The Orchive : A system for semi-automatic annotation and analysis of a large collection of bioacoustic recordings

by

Steven Ness
B.Sc., University of Alberta, 1994
M.Sc., University of Victoria, 2010

A Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Computer Science

© Steven R. Ness, 2013
University of Victoria

All rights reserved. This dissertation may not be reproduced in whole or in part, by photocopying or other means, without the permission of the author.

The Orchive : A system for semi-automatic annotation and analysis of a large collection of bioacoustic recordings

by

Steven Ness
B.Sc., University of Alberta, 1994
M.Sc., University of Victoria, 2010

Supervisory Committee

Dr. George Tzanetakis, Supervisor
(Department of Computer Science)

Dr. Margaret-Anne Storey, Departmental Member
(Department of Computer Science)

Dr. Peter Driessen, Outside Member
(Department of Electrical Engineering)

Supervisory Committee

Dr. George Tzanetakis, Supervisor
(Department of Computer Science)

Dr. Margaret-Anne Storey, Departmental Member
(Department of Computer Science)

Dr. Peter Driessen, Outside Member
(Department of Electrical Engineering)

ABSTRACT

Advances in computer technology have enabled the collection, digitization and automated processing of huge archives of bioacoustic sound. Many of the tools previously used in bioacoustics work well with small to medium-sized audio collections, but are challenged when processing large collections of tens of terabytes to petabyte size. In this thesis, a system is presented that assists researchers to listen to, view, annotate and run advanced audio feature extraction and machine learning algorithms on these audio recordings. This system is designed to scale to petabyte size. In addition, this system allows citizen scientists to participate in the process of annotating these large archives using a casual game metaphor. In this thesis, the use of this system to annotate a large audio archive called the Orchive will be evaluated. The Orchive contains over 20,000 hours of orca vocalizations collected over the course of 30 years, and represents one of the largest continuous collections of bioacoustic recordings in the world. The effectiveness of our semi-automatic approach for deriving knowledge from these recordings will be evaluated and results showing the utility of this system will be shown.

Contents

Supervisory Committee **ii**

Abstract **iii**

Table of Contents **iv**

List of Tables **viii**

List of Figures **ix**

Acknowledgements **xi**

Dedication **xii**

1 Introduction **1**

2 Personal Visualization and Personal Visual Analytics **8**

2.1 Background 8

2.2 Review Method and Process 10

2.3 Design Dimensions and Research Interest to Date 11

2.3.1 Design Dimensions 11

2.3.2 Research Interest to Date 15

2.4 Design Challenges in PV&PVA 18

2.4.1 Fit in Personal Routines and Environments 19

2.4.2 Recall of Relevant Context for Reasoning 19

2.4.3 Defining Appropriate Baselines 20

2.4.4 Sharing and Privacy 21

2.4.5 Evaluation 22

2.5 Contribution and Limitation 23

3 Related Work **24**
3.1 Feedback Design 24
3.1.1 Feedback Applications 25
3.2 persuasive design 27
3.3 Ambient Visualizatoin 28
3.4 Context use in Feedback Design 29
3.5 Evaluation of Personal Visualization in Everyday Context 30

4 Visualization Design **32**

5 Research Methods **36**
5.1 Research Methods 37
5.1.1 Lab Experiment 37
5.1.2 Field Study 38
5.1.3 Survey 38
5.1.4 Interview 39
5.2 Summary 39

6 Viability Study **40**
6.1 Background 40
6.2 Experiment Design 41
6.2.1 Apparatus 41
6.2.2 Participants 42
6.2.3 Experiment I: Calendar Tasks 42
6.3 Experiment II: Visualization Tasks 42
6.4 Procedure 43
6.5 Experiment Results 43
6.5.1 Experiment I: Visual Interference 44
6.5.2 Experiment II: Graphical Perception 45
6.5.3 Post-study Questionnaire 47
6.6 Discussion of Lab Experiment Results 51
6.6.1 Interference (Experiment I) 51
6.6.2 Perception (Experiment II) 51
6.6.3 Design Implications 52
6.6.4 Attentional Ambience 52

7 Implementation **54**

8 Pilot Studies **57**
8.1 Household Energy Consumption 57
8.2 Personal Fitness 59
8.3 Summary of Pilot Studies 61
8.4 Design revision 62

9 Field Study **64**
9.1 Participants 65
9.2 Conditions 66
9.3 Procedure 66
9.4 Data collection 67
9.5 Results 68
9.5.1 Physical activity levels 68
9.5.2 System Use 69
9.5.3 A Model of the Behavior Feedback Process 71
9.5.4 Effects of the On-Calendar Visualization 75
9.5.5 Context for Reasoning 78
9.5.6 Encouraging Ongoing Use 80
9.5.7 Feeeback Tools for Seniors 81
9.6 Discussion 82
9.6.1 Reflection on the On-Calendar Visualization 82
9.6.2 Understanding versus Coersion 85
9.6.3 Feedback Model and Related Models 86
9.6.4 Design implication 87
9.7 Limitation of the study 89

10 Future work **90**

11 Conclusion **92**

A Lab Experiemtn Tasks in Viability Study **94**

B Post-Experiment Questinnaire in Lab Experiments **95**

C Interview Outlines in Pilot Study **96**

D Screenshots of On-Calendar Application (Final Version) 97

E Background Questionnaire in Field Study 98

F Interview outlines in Pilot Study 99

G Screenshots of Fitbit Online Tool 100

H International Physical Activity Questionnaire (IPAQ) 101

Bibliography 107

List of Tables

Table 2.1 Design dimensions, levels with examples from the literature . . . 12
Table 2.2 Summary of surveyed papers 13

Table 6.1 Accuracy Rates (%) 44

Table 9.1 participants in Fitbit field study 65
Table 9.2 Most frequently used contextual information for reasoning (normalized as frequency per participant). 78

List of Figures

Figure 2.1 PV&PVA design dimensions (parallel axes) and surveyed tools (first axis). Box sizes indicate the number of tools with each classification. Linked highlighting enables cluster exploration. . 14

Figure 4.1 Design alternatives displayed overlaped (top: line graph; middle: colored region; bottom: luminance) 34
Figure 4.2 Design alternatives displayed side by side (top: line graph; middle: colored region; bottom: luminance) 35

Figure 6.1 Boxplots showing task time of Experiment I) 46
Figure 6.2 Boxplots showing task time of Experiment II) 47
Figure 6.3 Visual distraction reported by participants (-2 is "very distracting", 2 is "not distracting") 48
Figure 6.4 Graphical perception reported by participants (-2 is "very difficult" and 2 is "very easy"). 48
Figure 6.5 Aesthetics reported by participants (-2 represents "very poor" and 2 represents "very good") 49

Figure 7.1 web application of on-calendar visualization using Google calendar API and displaying household smart meter data. 54
Figure 7.2 Data flow of the on-calendar visualization system. 55
Figure 7.3 Architecture of the on-calendar visualization system. 56

Figure 8.1 Westhouse: eco-friendly smart home for the case study. 58
Figure 8.2 Fitbit trackers use in case study of personal fitness 60
Figure 8.3 On-calendar feedback application used in the study (week view with Fitbit data displayed as a line graph overlapped with calendar events). See more screenshots in Appendix **??**. 63

Figure 9.1 Study procedure 66

Figure 9.2 Change in MET values from weeks 1-2 (baseline) to weeks 3-8 (intervention) for individuals in control and experiment groups. Each mark represents one participants change in average MET scores. 68
Figure 9.3 Change in MET values from weeks 1-2 (baseline) to weeks 3-8 (intervention) for individuals in control and experiment groups. Each mark represents one participants change in average MET scores. 69
Figure 9.4 Preferred visualization settings (experiment group). The most popular visual encoding was a grey line chart overlapped with the calendar data in week view. 70
Figure 9.5 Model of behavior feedback process. 71

ACKNOWLEDGEMENTS

I would like to thank:

DEDICATION

Chapter 1

Introduction

Feedback tools have been attracting great attention in the recent years as the sensering technology is getting cheaper and more accessible. A large variety of tools are designed and developed to assist people to collect data about themselves. Computing technology has never been so close to everyone, who is and is *not* an expert. Designs of the user interface of the feedback application has been substaintially devoted to coerce users into behavior change, for example, taking responsible energy behaviors at home, engaging a healthy life style, etc. With those tools and applications, people could know theirselves better, be more informed about behavior choices and take more control of their lives than before. However, it is not perfect as expected with a close look. The behavior change does not necessarily result in understanding the cause and impact of behvior chioces. In many cases people may take the behavior suggestion even without understanding the behavior itself [51, 71]. Lack of context for reflection has been reported as one of the major barriers for personal information systems (Chapter 2). For example, evidence from a recent study [87] showed that it is very difficult for people to link their daily routines to their residential energy consumption. Meanwhile, feedback design also has to face the problem of low usage after passing the curiosity phase [39]. It is not uncommon that people keep trying new applications and drop them after losing interest. Lack of time was reported as one of the substainal barriers of ongoing adoption for personal informatics tools [74]. Temporal drop-offs in use are quite common to feedback systems [16, 61, 15, 5, 89]. The challenge is how to supoort interpretation and the development of knowledge, and meawhile faciliate the easy access to data within the context that requires minimal learning and management cost, sustaining the ongoing use.

with respect to the challenge, questions are how to provide context in designs to

help people understand their personal data and how to incorperate design components to encourage ongoing use of feedback applications. First, designs should make the information fit naturally into peoples life patterns, leveraging their current information use habits. Second the tools need to provide sufficient contextual information to understand not only the data patterns but also their causes. Finally, the effort required to gather and to integrate the various streams of data should be minimized: simply adding yet another app defeats this goal.

Towards this end, I propose to integrate visualization of personal data (specifically data about household energy use and personal fitness in this thesis) into people's existing information eco-system, combining them with common information tools that represent the desired context (specifically, personal digital calendars in this thesis). Thus, *the goal of the thesis is to investigate if the mash-up approach, integrating visualizaiton of personal data on a digital calendar, is a viable design approach that is comprehensable and not interfere people's existing information use habit (that is, regular calendar use), how people interpret personal data with the context provided by their calendars, and how people react and use the on-calendar visualization in the field as a feedback tool.* That is, this goal is not to develop novel visualizations per se, but rather to augment what people already use, thereby leveraging familiarity, reducing the overhead of using a separate application, and encouraging frequent observations of the data and ongoing use of feedback applications.

In this thesis, the work focuses on applying the on-calendar visualization in feedback applications, specifically feedback for personal fitness and household energy conservation. Considerably, these data are common in people's life, and understanding these data is important for individuals. It can provide ongoing feedback of one's behaviors. Better understanding these patterns can help individuals make better behavior choices to influence their lives. Visualization of these feedback data typically focus on answering three types of questions: *what* ("what is the current status or progress"), *why* ("why are the data patterns like this", "why did an anomaly happen"), and *how* ("how could I improve"). Self realization and improvement depend on first understanding what has happened in the past, then reasoning about why the past data is the way it is, and finally using that insight to identify realistic changes that can make a difference. The commonly used persuasive design strategy in this field usually focuses on *how*, promoting action in the moment. In contrast, my interest is to investigate how visualization design could enhance the reflective understanding of ones behaviors in a non-persuasive way. The general philosophy centers around

helping people know their behavior with operational understanding [16] rather than digitally lecturing them into behavior change. In this work, I view understanding as a first (but not the only) step towards encouraging new habits; therefore, tools that can succeed at this step have made progress.

Meanwhile, onging adoption is a commonly encountered barrier in personal information systems, especially in feedback designs, see chapter 2. People have lower usage of the feedback tools although with a high percentage of people owning at least one. Studies showed people need to feel control of using their tools, and this impact need to be recognized by oneself of others [5]. Adoption implies multiple phases in ongoing use. Kim identified two stages of how people adopt everyday technologies [64]: in the early stage, interest is the main motivation; then gradually the tool is adopted into daily routines. In a later stage, peoples practices with the tool become "rational reasoning rather than from an unconscious and habitual reiteration"; that is, using the tool becomes part of their routines. Peoples goals are mostly realized in the latter stage; however, the transition to this stage takes time. Whether the transition occurs at all depends on how easily the tool fits into the persons life.

Calendars are both a planning tool and a record of peoples daily activities, and thus may be useful in helping people understand data that are related to activities and locations (such as resource use, medical tracking, calorie counting, or spending). Calendar views have several advantages: the timeline aligns with the temporal data, time can be adjusted into different time scales to provide an overview and details as needed, and calendars represent the periodic nature of time-based data [109].Using a digital calendar is a common practice nowadays. The familiarity of the calendar views and the routine encountering should require low learning and managment cost, easily fitting in one's life routines. More importantaly, calendar events provide rich contextual information about one's daily life, which could be relevant to understanding personal data. Evidences showed that people inclined to check personal calendars for reasoing purpose in everday life [21, 72, 80, 87]. Thus, a personal digital calendar can provide appropriate contextual framing in helping people recognize and understand information patterns.

However, displaying addtional visualization layer on a calendar needs to be approached with care. When we mash-up information sources in a familiar tool, we need to ensure that the additional information is attentionally ambient to avoid interfering with the primary function of the application (digital calendar tasks in this case). Thus when adding data to an information space that may already be dense (the calen-

dar), the visualization design has to balance visual interference, with normal calendar tasks, with perceptibility and legibility of the added data stream. The design needs to minimize visual interference while supporting effective data perception. Meanwhile, reasoning about personal data and develping self knowledge is a long-term process, and using feedback tools in everday life is quite different, compared with using visualization tools in an occupational situation **??**. Especially, investigating how people use the feedback tool and reason about the data with the on-calendar visualization requires a real life context.

