한 주의적응형 PAAIV의 디자인 지침사항(design guideline)을 도출하였다.

본 논문의 핵심 성과는 DataHalo를 설계 및 구현하여 얻은 기술적 산출물, 현장 배포 연구를 통해 개인별 맞춤형 앰비언스 설계 경향을 실증적으로 분석한 결과, 그리고 실시간 주의적응형 기술을 앰비언트 정보시각화에 적용하기 위한 기술 및 사용자 관점의 분석 결과이다. 이를 통해 다양한 종류의 스마트 기기 환경에서 어떻게 맞춤형 앰비언트 정보 시각화 시스템을 설계할지 지침을 제공하고, 사용자를 방해하지 않으면서 더 풍부한 정보편익(informational benefit)을 제공하는 조용한 기술(calm technology)의 실현에 기여한다.

주요어: 앰비언트 정보 시각화; 개인화; 스마트폰 알림; 현장 데이터; 뇌–컴퓨터 인터페이스; 뇌전도

학번: 2013-20899

감사의 글

긴 학위 기간 이루 말할 수 없을 정도로 많은 분께 신세를 졌습니다. 특히 서진욱 교수님과 인간-컴퓨터 상호작용 연구실의 선후배들께 감사합니다. 영민하면서도 따뜻한 마음을 지닌 분들과 함께했던 시간 덕분에 연구자로서뿐만 아니라 인간으로서도 더욱 성숙할 수 있었습니다. 가족들에게도 고마운 마음뿐입니다. 제 선택을 지지해 주고 한결같이 응원해 준 덕분에 불안을 이겨낼 수 있었습니다. 여러분께 받았던 소중한 마음을 가슴에 새기고 새로운 인생의 장으로 나아가겠습니다.

2023년 8월

한구현