Thus, the study is divided in multiple phases: a systemactic literature review, a viability study to confirm the design goal and narrow down design alternatives, case studies to test the design concept and the implemented pototype, and a longitudinal field study to deeply investigate how the on-calendar feedback tool is used in everday life context. First, I reviewed the literatures of visualization designs used in everday life to identify design dimensions and challenges in personal visualization. I proposed the design approach and alternatives, aimed to tackle two challenges in literature: providing contextual informtion to enhance understanding and easy access to encourage on-going use. After that, I conducted a lab study to evaluate design alternatives with respect to the interference and the perceptability. With the selected visualization alternatives based on the study results, I implemented an on-calendar visualization linked Google Calendar with live personal data streams (from residence smart meter and Fitbit devices, respectively). The implementation was later improved based on two case studies: household electricity conservation and personal fitness. Finally, I deployed it as a fitness feedback tool in an eight-week field study, investigating how people interact with the on-calendar design and how they use context to reason about their Fibit data in the field.

These studies are designed to answer the following research questions:

- what are the current challenges in designing visualization used in everyday life
- are the on-calendar visulization comprehensable but not inteferring primary calendar tasks?
- what are the design options that can balance the interference and perceptabiity for the on-calendar visualization?
- To what extent can people use calendar events as context for reasoning about their personal data?

- could people improve their understanding of their data and behaviors better with the context provided by personal digital calendars?
- how do peopel react and use the on-calendar visualization as the feedback tool in everyday life?is the mash-up feedback tool easy to learn and manage in everyday life?

This thesis describes the following significant and novel contributions:

1. the first systemactic literature review of Personal Visualization and Personal Visual Analytics. It provides the starting point for researchers and designers to re-think about new design requirement and design dimensions of applying visualization design in everday life context.This work brings together research that was previously scattered in different deciplines and research fields.
2. a design approach that integrate visualization of personal data (personal fitness and home energy consumption) into people's existing information eco-system (a personal digital calendar). It can provide daily-life context for people to better understand their fitness data and home energy use. This approach incorperates people's existing habit of information use, requiring low learning and managing cost and engaging ongoing use.
3. defining *attentional ambience*, extending ambient visualization from spatial location (displayed in peripheral location) to attentional demand, with which ambient visualization can be centrally located as an addtional visual layer and the perception transiting between foreground and background depends on the degree of attention. This perspective changes "environmentally appropriate" to "attentionally appropriate" with respect to attentional demand.
4. quatitative lab experiments to evaluate visual interference and perceptability of design alternatives of on-calendar visualization. It showed the promise of the proposed mash-up approach that addtional visualization layers could be perceived ambiently with proper visualization choices.
5. qualitative field deployments to investigate how people react to the mash-up approach and use contextual information from personal calendars to reason about personal data in real life context. The studies confirmed that personal digital calendars could provide rich contextual information for people to reason

about their data. Particularly, it is helpful to identify and reason about data patterns and anomalies. Meanwhile, the on-calendar visualization made the revelant information easy to access.

6. a new feedback model of design compoments in feedback designs based on the study results. The model offers a structure to investigate connection of design componets in feedback design, and also indicate possible evaluation dimensions of feedback tools. It provides a starting point for thinking about how design characteristics might influence feedback tool use, including the likelihood of ongoing adoption and behavior change.

Although this work focuses on feedback scenarios, the on-calendar visualizations could be applied to many other types of data, at either individual or organization levels. For example, individuals might monitor physical activity, ratings of mood, or physiological indicators such as heart rate. Organizations might visualize network traffic or density of people present in a facility, and relate these to group activities or public events. Thus our investigation of on-calendar visualizations provides a foundation for exploring a potentially large design space of attentionally ambient visualizations.

In the following of this thesis, a systematical review across several fields is devoted first to identify design unique requirements, design dimensions and challenges of personal visualization & personal visual analytics (PV&PVA) used in everday life in Chapter 2. Among all the challenges, this work focuses on how to provide context for people to reason about personal data and how to make it easily fit into everday life routines. A review of related work from the literature is presented next in Chapter 3, in which I discuss examples of feedback design and the commonly used design strategies - persuasive design and ambient design, together with evaluation methods in visualization design and Human Computer Interaction (HCI). I propose the design approach and the visualization alternatives in Chapter 4 based on visualization design principles. After that, I summarize the the research method in this thesis to investigate the on-calendar deisgn approach (in Chapter **??**). First, a viability study were used to evaluate the design alternatives with repect to visual interference and perceptability, in Chapter 6. As the design concept was confirmed in the viability study, a working prototype (a web based application) was implemented with selected design alternatives from the viability study (in Chapter 7). To test the early vesion of implementation in real life context, I deployed the application in two case studies (home energy conservation and personal fitnesss) as pilot studies (Chapter 8), where

the application was connected with real time data (household smart meter and Fitbit devices respectively) and participants' Google Calendar. After the revised version of on-calendar visualization was implemented, I deployed it in an eight-week field study (in Chapter 9), aiming to investigate how people react to the integration approach and how they reason about their data with contextual information from personal calendars in everyday life. Some future work are then presented in Chapter 10, along with conclusion in Chapter 11.

(appendix) edit this part in the end ********** Supplementary information including a list of ...

****

Chapter 2

Personal Visualization and Personal Visual Analytics

Visualizing personally interested or related data have catched a great amount of attention in recently research and practice. However, pieces of this type of work scatter in several research comminities, for example, enironmental psychology, personal informatics, ambient visualization, information visualization and human computer interaction. There was no work before that provides systematic review of visualization work and bridge this research among these communities. Meanwhile, knowledging of the design patterns and gaps in previous practice is the first step before the design phase. Thus, a literature review of existing personal visualization and personal visual analytics work is necessary to identify design dimensions and challenges in this field.

In the following section I start with the overview of current research in this field. I then describe the review method and process. After summarizing the current design dimensions and patterns, I discuss the challenges the future reseach needs to face.

2.1 Background

We are surrounded by data in our everyday lives. Data relevant to our personal lives are increasingly available, often due to advances that make data easier to access or collect. These data cross a broad spectrum of domains and scope. For instance, due to Open Access policies, we now have immense data stores about our communities (e.g. census data); due to commercial availability of sensors, our health and fitness (e.g. exercise logs, pedometer data), and even our resource usage (e.g. utilities, such

as water, energy use) data is easily available to us. Some of this data collection is done without our explicit attention: our mobile devices, for instance, track our use of digital and social media to organize, plan, and connect with others; similarly, our web browsers collect digital traces of our online activities. Increasingly though, we are going out of our way to collect about ourselve (e.g.,the "Quantified self" movement). These data are relevant to our personal lives they enable us to explore information about ourselves, our communities, and issues that are personally relevant and important to us. As the data volume grows substantially, exploring and analyzing these data becomes a common daily demand, especially for people who are not data experts in their lives.

The value of data comes only after transforming into insights. Visualization is a powerful practice in this transfromation. However, research of this practice has been mostly focused in professional situations. As the interest of applying visualization in everday life to explore personal data or personal relavent data is growing, how to better visualization design to meet this fast growing need? What are the new requiremetns and challenges researchers and designers have to face with? Since the revelant research is currently scattering in several communities, to establish a common language of such visualization design I describe previous and future work as being part of a new field and research community called personal visualization and personal visual analytics.

Personal Visualization (PV) involves the design of interactive visual data representations for use in a personal context, and Personal Visual Analytics (PVA) is the science of analytical reasoning facilitated by visual representations used within a personal context.

The main question that PV&PVA is concerned with is: How can the power of visualization and visual analytics be made appropriate for use in personal contexts including for people who have little experience with data, visualization, or statistical reasoning? There is enormous potential for us to use data to make positive changes in our personal lives and the lives of others, but as visualization and visual analytics experts are well aware, greater availability of data does not on its own lead to new insights. Data must be accessible, understandable, and interpretable before interacting with it can lead to insights or actionable knowledge. Adoption of PV&PVA technologies also depends on how well those technologies fit into peoples daily environments and routines. PV&PVA builds on work in visualization (Vis) and visual analytics (VA) and aims to empower everyday users to develop insights within a personal con-

text. Personal context implies non-professional situations, in which people may have quite different motivations, priorities, role expectations, environments, or time and resource budgets as compared to professional situations. Because of these differences, PV&PVA designs necessarily have new requirements and challenges that bring new opportunities for Vis and VA research.

I reviewed existing PV&PVA literature across several fields to identify common approaches, findings and gaps. This work helps to establish an initial set of design dimensions to characterize this space and provide a common vocabulary that will make it easier to relate and share information between fields, which offers a starting point to learn about the challenges that arise when designing for personal contexts.

2.2 Review Method and Process

In this section, I describe the review method and process, with which a set of design dimensions characterizing the design space of PV&PVA research was articulated. These dimensions relate to data, context, interaction and insights (in Table **??**). These dimensions provides a common ground to identify cluters of work and also challenges in this area.

As part of our iterative process, I conducted an extensive literature survey of PV&PV A designs. First, I systematically identified relevant papers from TVCG (including VIS papers), CHI papers and notes, UBICOMP, INTERACT, AVI, EuroVis, and PacificVis from 2008-2012, plus CHI 2013 since those papers were available at the time. Relevance was based on meeting all of the following criteria:

- Data were about oneself, ones family or community, or relevant to personal interest or needs.
- Designs were intended to help people develop insights within a personal context (system design goal). Visualizations designed to support occupational roles or tasks were excluded (e.g., those specifically designed for domain experts or analysts in an occupational role).
- Designers intended the visualizations to be viewed from a non-professional perspective (e.g., from a self-centered viewpoint). Visualizations intended for use from an analysts perspective or that required professional training were excluded.

- Research focused on visualization and interaction design or their evaluation. Papers that focused on system architecture, system optimization or algorithms were excluded.

This resulted in a set of 59 papers that was complemented with a list with an additional seven papers encountered in other domain specific venues(see the summary in Table 2.1). The development of the taxonomy of PV&PV A design dimensions was based on an iterative, bottom-up approach. The objective was to identify and articulate a set of dimensions that could adequately describe and distinguish among PV&PVA tools in the literature. The dimensions evolved much like a process of open coding, where a set of dimensions and levels was iteratively mapped and adjusted with the literature. I considered this set of dimensions and levels to be complete when the collaborators reached consensus and could adequately represent the unique attributes of PV&PVA tools in the collection. The final set of dimensions are in Talbe 2.1.

To reduce redundancy while discussing the design dimensions (Table 2.1), when more than one paper was based on the same visualization tool this work refer only to their latest version of the tool. This reduced the set of 66 papers to 59 designs 2.2.

2.3 Design Dimensions and Research Interest to Date

With the collection of 59 designs,I coded and mapped them to an interactive parallel plot of design dimensions that helped to identify design trends and research interest to date. The most prevalent characteristics of existing PV&PVA tools can be observed by looking at the sizes of the boxes (representing levels within the dimensions) in Figure 2.3. As exploring the interactive parallel plot, I articulate these dimensions in the design space and discuss the research interest to date in this section.

2.3.1 Design Dimensions

From a data effort level, the most common situation was that data were either provided, requiring no effort (e.g., by web servers or system logs) or non-intrusively collected by sensors (with partial personal control). If data collection required no effort or people had little control over the data collection, the tools mostly had low

Table 2.1: Design dimensions, levels with examples from the literature

Categories	Dimensions	Definition	Levels	Examples
Data	Data Scope	Who the data is about	self	Sleep quality [7]
			family	Internet bandwidth shared at home [14]
			peers	Relationship with friends [37]
			community	Hang-out patterns on campus [75], online conversations[29]
	Data Effort	Amount of effort that is expended in data collection	none	Online search history [6]
			sensor	Nonobtrusive sensing devices, e.g, wearable sensors [31]
			manual	Manual logging pictures and annotations [62]
			mixed	Combination of sensor recording and manual input [58]
	Data Agency	The degree of control a person has over what data is collected, and when and how it is collected	no control	Online conversation logs [29]
			partial control	Users have control of whether or not to collect the data, but cannot customize what data they would like to collect [31]
			total control	Manually recorded childrens photos and growth progress [45]
Context	Design Context	Who designed and developed the application	self	Visualization designed by oneself
			group	Tools designed by a study group to chart their progress
			participatory	Using an online survey to get feedback on early visual design concepts [31]
			third party	Visualization of music listening history designed by the researcher [9]
	Settings	In what situation the tool is used and how it is used	personal	Personal laptop (mostly non-mobile but used by oneself) [65]
			domestic	Ambient display at home (mostly non-mobile)[32]
			mobile	Used on a mobile phone while on-the-go [31]
			shared	Visualization of physical activities viewed by co-workers [60]
			public	Visualization to promote energy conservation presented in a public space [41]
			mixed	Combination of above, e.g., visualization of residential energy on a personal computer and a mobile phone [5]
	Influence Context	Who the application is intended to inform	self	My physical condition [19]
			family	Childrens growth progress [45]
			community	Inform public about elections [84]
			mixed	Encourage water drinking for peers and oneself [16]
	Attentional Demand	How much attention is required to interact with the tool	low	Cell phone wall paper [7], ambient display [32]
			mixed	Pedometer counter on a cell phone and the historical data on a desktop [58]
			high	Exploration of music listening history with focused attention [9]
	Explorability	The ability to explore the data	low	Visual metaphor representing ones physical activities [31]
			mixed	Data overview on a cell phone and exploration of history data on a client service [5]
			high	Exploring historical data with dynamic queries [9]
Insight	Actionability	Degree to which the insight gained from using the tool can guide future actions	low	Does not inform further action, e.g., supports reminiscing about the past [68]
			mixed	Smartphone app that gives reminders about when to exercise and enables reflection on progress via historical records [58]
			high	Engages a certain behavior, e.g., encourage energy conservation [50]
	Automated Analysis	Data mining or other automated analysis methods are employed	Yes/No	Classify physical activities to types of transportation [31]

	VIS Community	HCI Community	Other	Total
Before 2008			2	2
2008	2	6	1	9
2009	1	7	0	8
2010	3	6	3	12
2011	2	8	0	10
2012	2	16	1	19
2013		6		6
Total	10	49	7	66

Table 2.2: Summary of surveyed papers

Figure 2.1: PV&PVA design dimensions (parallel axes) and surveyed tools (first axis). Box sizes indicate the number of tools with each classification. Linked highlighting enables cluster exploration.

actionability (34 out of 45 cases, seen by hovering the mouse over no effort or no control in the interactive visualization 2.3). The literature showed that people could achieve total control over data collection (what, how and when to record) only when they manually recorded or organized data; this occurred mostly in applications for personal health (9 out of 13 cases). Much less control was provided if data were collected through public channels such as social networks.

The practice of PV&PVA are mediated within personal context. In activity theory, Nardi [86] argued that context is "both internal, involving specific objects and goals and, at the same time, external to people, involving artifacts, other people, and specific settings". Internally, context could be "abstract artifacts" [60], such as goals, skill sets, preferences, experience, etc. Externally, context could be either physical constraints (e.g., physical environments or devices) or social influence (e.g., norms in a community or division of labor).In a personal context, people may look into their data with different goals, backgrounds, and expectations (i.e., internal context), highly influencing how they interact with the designs and what information and insights they could get from data.External factors that may characterize personal context include devices, use context and social influence. From the literature, most of the tools were intended to develop insights for ones family or oneself. We observed that nearly all

PV&PVA tools were designed by third parties (we reflect on this design perspective in section 6.5). However, the literature suggests that involving participants in the design process (participatory design) might be related to higher actionability (all 5 participatory designs achieved high actionability). Meanwhile, the tool set in our selection covers most use contexts: ambient displays at home, mobile devices on the go, personal computers or laptops used in a personal space, shared views with others, and displays for the public. It seems that applying mobile devices and shared views aimed to acheive higher actionability (14 out of 16 cases).

PV&PVA designs also covered a wide range of interactions, facilitating diverse attentional demands and explorability . Many of the tools, mostly with mobile devices or ambient displays, did not require focused attention (25 out of 59 cases).

From an insight perspective, not all PV&PVA designs were intended to reveal actionable knowledge (low actionability: 27 out of 59 cases). People also used these tools to satisfy their curiosity (e.g., exploring census data), to reminisce about experiences, or to share with others (e.g., exploring activity traces at home). Interestingly, although automated computational assistance (e.g., classification algorithms) is common in visual analytics generally, this type of analysis was not common in tools that we surveyed (14 out of 59 cases). Examples included sentiment analysis and classification of physical activities.

2.3.2 Research Interest to Date

Reviewing the design collection and exploring the parallel plot revealed the emerging interest in this field (What people have been working on) and possible gaps (i.e., research opportunities).Note that the clusters were not meant to be mutually exclusive or systematically categorize the design space; instead, they illustrated interesting relationships between design dimensions and highlighted some research trends to date.

Attentional Demand (high), Explorability (high), and Actionability (mostly low)

The first trend is designs for *enabling exploaration for curiosity* that requires high attentional demand and supported a high level of explorability. Insights obtained from using the tools were typically not very actionable, mostly was used to understandsomething rather than to support taking specific actions or making changes. Tools in this category were similar to traditional visualization tools, but usually had

a self-centered focus (my documents [9], my computer usage [14], places I have been to [59] or my finance [98]). These tools enabled user exploration facilitated by typical analytical tasks such as select, reconfigure, encode, elaborate, filter, connect, etc. For example, by interactively exploring (like traditional InfoVis techniques) with music listening history [21] people could investigate their listening patterns or re-experience a special life event in the past (as musical experiences are usually associated with events). Interaction techniques for tools in this cluster supported exploration (high explorability) that may help people narratively develop stories from their data. This might be the first phase of adoption .

For many tools in this cluster, personal knowledge and experience played an important role in the data interpretation process. For example, whether or not someone listens to music on a particular date depends on daily routines and special events [21]. Spending data can be explained by relevant routine activities, e.g., coffee drinking habits [98]. This implies that effectiveness of tools in this category could be dependent on highly personal factors. Yet most evaluations of tools in this cluster (12 out of 16) involved lab studies measuring task efficiency and error rate on experimentally controlled tasks with hard-coded use contexts. While such laboratory studies are common practice in VIS and VA research, they have limitations for evaluating PV&PVA applications.

Attentional Demand (low), Explorability (low), and Actionability (mostly high)

Another common design trend is to *provide in-the-moment or on-going awareness with respect to personal behaviors*. This practice was mostly applied in personal health or energy conservation, where they were expected to avoid interrupting life routines by combining low attentional demand and just-sufficient salience, for example, through a strategy of ambience (e.g., cell phone wallpaper). For example, through cell phone wallpaper, ShutEye indicated sleep-related activity [19]. Interactions with tools in this cluster tended to be simple to fit in the on-the-go or ambient context and to efficiently provide key information as needed.

Some tools in this cluster used machine learning or data mining algorithms to assist with data aggregation or disaggregation, e.g., classifying accelerometer data into physical activities [30] or disaggregating water consumption based on water use behaviors [48]. Some used graphical metaphors to remind people of the potential

impact of their behavior. For example, to encourage people to take exercises and green transportation [46], a polar bear on a piece of ice was displayed; the ice began to melt if the users behavior was not environmentally friendly. This example also reveals the special PV&PVA requirements in terms of aesthetics and emotional engagement.

Social influence was often used as a persuasive strategy to engage behavior change (e.g., for drinking more water [27], staying physically active [77] or encouraging recycling [106]). However, social engagement and comparison may also raise other problems: inappropriate social strategies actually made the design less effective or caused undue stress [106], and viewing other peoples personal data also brings up privacy concerns.

Data scope (family), influence context (family), and setting (domestic)

Systems designed for families focused on *data about family members or the home environment, and were used or deployed domestically.* Some applications used decorative ambient displays to make the technology less intrusive and better fit in the home environment [48]; others ran on a personal computer, enabling close exploration and organization of family data to track progress [62]. These applications were designed to monitor or engage behaviors towords family health, energy conservation, domestic resource sharing, or social interaction of family members.

In many cases, visualization designs consider individual differences among family members, for example, customizing views to adapt to difference cognitive levels (children v.s. adults) in th family [48]. Additional contextual knowledge was also provided in visualization to help people interpret the data, for example, by naratively depicting quantitative measures [48], which facilitate a better understanding of the family data. Meanwhile,interaction and sharing within families can bring up issues of competition, cooperation and privacy. For example, a visualization of Internet traffic [26] was designed to educate family members about their shared Internet usage. Family members could view each others online activities and bandwidth usage could be prioritized with respect to social roles. Here, some family members noted an unwelcome intrusion on privacy. The callenage is how to balance the diversity of users in a family, with respect to cognitive capabilities, skills, and social roles.

Data scope (community), data effort (none), data agency (no control), and influence context (community)

Beyond individuals and their families, research interest also revealed that people are also curious and care about the *communities* they live with. These designs were usually intended to inform the public or a certain social group, e.g., raising public awareness of elections [110], supporting easy exploration of survey data [36], or revealing topics evolving from social networks [35, 42]. In a few examples, they were also used to encourage behaviors valued within the community, for example, ambient displays deployed in a department lobby to encourage energy conservation and physical activities[53].

Tools in this cluster mostly supported focused data exploration tasks, similar to other Vis and VA applications; they employed many traditional visualization techniques to facilitate deep analysis, usually requiring high attentional demand. In several cases, automated computational analysis was used for mining large data sets from social networks (4 out of 11), e.g., peak-finding [79] and sentiment analysis [35]. Traditional Vis and VA techniques may work well to support reflection on community data. However, since public data may not be too personally relevant, such tools may benefit from employing additional engagement strategies or novice interaction techniques to enhance interpretation. Examples include supporting exploration from different perspectives to capture relevant context [99] and employing non-traditional representations to compensate for the limited analytics skills of non-experts [36].

2.4 Design Challenges in PV&PVA

PV&PVA brings forth a set of new design and research challenges because of the unique nature of personal context (e.g., role expectations, environments, and related activities). For example, PV&PVA systems may need to support people with limited visualization literacy and analytics experience, fit into personal life routines and physical surroundings, support fleeting and short term use, support recall of relevant events, and apply appropriate baselines to support reasoning about data. While some of these challenges are not completely new , PV&PV A introduces a unique perspective on these challenges, and emphasizes their importance. In this section, I articulate the key challenges based on the literature review.

2.4.1 Fit in Personal Routines and Environments

Any tool needs to be designed to fit within its physical environment and context of use. In a personal context, physical environments and activity routines can be quite different from those in professional contexts, leading to new design challenges. For example, we may wish to support fleeting use of a fitness tracking application without interrupting ones life routines, or customize a visualizations appearance so that it matches the aesthetic of a living room where it will be deployed.

Fitting into peoples lives means that designers should consider availability, ease of access and ease of use for long-term adoption. Kim identified two stages of how people adopt everyday technologies [65]: in the early stage, interest is the main motivation; then gradually the tool is adopted into daily routines. In a later stage, peoples practices with the tool become rational reasoning rather than from an unconscious and habitual reiteration; that is, using the tool becomes part of their routines. Peoples goals are mostly realized in the latter stage; however, the transition to this stage takes time. Furthermore, whether the transition occurs highly depends on how easily the tool fits into the persons life.

There are many barriers that limit the adoption of PV&PVA tools. One way to reduce these barriers is to consider the context of use; for example, designers can reduce the effort required to collect, organize and access data, so tools can be used with minimal effort or at-a-glance. Visualization designs can be integrated with tools or devices that people use or encounter regularly in their daily routines, in line with one's existing information use habits. For instance, a visualization integrated into mobile phone wallpaper would be frequently encountered as people use their phones.

Aesthetics of a PV&PVA tool (how it looks, how it is to be used, even its physical manifestation) must suit not only personal taste, but also its place context. Most notably, ambient visualizations that will be integrated into peoples environments, especially their homes, present additional design challenges. Such displays may need to emphasize visual appeal and customizability as well.

2.4.2 Recall of Relevant Context for Reasoning

A challenge in PV&PVA is that the appropriate context for interpreting the primary data may not be in the form of data that is easily accessible. Activity theory [6] has recognized that peoples understanding and use of information artifacts are strongly influenced by the context (experience, preferences, competencies, values, etc.) Rele-

vant context for interpreting data in a PV&PVA tool might be the knowledge of ones own past activities, feelings, and interactions with others. For example, Understanding temporal patterns of household energy use may be difficult without knowing what one was doing at certain times of the day.

Some of this necessary context is in the form of memories that are recalled to explain past behaviors. Lee and Dey conducted a study with older people on pill-taking [72]. Participants tended to explain anomalies of pill taking (i.e., forgetting to take pills on time) with routines and their subtle variations, mostly by digging into their memories. But memory is fallible and imprecise, particularly for older people in this case. Adding additional data from other sources (e.g., with help from context-aware technologies) may help to trigger peoples memory and enable them to better make sense of the primary data. In the same example, those seniors many times referred to the events marked on their wall calendars for relevant information that can explain the anomalies.

Overall, relevant context can relate to individual differences, personal experiences, view perspectives, and social encounters. One challenge is that the appropriate context may vary for different people and in different situations. Identifying types of contextual data that will be more generically useful, and devising flexible mechanisms to enable people to recall or recognize contextual data that they consider relevant, may help to enrich the inferential knowledge that people bring when using PV&PVA tools, supporting richer insights.

2.4.3 Defining Appropriate Baselines

Making comparisons is a fundamental way to gain insights from data, and this is equally true for PV&PVA applications. For example, parents could compare their childrens development to milestones provided by a pediatrician [62], family members could compare their water uage with each other or among different rooms [48]. People could learn about nutrition from a national food guide. In other words, people often need a reference (or baseline) to understand and assess their current situation.

But what baseline should be used for comparison? One challenge is to understand what makes an appropriate comparison set. Should a persons energy usage data be compared to their prior usage levels? Should it be compared to a national average? Should it be compared to their peers data or data from demographically equivalent people? What does demographically equivalent mean? Appropriate baseline is an

elusive idea, mainly because it depends so heavily on the context of use, goals, and also on each persons values. For instance, many people may be interested in leading healthy lives. Yet, what constitutes healthy may differfor one person, it may be the absence of stress; for another, whether he is sleeping well; for another, her adherence to a national food guide. It is unlikely that we could define a single baseline to satisfy all these goals and values. Moreover, the appropriate baseline is likely to change along with the questions the person is trying to answer. PV&PVA designs might need to make people aware of the variety and varying nature of baselines, and also provide flexibility for a person to choose and adjust baselines depending on their own situation.

2.4.4 Sharing and Privacy

Sharing experiences and spaces with others (family, friends, social groups, etc.) is an important aspect of everyday life. Already there are many PV&PVA tools with an influence context beyond the self. Examples include tools for sharing memories and experiences among family members or friends [91, 107]. One intriguing space is to apply social interactions to enhance motivation or persuade behavior change, for example, setting group goals [77], comparing your own progress to others [27], or even interfering with social surveillance [106]. However, this approach should be applied carefully, since social interactions may also evoke negative emotions such as stress or guilt. Moreover, because sharing may enable people to see each others data (e.g., when using data from peers or the neighborhood as a baseline), privacy must be considered.

For displays of personal data (data about oneself), people may desire even more privacy. In some situations one may actually want to have a display that cannot be easily interpreted by everyone; it may be important to deliberately design visualizations that are incomprehensible to everyone but the owner. Such designs may be particularly important when personal interest is intrinsic and where privacy may be a concern. In such situations, highly personal-data encodings may be an essential design feature. One example is UbiFit, which provided a view of ones physical activities over the past week on a mobile phone with an abstract visualization of flowers in a garden, making the data difficult to read by any other person. This kind of approach is important, since our personal data may be in public view (here on a mobile phone, but perhaps alternatively as an ambient display), and we may want to be selective

about to whom we reveal the meaning of the display. The possible focus on visualization that is both revealing and insightful to a single viewer and concealing or at least neutral to others is a design approach that has not previously been considered in Vis or VA.

2.4.5 Evaluation

Evaluation of visualization and VA tools has been an ongoing research discussion for several years. PV&PVA is no exception, and in fact, presents some unique challenges for evaluation. Designers often aim for PV&PVA tools to integrate seamlessly into peoples life routines, physical environments, and social situations; these contexts of use would be very difficult to simulate in a controlled lab study. Moreover, researchers also need to re-consider the metrics that are typically used to assess VA or Vis systems. Time, error, and insights are not the only relevant metrics for evaluating PV&PVA tools, and often may not be the most important ones.

Ease as a conceptual metric could be used as one basis for evaluating PV&PVA tools. That is, how easily does the tool fit into ones daily life, habits, and routine? Can one *ease* into the use of the tool without effortfully breaking from ones current activities? Can one easily answer the questions they might have of their dataset? Can one easily interpret and understand a visual presentation? Can one easily grow with the tool, moving towards more sophisticated analysis as they gain experience? This concept of ease goes far beyond the traditional "ease of use" metric. While ease of use is one relevant aspect, the concept of ease in PV&PVA goes much more broadly. Ease can be considered analogous to "comfort": whether a tool fits comfortably into peoples environments, routines, habits, and social experiences and how this comfort level evolve and adapt over time. Obviously, the flip side of ease is unease: what are the barriers to ongoing use? Yet, only a few studies have addressed this adoption issue [49, 70]. Dedicated applications had to face the low usage issue [16, 61, 39, 108]. How to encourage ongoing use is a critical factor in PV&PV A research.

While operationalizing this concept of "ease" is challenging, it should be clear that conventional metrics used to evaluate visualization tools (i.e., task completion time, task errors, and even insights) are not only insufficient, they may be the wrong metrics to use altogether for many scenarios. One unique characteristic of PV&PVA tools is that they may be used to fill the gaps in time when one is bored, curious, or doing something else [107]. In contrast, our canonical view of VA tool use is one of a

focused information worker actively seeking information or insights. While someone using a PV&PVA tool might be focused on discovering complex insights (e.g., tracking health symptoms), they might be equally likely to use it for purposes such as fun or awareness. Appropriate evaluation methods and metrics for assessing PV&PVA tools are urgently needed to support future research.

2.5 Contribution and Limitation

PV&PVA brings unique design requirements because in everyday life, data interpretation and insight development are mediated by personal context, including environments, settings, personal experiences, skill sets, prior knowledge, and social influences. It offers the chance to identify new challenges in visualization design used in everday life and a taxonomy of design dimensions to provide a coherent vocabulary for discussing Personal Visualization and Personal Visual Analytics. particularly, this on-calendar design approach proposed in this thesis mostly aims to tackle two of challenges: fitting in personal routines and providing relevant context for reasoning. It indicates directions to investigate the ongoing issue in visualization design of feedback applications.

I am also aware that the taxonomy is based on the literature data and intended as a starting point to foster PV&PVA research that will evolve and expand as PV&PVA becomes established as a field and research. I would also like to acknowledge the rapid growth in industry practice related to PV&PVA, most notably the Quantified Self movement. Wearable devices are widely used to collect daily activities in an unintrusive way, e.g., Fitbit, Nike Fuelband, Jawbone, etc. Mobile apps that connect to sensors embedded in smartphones provide similar functionality. Applications such as open.sen.se and Exist help individuals integrate and manage data from all their devices and apps. The substantial industry interest in PV&PVA sugggests promise for future growth in this field. As well, these industry tools share common ground with academic work. However, for tractability, the comprehensive review focuses on the academic literature.

Chapter 3

Related Work

This thesis focues on perosnal visualization used in everday life, specifically used as feedback tools to learn behavior choices. With the design requirements and challenges identified in the systematic literature review, this chapter presents the design examples, strategies and evaluation methods in previous work related to this work. A selection of design examples were analyzed first, followed by the discussion of the common design strategies in this domain - persuasive design and ambient visualization. Meanwhile, I also discuss about evaluation methods that have been typically used in visualization deisgn and HCI research that inspired the study design in this work.

3.1 Feedback Design

The definition of feedback is dated back from behavior science in organizaiton learning and management theory [94, 12, 56]. Ramaprasad's definition, "information about the gap between the actual level and the reference leval" [94], points to the three key components in feedback: current status,reference and gap. Herold and Greller also claim that feedback information would also reflect behaior influence, indicating appropriate behaviors with respect to the reference (e.g., a goal) and how well these behaviors have been excuted [56]. Later, the concept of feedback was introduced in mechanical systems, where it is used to control and adjust the system behavior by monitoring the output and feeding it back to the system. Current studies of feedback systems in human computer interactin probably extended from environmental psychology where people are assumed to lack awareness and understanding their behaviors in everyday

life could lead them towards sustainable living. [47].

In feedback systems information were typically provided as antecedent or consequence interventions to affect behaviors[4]. For example, goal setting, commitment or public media campaign is typical antecedent intervention, and feedback or rewards is popular consequence interventions. Specifically, feedback, also known as eco-feedback in this application domain, is amied to show people the outcome of their behaviors and what are their behavior choices.

However, different from environmental psychology, resarchers in HCI are more interested in designing and engineering interactive feedback systems, exploring design approaches, and experimenting how to apply designs to help people [34, 90, 16, 47, 100, 103]. Disolvo et al. reviewed work in sustainable HCI and analyzed the genres and scope of research topics in this merging field [34], offering a multi-facet perspective to re-think about emerging issues in sustainable HCI. Moreover, Pierce conducted a review focusing on Electricity Consumption Feedback(ECF) [90], in which they outlined the research in an even broader scope beyond interactions. They also suggested energy-related HCI research to inspect design strategies that increase awarenes and engage individuals in practice rather than narrowlly focusing on behavior change.

On the design perspective, Froehlich suggested design dimensions of feedback tools:frequency,measurement unit, data granularity, push/pull, presentation medium, location, visual design, recommending action, comparison, and social sharing. [45]. For example, the frequency of updating feedback matters: the more frequent feedback is given, the more effective it influences behaviors [4].

Related practice and research of personal visualization in feedback design are common in two application domains: energy conservation and personal fitness. This thesis focused on these two fiels as example to investigate the on-calendar design approach. In the next section, I present examples of feedback design in these domains with respect to design dimenisons of PV&PVA.

3.1.1 Feedback Applications

Feedback tools are commonly used in home conservation and personal fitness. From data perspective, they are mostly providing information about individuals (e.g., one's transportation behaviors [28], sitting postures [50], physical activities [76, 5], water drinking habit [27]), families and home (e.g., electricity use [95, 97], water use [39, 48], air quality [64]), aimed to engage responsible or healthy behavior change. This

approach was taken to benefit communities as well in some cases. For example, co-workers [77], or general public [11, 53].

As the resualt, the use situation varies. Kjeldskov [68] developed a mobile phone applicaiton, Power Advisor, enabling users to watch on the electricity use on one's smart phone. Froehlich designed and built an ambient feedback display [48] installed in particpants' home to educate them about water usage shared by family members. Schwart et al. designed the feedback to display on household's TV, so users can review their electricity consumption during comertical period. Similarly,such energy use feedback display could be used in public buildings for the building occupants and visitors [11, 53]. Meanwhile, sensor-based warable devices and mobile phones are commonly applied in this area, enabling to collect or show data in situation [28, 46, 19].

Feedback designs also support rich interaction. In-the-moment tracking requires low attention, and engages people to take an action in the situation. Upstream [71] indicated the water usage during one's shower. Similary, waterBot [10] showed users the water usage over the sink. Rogers [95] applied artistic visualization to show electricity intensity when the appliance is in use. A real-time updatd visualization can alert an inapropriate sitting posture when it happens[50]. As smart phones are getting popular, visualizatio of feedback has been integrated into the phone use [28, 46, 19], for example, as the phone wallpaper. It increases the chance people encounter their data with minimal attention and effort. On the other side, people are also allowed to explore historical data to learn patterns, causes and influence of their behavior decisions. Replacing the paper bill, Dubuque [39] was a web porrtal, where users can review data patterns of water consumption. Li proposed a design that incorperate people's geolocation and social interaction while displaying pedeometer data [76], hoping to improve people's interpretation of their fitness level. More interestingly, such reflectin could also supported by artistic style of visualization [41].

One of the common design component to develop insights and knowlege through the interaction is to support comparisons, with peers, families, neighboors or oneself. Comparing with similar household in the neighborhood may help people have better understanding of their energy consumption level [39, 84]. Many designs employed social comparison to engage behavior change, in which one can share and compare progress with friends, co-workers or peers [88, 20, 77, 27]. Lin introduced group comprison to engage people in physical activities [77]. VERA [20] enables users to share photos related to their healthy decisions (e.g., eat healthy, take exercises),

hoping to enhance the social awareness that would persuade one towards healthy life.

Meanwhile, computer assited analysis has been commonly use to enhance the insight development, especially in large amounts of data where patterns are not easily recognizable visually. For example,Ubifit, by classifying ones movement data, is able to inform users both the amount and types of exercise one has done [28]. Similiarly, transportation behaviors could be recognized by programatic classifiers [46]. Water usage could be automatically categorized and aggregated by location of water use behaviors (e.g., kitchen v.s. washroom). However, automated techniques inevitably have flaws (disscussion in Section 3.2).

Among the examples in previous work, two of the common design strategies are persuasive design and ambient visualization.

3.2 persuasive design

In many feedback designs, persuasion is the mostly used design strategy. Persuasion technology usually refers to Foggs model [43] that described strategies that people would employ to change behaviors or attitudes, even without their understanding of the problem or behavior, e.g., by computer generated suggestion [7]. Meanwhile, behavioral models were brought in from environmental psychology to study behavior choices in feedback designs. Mostly these methods are based on rational choice or norm activated [4, 47, 103, 54]. The former method assumes that people take an action (e.g, towards healthy life or proenvironmental living) based on rationally purposive plans by evaluating personal gain and cost. For example, people save energy for financial purpose. On the other hand, norm activated theory consider the social and culture environment people are situated in. For example, immediate feedback is used to engage behavior change in the moment (e.g., upstream [71] provides immediate feedback on water use to encourage people to take shorter showers), even without poeple fully understand the behaivor itself.

The persuasive technology approach has also been criticized [23, 102, 111] . Strengers et al. questioned the value of in-home display approaches [102]. They found that household behavior change related to energy consumption cannot be modeled by one or two variables; instead, it is mediated in the context of everyday life, socially, culturally and institutionally. Thus, rather than engaging immediate action, it may be more important to encourage people to reflect deeply on the problem according to their own situations. Brynjarsdottir et al. similarly argued that sustainable behaviors

are not isolated, rather they are situated in the social-cultural environment [23]. They questioned modernist approaches that human behaviors can be predictable because the essential aspects of life are truly reflected by calculable and formal measures that are captured with systems. However, the limited aspects captured with systems fails to consider the complexity of reality and limit the focus on aspects that can be clearly measurable, narrowing peoples vision. Thus, they suggested a shift from prescription to reflection. Baumer [20] suggested an approach of "open-ended" awareness because people have varying definitions and assessment of "being healthy". The findings of the study [103] suggested that eco-feedback research need to help people to understand technology ethnographically, identifying the origions, courses and influence of their behaviors with "contingent context". For example, Dubuque is an open-ended system that engages people to reflect about household energy consumption within their own lives [40]. The study showed that the reflection oriented design helped the majority of participants increase their understanding of water consumption and also encouraged social conversation about water conservation. Similarly, in my design, I focus on the role of visualization in supporting reflection instead of behavior coercion.

3.3 Ambient Visualizatoin

The goal of the on-calendar design in this theis is to publish information in an unobstrusive way in one's existing information use eco-system (personal digital calendar in this case). It requires that the visualization must be ambient and faded into the calendar frame most of the time, although it is always displayed on the background. Kim et. al discussed about design requirements for ambient display, and its advantage, comparing with direct information presentation, of providing feedback of personal activities in engaging sustainable life [66]. Their analysis indicated the essential characteristics of ambient display in feedback design: subtle and non-interferring with primary tasks (e.g., time management tasks with digital calendar in this work).

With a systematic survey of reearch systems using ambient visualization [92], Pousman and Stasko defined ambient information systems by several criteria: important but not critical information; they can move from the periphery to the focus of attention and back again; use subtle changes to reflect updates in information (should not be distracting); and are aesthetically pleasing and environmentally appropriate. For example, an artistic splash display over the sink as part of one's kitchen can indicate energy usage in the hosue [18]. As the decoration at home, an aesthetically

tableau machine indicated the occupants' trace at home [91]. Ambient displays also could use physical objects as visualization artifact, e.g., to show energy consumption of a public building [53]. Other than displays embeded in physical environemt, designers also publish visualization in digital environment (e.g., one's mobile devices), hoping to enhance one's awareness of the data as it would be often encountered. Ubigreen encoded fitness data with story-telling graphics and was implemented as the background paper of mobile phones [46]. Antoher example is ShutEye [19], with visualization of sleep-related data as the wallpaper of one's phone.

In this thesis, I extend this notion of "ambient" from that of spatial location to attentional demand; in other words, an ambient visualization need not be physically located in the periphery, it can be centrally located as a secondary visual layer that is not visually demanding. That is, the information encoded with the secondary visual layer could be attentionally "centralized" and "de-centralized" by slightly changing the degree of attention as the task requires. This perspective changes "environmentally appropriate" to "attentionally appropriate" with respect to the representation of the primary task; in this case, typical tasks with a digital calendar.

3.4 Context use in Feedback Design

Providing relevant context for in feedback design could help to develop insights and enhance the reasoning (Chapter 2).Contextual information is a critical factor in helping people recognize and understand information patterns [16, 33, 103]. Strengers [103] also suggested that energy use was mediated by everyday life context ("social, cultural, technical and institutional dynamics"), with which people can better understand their behavior and evaluate their alternatives. The study of Ahtinen [5] showed the importance of "personally relavent" contextual information for people to understand their exercise logs. Without sufficient context to understand the data, it can be very difficult to make productive behavior changes. Lack of personally related context for relecting on one's data can also prevent further use of the feedback tools [63].

Location has been often used in design as contexual information. Li [76] and Nadalutti [85] used geo location as context to help people understand their fitness exercises. Personal Knowledge of water unit could be improved by narrative graphics [48]. Placing data in social context could also convey a better picture, e.g., compare water consumption among family members[48], or with neighbors [40]. Another way

of providing context in visualization is through linking related data, e.g., at-home activities and household energy consumption. Elleg[top]ard [38] display both occupants' activities at home and electrisity usage together along the timeine to indicate the cause of usage patterns.

Meanwhile, a calendar has its advatage of providing context: the timeline aligns with the temporal data, flexibility of data granularity by switching time scale (day, week, month), and calendars represent the periodic nature of time-based data. Van Wijk revealed the patterns and the possible causes of energy use in the work place, by clustering the consumption data within a calendar [109] that indicated employees' work patterns through the weeks and holidays. AffectAura visualized an employee's affect data in a calendar-like view (only week view was presented in the visualization, since the study only lasted for a few workdays.), together with her other work related data[80] (e.g., work schedules, emails,documents used at work), aimed to help employees reflect on their emotional level and productivity over time. Events in one's personal calendars may capture part of daily activities for later recall. In the study of the pill-taking reminding system [72], seniors inclined to look up events on their wall calendars when they were ask to reason about situations that they forgot to take pills.

3.5 Evaluation of Personal Visualization in Everyday Context

Feedback applications have been studied extensively, most often in lab evaluation and field studies. As Freohlich point out, HCI researchers usually have different methodology compared with those in environmental psychology [47, 4]. HCI research primarily focuses on feedback artifact itself more than effect of behaviors, because of its motivation for iterative design. Thus, they ended up with quick feedback from lab evaluation, and some of them employed field studies but in a small scale and short duration [47]. The natural of HCI evaluation is to inform designs, which makes it unrealistic and unnecessary to to apply the method of Randomized Controlled Trail that is commonly used in environmental psychology and medical informatic systems.

Evaluation is one of the major challenge of personal visualization designs (see section 2.4.5). Such visualization tools are situated in everday life, and largely influenced by personal context, making it difficult to simulate use scenarios in a controlled

lab. Thus, ease and understandability could be used as criteria in evaluating personal visualization tools in everday life: how people react, understand and use the artfifact with personally relevent context in one's life; how easy these artifact fit into one's life routines, physically, socially and cultrually.

In recent feedback design, researchers showed increasing interest of field study as the evaluation method. Many of these studies maily focused on evaluating specific systems or designs, with measuring behavior change or its outcome in a short duration, for example, emphasizing mostly the fluence of decreasing energy usage or increase exercises [71, 106, 28, 77, 27]. However, measuring behavior change might be inaccurate and unnecessary [69]. The behavior change involves many factors and usually takes a long-term ongoing process. With the typical study scale in HCI research, it is difficult to make the claim that participants' behaviors have truly changed. Instead, it would be more helpful to understand how and why the system or tool works. Such information could imply future design, especially for novel technologies at an early stage [103, 89, 32].

Considering that, the goal of this thesis is to understand how the easy access to contextual information (specifically, events on a digital calendar) can help people to understand their behavior data. Most similar to this work is the study by Li et al. [76] that demonstrated that contextual information can help people increase awareness of their data and Fritz et al. [44] that investigated long-term persuasive imapct of fitness trackers like Fitbit. However, unlike their work, my interest is to investigate how people react to the idea of integrating feedback data on their personal calendar, and how personal calendars can provide rich context for people to reason about their data.

Chapter 4

Visualization Design

What visual encodings could fit the on-calendar visualization, representing time-varying quantitative feedback data and creating an attenionally ambient layer on a calendar? The challenges are to make the visualization salient enough without interfering with the digital calendar use and meanwhile data patterns are perceiptible and comprehensible. In this chapter, I discribe the design alternatives that may fit these requirements based on the visualization design principles in literature. Their visual interference and perceiptibility will be evaluated in later experiments.

Common representations are line graph, bar chart, stacked areas and saturation. More complex visualizations are Gantt charts, horizon graphs, spirals and 3D visualizations. However, previous work showed that visualization novices typically prefer visualizations they are familiar with [49]. Meanwhile, effectiveness of different visual encodings have been ranked in previous empirical studies and perception theories [78]. Visual encoding of spatial position dominantly reflects the users mental model of quantitative data (i.e. people are accustomed to line charts), but the graph ratio could possibly influence its perceptibility within the limited space of calendar views. By contrast, luminance can be conveyed within a relatively small (but not tiny) area, so it could be an alternative to the line graph. Another reason for considering luminance is related to its effectiveness for visual aggregation. Individuals may need to estimate and compare summaries of their data across days or weeks. Although color (luminance in this case) is less accurate than position for estimating single values, it can be more quickly visually averaged to convey a summary message as compared to position (i.e., line graphs) [31]. Based on that, I consider line graph, colored-region (line graph with shaded color), and luminance (I avoided hue to prevent possible color interference effects, since calendar activity entries are usually colored) would be the

most appropriate for quantitative data to encode quantitative data in the on-calendar visualization.

As showed in Figure 4.1 and Figure 4.2, temporal data (e.g., of energy use or personal fitness) is visually encoded by line graph, luminance and colored-region (a combined case of position and lumninance). For line graph (Figure 4.1 top), temporal data is visually encoded by position along one axis, and time is represented on the other axis. Colored-region (Figure 4.1 middle) is an alternative version of line graph with filled solid color in the contour. Luminance (Figure 4.1 bottom) encodes data by gray scale. (hue is avoided to prevent possible color interference effects, since calendar activity entries are usually colored.)

The three visualizations were designed respectively for the typical calendar views: day, week and month (Figure 4.1 and Figure 4.2). Day view is an identical visual encoding to week view. In month view with luminance, only daily average values were encoded due to the constraint of showing luminance in a small region. In contrast, line chart and colored-region always showed continuous data in all calendar views.

Considering that the overlapped visualization that shares the same space with calendar events and might cause visual interference, another layout alternative is side-by-side display (Figure 4.2) with an addtional band paralell with calendar cells, aimed to preserve effective graphical perception while minimizing interference with normal calendar tasks. The width of the visualization band (side-by-side condition) was made smaller than the calendar cell, because the visualizations should not visually dominate. To enhance the visibility of luminance in overlapped layout, the width of activities (green and orange boxes) is squeezed a little bit compared to other representations to leave a small gap from the boundary (Figure 4.1 bottom).

The early design alternatives were prototyped respective to three criteria: display location (overlaped or side-by-side), visual type (line graph, line graph with shaded color, luminance) and calendar scale (day, week and month views). However, their viability of supporting attentional ambient in the on-calendar design need to be evaluated and confirmed first (Chapter 6) before they are applied in the field (Chapter 9).

Figure 4.1: Design alternatives displayed overlaped (top: line graph; middle: colored region; bottom: luminance)

Figure 4.2: Design alternatives displayed side by side (top: line graph; middle: colored region; bottom: luminance)

Chapter 5

Research Methods

Providing context for reasoning and fitting in everyday life with ease effort were identified as main challanges in personal visualization design (Section 2.4). With respect to that, this thesis explores the design concept of integrating personal visualization into one's exisisting information use eco-system. Specifically, I investigates this concept in the example of integrating visualization of personal feedback data (data about home energy consumption and personal fitness) into one's personal digital calendar, hoping to provide better context to help people interpret their data and lower the barriers involved in accessing these data. This goal leads to the following research questions that were investigated in the rest of this work.

1. Are the design alternatives comprehensable but not inteferring primary calendar tasks?

2. How do peopel react and use the on-calendar visualization as the feedback tool in everyday life?

3. To what extent can people use calendar events as context for reasoning about their personal data?

4. Could providing addtional context in visualization improve people's understanding of their data?

Note that behavior models and their inflence on behavior change are out of the scope of this thesis, and the effect of behaivor change is not the assessment of the on-calenar approach in this work. The general philosophy centers around helping people understand their behavior with operational understanding [16, 103] rather

than digitally lecturing them into behavior change. We view understanding as a first (but not sufficient) step towards encouraging new habits or taking actions; therefore, tools that can succeed at this step have made progress.

5.1 Research Methods

On the basis of the research questions above, I mainly applied four research approches in this work: lab experiment, field study, interview and survey. The analysis methods include quantitative statistic analysis and quanlitative open coding based on ground theory. Other than that, two case studies were conducted as pilot studies for the later field study. The research methods here were considered according to the research phases, contextual factors and related work in previous research. I will discuss these methods with their benefits and limitation in the following sections, and the further detail are in subsequent chapters.

5.1.1 Lab Experiment

Lab experiment has been commonly used in evaluating information visualization design, in which the often used metrics are completion time, task errors, and even insights [96]. The first phase of this research is to evaluate erceptibility and legibility of the design choices (Chapter 6) in facilitating the attentional ambience; that is, visualization on a calendar is perceptable but not interfering with normal calendar use. This is typical perception evaluation in informtion visualization systems, for example, to evaluate effectiveness of visual channels [78], perception of new visualization techniques [55], or visual perception interferred with background grid [17].

Online experiment mayb also meet the requirement [24]. However, it may not able to provide contextual information about participants, e.g., through observation. Amazon Mechanical Turk [1] is also out of the choices due to the ethics protocol.

One concern is that visualization mock-ups (rather than interactive application) were used in the experiment (see Appendix **??**), limiting the interaction of participants with the design. Data encoded in the iamges were "hard corded" (see Seciton 6.2) may not reflect real life data patterns. However, the data pattern and the tasks designed for the tests suit the experiment in perception and interference evaluation, helping to verify the design concept in the early stage.

[1] https://www.mturk.com/mturk/welcome

5.1.2 Field Study

Personal visualization has been oftern investigated in field study, particularly in HCI community **??** and most studeis last from one week to one month. With the working prototype, my interest is to investigate how people react to the on-calendar design in a real life context. Lab experiment is obviously not suitable in this case (see Section 2.4.5). Engaging with personal visualization is an ongoing process; it does not require participatns' continuous focused attention all the time like usability tasks. The interaction and reasoing process are embeded in everyday life context that is impossible to simulate in a labl environment. With contextual interview only, the ongoing nature would be impossible to investigate. Randomized Controlled Trail is not proper choice for this research neither, since the research questions are open-ended questions about participants experience rather than veryfing (or comparing) design solutions (Section 2.4.5).

This method also has its limitation. The small scale of participants may not enough to represent a statistical effect. The way people engage with the on-calendar design may vary in different application domains (the analysis focused on personl fitness in this thesis). Thus, I employ a qualitative method to analyse data from the field study.

5.1.3 Survey

Survey questionnaire is an unintrisive method to collect data through field study, compared with interviews. Particularly, self-report surveys for assessment of behvior in natural settings have been widely used [57], especially in personal informatic systems [74]. The short questionnaire was used as post-study questionnaire in lab experiments to collect participants' general experience referring to intererence, perception and aesthetics. The weekly survey, a standard instrument to evaluate people's physical activities, was used in field study to evaluate their weekly physical activities. These data served as supplement for physical activit data collected by wearable devices. However, self-report survey may bring biases (e.g., faulty memories), and participants may forget to fill in the questionnaire in expected time period. For that, I send reminding emails to participants at the same time every week, and the participants were adviced to fill in the questionnaire about the same time every week if possible.

5.1.4 Interview

I used interviews to collect data in the field study, gathering information about participants' experience, motivation, and opinions. It also offers the change to observe how participants use the visualization application, and investigate their cognitive process of develping insights from data, which might be hard to collected from the survey. The use of personal visualization is mediated in everyday context is a highly personalized environment, also interacting with culture and social factors (Chapter 2). Applying quantitative method in this case might limit researchers' insight on how and why certain design works. Particularly, in-deepth interview tend to work well with ethnographic methods, e.g., case study or field study [73]. It definitely provides advantage to investigate open-end quesions in research.

Inevitably, interviews usually would suffer from bias problems that could be from researchers (e.g., attitude biases) or participants (e.g., recall biases). Also in field study, the interview itself may bring intervetion effect. Considering that, times of interviews in both research groups (Chapter 9) were keep the same to balance the effect.

The interview data were analysed with quantitative approach based on ground theory through open coding process [3]. Transcrition and observation data were mannually tagged. As patterns and trends of tags were investigated, tags were merged into higher level categories. As this process is repeated, new models or theories were derived. This approch is good at analysing unstructed data, e.g., from interviews, observations, etc., but the caveat is that the coding process may bring in bias from the researcher, e.g., the perspective that she interpret the data.

5.2 Summary

Generally, evaluating visualization used in everyday life is difficult and challenging (Section 2.4.5). The choices of research methods considered in this thesis are based on my research questions. Specifically, this research is mainly to investigate a reflective approach of visualization design in feedback applicaitons. The research methods used in this work are based on this focus. Potentially, it would be helpful for future research in PV&PVA. In the next chapters, I will discuss the two primary studies where these methods applied.

Chapter 6

Viability Study

As the on-calendar design alternatives are developed based on visualization principles, whether they valid as expected is still unknown. To support attentional ambience as embeding a visualization layer into one's calendar, designs need to be visually salient but perceptible. Controlled lab experiment is a good option to evaluate these perception and legibility criteria (Chapter 5). In this chapter, I describe the viability study, where I employ lab experiemtns to evaluate visual interference and perceptability of design alternatives (Chapter 4). This stage of my research is to evaluate the idea of on-calendar design approach with proposed alternatives and suggest design options for later implementation.

6.1 Background

The on-calendar approach directly integrates personal time-varying data into a personal digital calendar (i.e. the same calendar that people already use for managing their personal appoitnments) (Figure 4.1,4.2). Our design goal is to present information in an unobtrusive way, informing people while blending into the environment and requiring low attentional demand [83]. In other words, we mash-up information sources in a familiar tool (a digital calendar); that is, the additional information is attentionally ambient to avoid interfering with the primary function of the application (i.e., tasks of schedule management). Meanwhile, calendar events can provide context to reason about data patterns that are aligned with them.

I am exploring the viability of this approach as the first step to investigate the effectiveness of visualization in this design approach. Will people accept the idea of

having additional data added to their digital calendars? Can this data be integrated in a way that does not interfere with normal calendar activities, yet enables the data to be perceived? These questions are the focus of this step. Thus, I investigated visual interference and perceptibility of visualization alternatives with lab experiments, and meanwhile aimed to narrow down the design alternatives of on-calendar visualizations for later development. The goal of the experiments here was to identify visual encodings that minimize visual interference with normal calendar tasks, while supporting effective data perception. The experiments in this chapter focuses on these basic design issues related to interference and perceptibility rather than other factors influencing sustainable behaviors, such as motivation and social interaction.

6.2 Experiment Design

two experiments explored how the following factors that influence visual interference with normal calendar tasks and graphical perception of the quantitative data:(all within-subjects factors in the study):

- Display location: overlaped or side-by-side.
- Visualization type: line graph, colored-region or luminance.
- Calendar scale: week[1] or month.

6.2.1 Apparatus

The experiments were conducted in a controlled usability lab, and images were displayed on a 21" inch monitor at 1280*1024 resolution. Static visualization images were created for each combination of the above factors. I chose static images in the experimental study to minimize the influence of any system delay and to ensure that trials had a clearly defined correct answer. Trials were presented in random order by a custom Java-based slide-show and data collection program. Multiple-choice answers were presented in a new page after the participant had done the visual search or data interpretation, so time to input the answer was not included in the timing measures. The schedule in the experiments was a sample schedule from a faculty member, and the energy data (electricity usage) were "hard coded" according to each of the tasks.

[1]In the viability study, day view was not considered because it would use an identical visual encoding as the week view.

To prevent color effects, all visualizations were in grey (or with luminance varied). The colored calendar event blocks on the calendar were semi-transparent (with alpha level at 0.6).

6.2.2 Participants

Thirty one participants were recruited (14 female, 17 male) including undergrad and graduate students, with a diversity of backgrounds (computer science, engineering, chemistry, biology, education, social science and political science). Fifteen of them participated in the first experiment and 16 participated in the second one.

6.2.3 Experiment I: Calendar Tasks

Experiment I investigated the interference of visualizations with normal calendar tasks. Tasks were designed to involve only visual search, to eliminate any individual differences in interaction speed, such as event editing or text input. We also chose to investigate visual search tasks because they are the most likely to suffer from interference from the addition of background data displays. Participants were asked to search for a single event (e.g., "What time is the group meeting?") or count repeated events (e.g., "How many seminars do you have in the week/month?"), without being informed about the additional data layer in advance (see the full list of task questions in Appendix **??**).

Experiment I was to test the folloing hypotheses:

- visualizations displayed in overlapped position would have greater interference than visualizations displayed side by side(H1.1).
- line graph would interfere with calendar activities the least and luminance would interfere the most (H1.2), since luminance would take the most space on the calendar background and line graph would take the least.

6.3 Experiment II: Visualization Tasks

The second experiment investigated the graphical perception of visualizations on a calendar; that is, how people interpret the meaning of data patterns. Participants were asked to complete tasks that involved perceiving general patterns from the visualizations, but not precise values. The tasks were derived from established temporal

data tasks described in the literature. Saraiya et al. suggested that people would get insights from overview, patterns, groups and details [96]. Furthermore, elementary (local patterns or extreme values) and synoptic tasks (overall estimate or distribution) are typical ways to explore temporal data [8]. Thus we included tasks that involved investigating local patterns and estimating overall summaries. For instance, participants were asked to interpret the energy consumption spikes (e.g., "What day do things start the latest in the morning?"), local patterns (e.g., "Which evening do you consume the most energy from 7pm to 9pm?") and compare summaries of days (e.g., "Do you consume more energy on Tuesday than Thursday?"). The full list of task questions is included in Appendix **??**.

Experiment II was to test the folloing hypotheses:

- visualizations displayed side by side would be easier to perceive than visualizations in overlapped position (H2.1), since side-by-side visualizations do not have calendar activities occluding the data representation
- colored-region would be easier to perceive than luminance or line graph (H2.2), considering position encoding is a stronger visual cue than luminance encoding and color difference would enhance visibility of the contour.

6.4 Procedure

Each of the participants completed a set of trials comprised of all combinations of the three factors (visualization type, display location, and calendar scale) for one of the two experiments. All trials were presented in a random order. Three practice trials were presented prior to the main trials. In Experiment I, an additional set of trials was included for a control condition that had no visualization on the background; control trials were also in random order, intermixed with the other trials.

6.5 Experiment Results

In the viability lab study, task time and task accuracy were set as the measures. Task time was defined as the time period only for viewing the images, and did not include the time of answer input (via multiple choice). Accuracy rate refers to the percentage of tasks completed correctly.

Display	Encoding	Experiment I	Experiment II
side-by-side	line graph	90	78
	colored-region	80	87
	luminance	87	84
overlapped	line graph	40*	44
	colored-region	53*	87
	luminance	87	97

*Significant difference compared with the control condition

Table 6.1: Accuracy Rates (%)

Data of the two experiments were analysed with respect to three factors: display location (overlapped or side-by-side), visualization type (colored-region, luminance or line graph), and calendar scale (month or week view). Task time was tested with repeated measures ANOVA followed by pairwise comparisons, as the data were confirmed to fit a normal distribution via Q-Q plots. All post hoc comparisons used Bonferroni correction. Accuracy was analyzed with Cochrans Q due to its binomial distribution, followed by Bonferroni-corrected McNemar's tests for pairwise comparisons. For the control condition in Experiment I, I used Bonferroni corrected paired comparisons to compare the control condition task time to each combination of Visualization Type and Display Location.

6.5.1 Experiment I: Visual Interference

Visualization Conditions vs. Control Condition

Accuracy rates are shown in Table 6.1. There was an overall significant difference in accuracy across conditions ($Q(6) = 34.08$, $p < 0.01$) in Experiment I. When comparing visualization conditions to the control, McNemar's tests showed that task accuracy was significantly lower with overlapped colored-region than the control condition ($p < 0.04$), and with overlapped line graph compared to the control ($p < 0.01$). The other conditions were not significantly different than the control. None of the visualization conditions was significantly different than the control condition for task time (shown in Figure 6.1).

Visualization conditions

Task time: The three-factor ANOVA analysis showed a significant main effect for Visualization Type ($F(2,28) = 6.00$, $p < 0.02$, $\eta^2 = 0.30$) and a significant main effect for Calendar Scale ($F(1,14) = 194.89$, $p < 0.01$, $\eta^2 = 0.93$) but no significant main effect for Display Location ($F(1,14) = 4.07$, $p < 0.07$, $\eta^2 = 0.23$). There were no significant interactions.

Overlapped visualizations (M=17.17, SD=7.84) took slightly more task time than side-by-side visualizations (M=16.16, SD=7.41), as shown in Figure 6.1. Pairwise comparisons between groups of Visualization Type showed that line graph took significantly more task time than colored-region ($p < 0.02$). Participants spent significantly more task time with month view than week view (not shown).

Accuracy: The side-by-side conditions had higher accuracy rate (86%) than overlapped (60%), ($Q(1) = 15.11$, $p < 0.01$, see Table 6.1). With both side-by-side and overlapped conditions, the accuracy rate was 87% with luminance, 65% with line graph and 65% with colored-region, ($Q(6) = 34.08$, $p < 0.01$). Pairwise comparisons with McNemar's tests showed that luminance had significantly higher accuracy than line graph ($p < 0.03$). Accuracy rate was significantly lower in month view (64%) than in week view (81%)($Q(1) = 6.82$, $p < 0.02$).

Summary of Experiment I

Colored-region and line graph visualizations caused minor interference with task accuracy when located overlapped, but luminance and side-by-side visualizations had no observable interference. Among the visualization types, line graph caused greater interference than the others as measured by both time and accuracy.

6.5.2 Experiment II: Graphical Perception

Task time: Response times for Experiment II are summarized in Figure 6.2. The three-factor ANOVA analysis revealed a significant main effect for Display Location ($F(1,15) = 10.51$, $p < 0.01$, $\eta^2 = 0.41$), and a significant main effect for Visualization Type ($F(2,30) = 10.72$, $p < 0.01$, $\eta^2 = 0.42$), but no significant main effect for Calendar Scale,($F(1,15) = 2.72$, $p < 0.12$, $\eta^2 = .15$). The interaction between Visualization Type and Time Scale was significant ($F(2,30) = 6.72$, $p < 0.01$, $\eta^2 = .31$). Other interactions were not significant.

Figure 6.1: Boxplots showing task time of Experiment I)

The overlapped condition (M=10.03, SD=8.29) took significantly less task time than the side-by-side condition (M=13.02, SD=8.81), as shown in Figure 3. Pairwise comparisons showed that luminance took significantly less task time than colored-region (p <0.01) and line graph (p <0. 01), but this difference occurred only in month view, not in week view. Recall that luminance in month view represented aggregated information (daily average), a possible explanation for this difference.

Accuracy: the accuracy rate was 76% with the overlapped condition and 83% with the side-by-side condition, but these were not significantly different ($Q(1)$ = 1.82, p <0.18, see Table 6.1). With both side-by-side and overlapped conditions the accuracy rate was 87% with colored-region, 90% with luminance and 61% with line graph, with a significant effect ($Q(2)$ = 20.44, p <0. 01). Pairwise comparisons with McNemar's tests showed that line graph had significantly lower accuracy than luminance (p <0.01) and colored-region (p <0.01). Accuracy was particularly low with line graph when it was located overlapped (44%). With respect to temporal scales, the accuracy rate was 85% in month view and 74% in week view, not a significant

Figure 6.2: Boxplots showing task time of Experiment II)

difference ($Q(1) = 4.17$, p <0.41).

Summary of Experiment II

The overlapped condition facilitated better graphical perception than the side-by-side condition. Luminance was easier to perceive than colored- region and line graph (as measured by task time), possibly because of the data aggregation in month view. Line graph was the most difficult to perceive (it had lower task accuracy).

6.5.3 Post-study Questionnaire

Participants were asked to fill in a questionnaire after the experiments, with respect to their experience of visual distraction, graphical perception and aesthetics. Those questions were likert scale ratings that were analyzed with repeated measures ANOVA after it was confirmed to match to a normal distribution via Q-Q plots. Post hoc comparisons used Bonferroni correction.

Figure 6.3: Visual distraction reported by participants (-2 is "very distracting", 2 is "not distracting")

Figure 6.4: Graphical perception reported by participants (-2 is "very difficult" and 2 is "very easy").

Figure 6.5: Aesthetics reported by participants (-2 represents "very poor" and 2 represents "very good")

Visual Distraction

As shown in Figure 6.3, participants in Experiment I rated the in-cell condition more distracting than the in-band condition, and line graph was rated more distracting than colored-region and luminance. We found a significant main effect of Display Location ($F(1,14)$=8.90, $p = 0.01$, η^2= 0.39), but not Visualization Type. The interaction was not significant. Participants also reported that they could not stop being visually distracted by sparkline, particularly with the in-cell condition.

Graphical Perception

As shown in Figure 6.4, participants in Experiment II rated the in-cell condition as easier to perceive than the in-band condition. Line graph was rated the most difficult visualization to perceive. We found a significant main effect of Visualization Type ($F(2,30)$=.04, p =3.72, η^2 =.20), but not Display Location. The interaction was not significant. Pairwise comparisons showed that colored-region was rated significantly better than sparkline (p <0.01).

Aesthetics

In-band visualizations were rated more appealing than in-cell visualizations, and sparkline was rated the least appealing visualization (Figure 6.5). There were significant main effects of Display Location ($F(1,30)$=5.15, p =0 .03, η^2 = 0.15) and Visualization Type ($F(2,60)$=26.19, p <0.01, η^2 = 0.47). The interaction was not significant. Pairwise comparisons showed that sparkline was rated significantly lower than colored-region (p <0.01) and luminance (p <0.01).

Summary

User ratings confirmed the performance findings: in-cell visualizations had greater interference with calendar operations than in-band ones, but data was easier to perceive with in-cell visualizations. Sparkline was rated the most interfering and the most difficult visualization to perceive. Considering aesthetics, in-band visualizations were preferred over in-cell visualizations, and colored- region and luminance were preferred over sparkline.

6.6 Discussion of Lab Experiment Results

6.6.1 Interference (Experiment I)

The hypothesis (H1.1) that in-cell visualizations would have greater interference with calendar activities compared to in-band was confirmed. In Experiment I, task time was slightly faster with in-band visualizations and accuracy was significantly higher. However, presence of quantitative visualizations on a calendar did not greatly compromise the regular calendar tasks. Although the results showed a drawback of in-cell visualizations for calendar task accuracy (mostly from line graph), there were no significant differences in task time between the control condition and the visualization conditions.

The hypothesis (H2.1) that in-band visualizations would be easier to perceive than in-cell ones was not confirmed. On the contrary, in-cell visualizations were easier to perceive than in-band ones. The presence of calendar activities in the same space did not compromise performance at graphical perception.

6.6.2 Perception (Experiment II)

As hypothesized, line graph would interfere with calendar activities the least and luminance would interfere the most (H1.2). This hypothesis was not confirmed; in fact it was directly contradicted. Task time with line graph was significantly longer and its error rate was significantly higher than the others. Participants also qualitatively reported that line graph caused greater interference than colored-region and luminance. I speculate that this interference might be caused by the similar color (grey) of the line graph and activity text, which could make reading the text more difficult. With colored-region, the filled color area tones down the interference to some degree. Thus, we eliminated line graph as a visualization option in our later implementation.

The hypothesis (H2.2) that colored-region would be easier to perceive than the others was not confirmed either. Luminance had higher accuracy than line graph and was faster than both other visualizations in month view. The advantage of color coding was observed in small graph scales (month view) but not in large graph scales (week view). I strongly suspect that this is due to aggregation: luminance in month view encoded daily average value while line graph and colored-region represented continuous daily data.

6.6.3 Design Implications

These results indicate the viability of compositing quantitative data into a typical calendar view without compromising the effectiveness of either the data visualization or the calendar itself, but they also highlight a number of design implications and issues. Foremost, it seems clear that there is not a "one size fits all" solution. Even in the small set of factors we considered in the study, certain visualizations fit better at different calendar scales and in different calendar codings. For example, the performance of luminance, in which we filled the entire cell with a gray scale representing a single aggregate value, was most effective in the month view; its superiority to the coloured-region in the week view was insignificant. The continuous encoding in week view provides a finer resolution of the data (and is reported qualitatively as moderately less distracting). However, luminance in the week view may interfere with the colour coding the user has applied to her personalized calendar entries, whereas these entries in the month view are typically presented as text, so filling the cell could be less interfering. In addition, there are a limited number of discriminable levels in gray scales, thus reducing the resolution of the visualization. Thus the usability principle of consistent coding may be ineffective here: adaptive visualization, where the representation takes a different form according to the scale of the calendar and the desired granularity of the data, better fits the goals of clarity and visual non-interference. Based on this, we offer some design suggestions to optimize visualization of quantitative data on a calendar.

In-cell visualizations support better graphical perception with the sacrifice of minor interference, compared to in-band visualizations. Therefore, in-cell approaches might be the better default design choice, especially for small devices where screen space is very limited. However, in-cell visualizations may not extend well to situations where the calendar is extremely dense (e.g. back to back meetings all day); in this case, in-band could be a viable alternative.

6.6.4 Attentional Ambience

This work introduced the concept of attentional ambience as an extension to ambient visualizations. Attentional ambience is defined by the degree to which the representation can exist in a visual middle ground where features can be pulled into the foreground or relegated to the background by slightly changing the degree of attention [17]. It is this capacity of our visual system that supports the kinds of information

mash-ups proposed in this paper.

Limitations in the study scope encourage further investigation. I did not, for example, consider device size nor context of use (mobile vs. fixed). Clearly these different conditions will influence the degree of visual saliency or subtlety that is most effective for attentional ambience. The focus of this study was to investigate the degree of visual salience that provides the best balance between ambience and perceptual efficiency for the two conditions (colored region and luminance), as they seem the most promising. The alpha level used in both was 0.6, but previous research in overlaid structures suggests that we can use a much more subtle level of 0.2 to provide a just attendable difference [17] between the two data sets. Possibly, a lower opacity of the quantitative data might have positively affected the calendar accuracy results without compromising the quantitative interpretation. Thus, the opacity of visualization could be customized to better support attentional ambience. Particularly in this case, attentional ambience would be influenced by the characteristics of individuals' calendar as well, e.g., the density of calendar events displayed, existing color used to code events, etc.

Moreover, the "appropriate" levels of attentional ambience introduces new design options. There may be thresholds at which the visual salience should be increased: for example, high blood glucose levels (diabetes data) or energy consumption spikes that exceed a daily average. We believe this approach may thus support both informed historical analysis and near-real-time monitoring and alerting tasks.

Chapter 7

Implementation

Following the lab study, I implemented a working prototype as an interactive web application (Figure 7.1). The lab study results suggest to remove line graph from the visualization alternatives because it caused significant interference between the visualization layer (additional data stream) and calendar events. The implementation kept the other visual encodings because they all seemed viable and I wanted to see which ones people would prefer in practice.

Figure 7.1: web application of on-calendar visualization using Google calendar API and displaying household smart meter data.

The web application basically works as an online digital calendar, synchronizing with calendar events (through Google API [1]) and also fetching live data feeds (from a household smart meter or Fitbit data API [2]) 7.2. The web application was implemented with PHP and Javascript, and the visualization layer was implemented with D3.js [3]. It could be run on desktop (or labtop), tablet and mobile phone with a browser, but the layout was designed for desktop brower size (it was not customized for mobile devices). The appliction was hosted on the server at Simon Fraser University.

Figure 7.2: Data flow of the on-calendar visualization system.

In the application, personal calendars are placed in foreground and the data visualization is displayed on the background (see the architecture in Figure 7.3). The calendar settings are controlled from one's google account (e.g., event colors). The on-calendar application facilicates basic calendar functions, allowing users to select calendars to display, edit their calendar events, and control the calendar view. The drop-down list from "Select Calendar" button enable users to select multiple calendars to display. The interactions of event editing are kept the same as Google calendar, for example, clicking event brings up pop-up dialog window for event details.

The visualization layer can be customized with the top control panel. For example, the chart can be displayed either overlapped or side by side. Users can also choose the visual encoding: either colored-region, as shown in Figure 7.1, or luminance. To balance the ambience of foreground calendar events and background data stream,

[1] https://developers.google.com/google-apps/calendar/
[2] https://dev.fitbit.com/
[3] https://d3js.org/

Figure 7.3: Architecture of the on-calendar visualization system.

the user is also allowed to adjust the settings of transparency, visualization color (with grey as default), and chart scale. Meanwhile, a script is running at backend to collection data of user interations, e.g., the time when the application is brought up, what buttons users click, what visualization settings users change, etc.

Chapter 8

Pilot Studies

In the viability study, visual interference and perceptibility were investigated with "hard-coded" data and tasks (Chapter 6). The implementation enables real-time data and allows users to interact with the personal data in real-life context. Meanwhile, I need to investigate if the implementation is follow the design goal and inline with the results in the viability study. Thus, I conducted two case studies as pilot studies with the early version of implementation: home energy conservation (with home power meter data) and personal physical activities (with data from Fitbit), where the data are highly related to personal context (e.g., time, locations and activities). In this chapter, I describle two case studies where I deoployed the prototype tool in an experimental eco-friendly town house and a small group of students, respectively. The cases studies were aimed to test the early version of implementation in real life situation and identify usability issues, which helps to revise the application for later field study. As well, I present the implementation revision in the end of the chapter.

8.1 Household Energy Consumption

In the first case study, the web application was deployed to people living in an eco-friendly smart home [1] (see Figure 8.1) with the data source connected to their electricity meter. This experiemntal townhouse was occupied by tenants, and was equiped with different sensors to collect data of various utility consumption sources. The data feeds synchronized with the calendar tool were the overall electricity consumption. Both of the tenants were musicians, so they have flexible working schedules and

[1] https://www.sfu.ca/westhouse.html

practice at home regularly even on weekdays. In this case, the calendar tool may be helpful to reveal the pattern of their at-home activities and their relationship between the consumptional behaviors and the electricity use.

Figure 8.1: Westhouse: eco-friendly smart home for the case study.

The study had two phases: baseline session and prototype deployment session. The first session was to understand baseline energy consumption, without asking the participants to use the web application. I conducted an interview with the participants after the first session. Afterward, the participants were introduced to the web application and asked to use the application in the way they use digital calendars in their daily lives for the next two months. At the end of the second session I conducted another interview with the participants. Questions of the interviews were mostly about the participants understanding of their energy consumption behaviors, their experience of using the calendar tool in everday life, and how they used the contextual informaiton from personal digital calendars to make sense of their consumption data.

By the end of the study, the participants reported that the calendar events were

helpful as a reference tool that helped them figure out couple of things, e.g., the baseline usage of the house, spikes caused by the fridge, the heater and the washing machine, etc. They could easily see what days they worked at home, went out because of work, went on vacation, etc. For example, the wife reported that she had noticed a continuous large consumption chunk one day in the past, and then from the calendar information (it was her concert on the next day) she figured out that it was when she was cooking a lot of food for a party after the concert. From the temporal data positioned along the timeline, she also realized that using stove took less power than oven because the oven usually took a much longer time. The participants also reported the general context provided by the calendar view helped them identify some gaps between their prior knowledge and current facts, even when the activity was not marked on the calendar. For example, they usually did laundry right after they got up before things started. The spikes caused by the spinning of the washing machine in the beginning phase were totally out of their expectation (I am surprised to see the spinning takes so much power.). However, the on-calendar visualization was in low usage after the first two weeks. The participants reported that they were more familiar with iCal more than Google calendar, and used iCal more often. The inconsistant looks and features prevented them using the on-calendar visualization as the daily default option. They hoped that the visualization could be integrated into their iCal application. Thus, I modified the design of the application and make the looks and features more close to Google calendar and also refined the recruiting requirements in the later field study where participants who were familiar with Google calendar were included (Chapter 9.1).

8.2 Personal Fitness

In the second case study, we deployed our calendar application to 10 undergraduate students as part of a Psychology seminar course. Different from the energy consumption case where the consumption is related to activities of all family members, I considered the fitness case because personal fitness data might be more relavent to individuals' activities and directly linked to one's personal calendar.

Participants (eight females and two males) in this study were undergraduate stdudents from Psychology department with age between 21 and 25. They were provided Fitbit devices 8.2 (all of them were novice Fitbit users), with which they were asked to track their daily physical activities. However, the online Fitbit accounts were not

Figure 8.2: Fitbit trackers use in case study of personal fitness

provided, so they can not use the default Fitbit online tool for feedback purpose. Instead, they can only use the display on the device for feedback information. The study last for 4 weeks. In the first two weeks, they can only rely on the device display for feedback, and the on-calendar web application were introduced after the first two weeks. In the later two weeks, they were asked to use the on-calendar application as the default feedback tool. Two interviews were conducted at the end of week 2 and week 4, respectively. In the first interview, partcipants were asked about their experience in the first two weeks. I also helped participants set up the device and gave tutorial. In the second interview, participants were asked about their experience of using the on-calendar application and how they use the calendar context for reasoning. (see outlines in Appendix **??**).

Participants in this study reported that they liked the idea of integrating Fitbit data into a digital calendar. It helped to reveal their school-life patterns, for example, start of the day, sitting most of the time for classes, walking (or running) during class intervals, studying during the weekend. It also helped to easily identify anormalies. For example, a participant found his active activities mostly happened at midnight. It is because he often worked in a restaurant with a night shift. The on-calendar visualization also revealed errors in data collection. One partcipant found the movement during her class event on her calendar, and suspected it was her hand movement even when she was sitting. Participants' comments about visualization preferences confirmed with the lab results in early viability experiments. They reported that colored region was easier to perceive than luminance ("I can see the time, where the peaks are, and the intensity"). Visualization display overlaped ("in-cell" option) was regarded easier to read because the chart was bigger, even though overlapped with calendar events.

However, part of participants' Fitbit data were often found missing on the visualization. They reported in the interview that they forgot to wear the device or charge the battery. The novice Fitbit users may not be familiar of using the devices or not comfortable to wear the device all the time. Thus in the later deployment, experienced Fitibit users were recuited (Section 9.1).

Most of the participants in this study had very sparse calendar events because the calendar events on students' calendar were very simple, mostly related to their school schedules (for example, course arrangement). The only exception was one participant who has busy calendar schedules (because he had to manage a few student organizations) and use digital calendar very frequently. During the study he used the on-calendar visualization more than the others. I suspect it might be a factor to influence the application usage in this case, so I tried to balance the diversity of participants in the later field study (Section 9.1).

Meanwhile, the same issue as the first case study was brought up as well. The overall usage of the on-calendar application was very low. Participants commented on the inconsistant look&feel and features compared with calendar tools that they used most often. Some participants suggested to include the visualization in their iCal application. Thus, a revision of the on-calendar visualization seems necessary.

8.3 Summary of Pilot Studies

These pilot studies were aimed to collect feedback on the design approach and identify usability issues with our early version of implementation. Overall, the results were encouraging. Participants liked the concept of integrating a data stream within a personal digital calendar. They believed it was an easy way to access and keep track of relevant data. They also found that the context provided by their calendars was helpful for interpreting patterns and abnormalities in the data. More interestingly, we found that people could easily see their life routines reflected in the on-calendar visualization, even though many of the relevant routine activities were not recorded in the calendar (e.g., cooking, laundry, showers, etc.). This suggests that the context from personal calendars can provide high-level information to help people understand personal data patterns without requiring extra effort for people to record their daily routine activities.

Meanwhile, the studies also revealed a few usability issues with the application. The logging scripts showed that the online web application had very low usage after

the first few days when it was introduced, which might be caused by some inconsistencies in functionality and look&feel as compared to the commercial calendars that participants regularly used (e.g., iCal or Google Calendar). These inconsistencies presented a barrier for participants and prevented them from routinely using the web application instead of the one they currently use. This meant that this tool was used more as a dedicated visualization tool for accessing the data stream rather than the way as intended, where the data would be viewed in a secondary background stream. However, participants hoped the additional visualization layer could be added into their own calendar application, showing merit for the intended design. These results suggest that the prototype needs to be revised before larger scale deployment, to make the features consistent with digital calendars that people use regularly. Moreover, I found that visualization customization and time scales may be subject to many personal preferences. Some preferred week view for bigger charts, and some preferred month view for overview. Participants reported that the advantages of colored-region was that the spikes helped her recall things for reasoning. However, luminance displayed in-cell seemed to be most irritating for them. For these reasons I keep the visualization settings adjustable.

8.4 Design revision

The revision of early version of implementation primarily focused on the consistancy of fuctions and look & feel with exiting calendar tools (specifically, Google Calendar in this case). The final version was implemented with PHP and javascript (D3.js for visualization layer), together with Fitbit API [2] and Google Calendar API [3] (see Figure 8.3).

First, to make the application easy to manage with, previous customized visualization settings are remembered for individuals. Poeple could customize the visualization to fit themselves, and the on-calendar visualization will be set with their previous preference when next time they log in.

The on-calendar application was expected to use as everday calendar tool, so I made the control panel (for adjusting visualization settings) more subtle on the top. I might reduce the distraction of non-calendar functions while people are performing scheduling tasks.

[2] https://dev.fitbit.com/
[3] https://developers.google.com/google-apps/calendar/

Figure 8.3: On-calendar feedback application used in the study (week view with Fitbit data displayed as a line graph overlapped with calendar events). See more screenshots in Appendix **??**.

The calendar layout was re-designed according to Google calendar [4]. Calendar browsing buttons (e.g., selecting time scale) were placed on the top of the main calendar view. A small calendar was added on the left for quick date selection. Below that was the calendar selection panel, with which people could decide which calendars to be displayed in the calendar view. Similar as Google calendar, all day events could be shown at the top of the day below the date label (only in day view and week view).

Other than that, people can customize the row heigh of calendar cells. With smaller height, a longer time period of the day can be seen (only in day view and week view), so it could provide a better overview. Considering privacy, functions of access control were enabled (e.g., sign out the application)

[4]https://calendar.google.com/calendar

Chapter 9

Field Study

Results from pilot studies confirmed that the on-calendar design could provide daily life context for people to reason about their data and support ambient attention. However, insights from pilot studies are limited. The early implementation have a few usability issues. Participants in the pilot study may not properly represent the target user as expected. After revising the application based on the feedback, I deploy it in a longitudinal field study. In this chapter, I discribe an eight-week field study with the latest implementation connected to Fitbit data. This study primarily focuses on two questions:

- To what extent can people use calendar data as context for reasoning about their fitness data?
- How do people react to the idea of integrating feedback data into their personal calendars?

This study compared experimental and control groups (who used our calendar prototype and Fitbits standard feedback tools respectively), aimed to investigate mash-up design approach and the influence of providing extra context for reasoning. The emphasis was on exploring peoples experiences with the on-calendar visualizations rather than measuring differences in behavior change, as suggested by early research [103, 89, 32, 69]. Therefore, I employed a qualitative approach with open-ended research questions rather than statistical comparisons between the groups.

Participants	Age Group	Gender	Fitbit experience (current tracker)	number of fitness trackers used before current one
p1	30-39	Male	2.5 years	0
p2	18-29	Female	1 month	0
p3	30-39	male	1 year and 1 month	0
p4	30-39	Female	3 months	0
p5	18-29	Female	3 months	0
p6	40-49	Female	11 months	2
p7	18-29	Male	2 Months	0
p8	30-39	Female	1 year and 1 month	2
p9	60+	Female	2.5 months	2
p10	18-29	Female	3 months	0
p11	40-49	Female	4 months	0
p12	18-29	Female	1 year	2
p13	18-29	Female	1 month	2
p14	30-39	Male	1.5 years	0
p15	30-39	Female	4 months	0
p16	60+	Female	1 year	2
p17	30-39	Female	4 months	0
p18	30-39	Female	3 months	0
p19	30-39	Male	3 weeks	2

Table 9.1: participants in Fitbit field study

9.1 Participants

I recruited participants among existing Fitbit users instead of providing Fitbit devices, considering that existing users, compared with new users, already had some motivation to use feedback tools. Existing users were also already experienced with Fitbits basic feedback applications, and for them, using a fitness tracker and its software would not itself be a novelty. Meanwhile, the participant screening required that participants be familiar with digital calendars and have a Google account (necessary to use the web application). In total, 21 Fitbit users participanted in this study with age ranging from 20 to 60+, 15 female and 6 male (see Table 9.1). Two of them (one female and one male) dropped out after the first two weeks. Seven of them used other fitness trackers before.

9.2 Conditions

Participants were randomly divided into two groups: Control (C1∼C9) and Experiment (calendar Visualization) (V1 ∼ V10). This design allowed us to investigate whether extra context from a personal calendar could improve peoples understanding of their feedback data. Participants in the control group used their baseline feedback application (i.e., that provided by Fitbit). Participants in the Visualization group used the baseline feedback application in the first two weeks; they were then introduced to our web-based calendar visualization after week 2. Visualization group participants were asked to use our calendar application as their primary scheduling and feedback tool; however, they were not prevented from also using their default calendar service (e.g., Google calendar or iCal) and Fitbits feedback tools. With the control condition, the use of original Fitbit application is equal in the two groups.

Figure 9.1: Study procedure

9.3 Procedure

Before the first week, we met participants and introduced the procedure. During the first two weeks baseline information was collected and participants were told to continue using Fitbit as they had done in the past (see Figure 9.1). I interviewed all participants in week 3, during which Visualization group participants were introduced to our on-calendar visualization.This interview was to investigate how the participants use their feedback tools currently, and set up the baseline of using feedback tool for reasoning to compare before and after the visualization tool was introduced to Visualization group. To investigate their initial experience and help the participants on technical issues of using on-calendar feedback application, I interviewed all participants again in week 5. participatns in Control group were also interviewed to

balance the influence of interview intervention in the Visualization group. Final interviews were scheduled in week 9, in which participants shared their experience of using the feedback tool in everday life during the study. During these interviews, participants were asked to review their Fitbit data with and without their feedback application, identify patterns and anomalies, and reasaon about the patterns and anomalies. These tasks were to investigate the general awareness of personal awareness, how they reason about data patterns and anomalies using feedback tools, and how they use inferential context in the reflection. Meanwhile, during the total eight weeks, participants were asked to fill in a weekly International Physical Activity Questionnaire (IPAQ) [2] through an online portal. Reminder emails with the survey link were sent to them on Friday afternoon every week. At the end of the final interview, participants in the control group were also introduced to the on-calendar application and asked for comments. The interview outlines are included in Appendix.

9.4 Data collection

The data collection included weekly surveys, application logs and interviews. Although participants Fitbit data were accessible to measure physical activity level, this data was not to used because of its incompleteness. (Fitbit cannot accurately capture activities such as cycling, spin class, and swimming. Sometimes they were misfunctioning.) Instead, the estimated physical activity (PA) were evaluated by the weekly IPAQ survey that is often used method to meanusre physical activities [5, 29, 22, 52, 25, 37]. V8s survey data were dropped from the analysis because only 3 surveys were submitted. The remaining 18 participants submitted at least 6 entries of the online survey.

Metabolic Equivalent(MET) is a commonly used physiological measure to assess physical activities [58]. METs of the weekly surveys were calculated according to the scoring protocol of IPAQ [2]. Meanwhile, interactions of participants while using the calendar visualization (e.g., change visual encoding or layout) were automatically logged. In the interview , the participants were asked to recall their PAs to explain their data patterns, their experience using the feedback tools, and the impact in their life. During the interview, they were also asked to bring up their feedback application and reason about their own data patterns. I observed how they interacted with the application and how they performed tasks to reason about their data. The following analysis focused on qualitative feedback about the on-calendar design approach. This

study was most interested in how the approach would influence peoples ability to reason about their feedback data, and to what extent they would find the on-calendar visualizations helpful and/or disruptive. Therefore, I employed a primarily qualitative analysis approach.

9.5 Results

9.5.1 Physical activity levels

I first examined the physical activity (PA) variation of the two groups before and after the calendar intervention. The results showed the two groups were not significantly different in MET measures (t(16) =0.53, p=0.60, Cohen'd=0.27). PA tended to increase more for the experimental group than for the control group, but this was overshadowed by individual differences (Figure 9.2). Participant comments suggested that behavior change (PA variation) was most influenced by other aspects in their lives, e.g., traveling (V2, V6, C5), relocation (V10, C6), facility service interruption (C3), or a training program (V5, C4). However, the influence of single intervention is difficult to qualified in behavior change and measuring behavior change was not the main goal of this study. Instead, I focused the majority of our analysis instead on system use, its role in the feedback process, and how it influenced peoples reasoning.

Figure 9.2: Change in MET values from weeks 1-2 (baseline) to weeks 3-8 (intervention) for individuals in control and experiment groups. Each mark represents one participants change in average MET scores.

9.5.2 System Use

Application logs showed 152 visits (user sessions) and 208 user interactions (setting and view changes) during the study. The peak usage was in the morning (around 10 am) and in the evening (around 9 pm). The application remained active for durations ranging from one minute to four days (M = 1043, SD = 2791), indicating that people used the application quite differently: some brought it up for a quick look while others continually kept the tab open. That means participants might keep the browser tab of the application open for calendar use while working with their computer. This might indicate the efficacy of the mash-up design approach: use addtional visualization layer to support attentional ambience.

Figure 9.3: Change in MET values from weeks 1-2 (baseline) to weeks 3-8 (intervention) for individuals in control and experiment groups. Each mark represents one participants change in average MET scores.

By default the visualization was encoded with Coloured Region and in grey color

(with 60% transparicy), and displayed with an overlapped layout in week view. When the application was introduced, participants were asked to try and explore all possible settings.The application was implemented to remember the customized settings, so the bias of default visualization settings was minimized. Application logs (Figure 9.4) also showed that all participants preferred Colored Region (line graphs) as the visualization setting. They reported that luminance as the visual encoding required extra cognitive effort to understand, and the color made the calendar look busy and interfered with calendar events (particularly when the calendar events were color coded). Grey color and overlapped display were used most often, suggesting that they were least disruptive. Only one participant chose to show the visualization layer in a separate band side-by-side with calendar events (kim). In the interviews, participants also reported that the visualization layer did not interfere with their use of calendar events (V1, V2, V9, V10), especially with the grey color. This suggests that with proper visual encoding, displaying data as an additional layer on a calendar need not interfere with regular calendar use. Most participants stayed on week view most of the time, and switched between week and month views (175 view switches were logged among 10 participants) when they explored data patterns with different time ranges and levels of detail.

Figure 9.4: Preferred visualization settings (experiment group). The most popular visual encoding was a grey line chart overlapped with the calendar data in week view.

9.5.3 A Model of the Behavior Feedback Process

I transcribed the interview recordings and conducted content analysis [82]. The coding process was facilitated by AQUAD (version: 7.4.1.2) [1]. First, the transcripts were open coded with a focus on how feedback tools influence understanding and reasoning about physical activity, what context the participants used for reasoning, interaction with visualization tools, how this understanding relates to ones goals and to changes in behavior, and barriers of current feedback use (example of coding??). Then those codes were clustered and organized into categories of state (current physical activity status), goal (personal objectives for using feedback tools), reasoning (how one makes sense of data patterns), insights and awareness (peoples understanding of their PA), behavior choice (choices about when and how to engage in physical activity) and emotion (what emotion could be evoked in the process). We then used the data to build an understanding of relationships between these concepts. This analysis resulted in the behavior feedback model illustrated in Figure 9.5.

Figure 9.5: Model of behavior feedback process.

State represents data with respect to current status that is collected and visualized with feedback tools; for example, the current activity level, progress during the day or the week, PA routines and change. Participants reported various data about state that they would read from feedback tools, including immediate measures in the moment (e.g., active minutes, heart rate, steps) and reflective progress measures (e.g., long-term trend, activity performance, calorie balance, daily and weekly progress towards goals, and sleep quality). Participants used both data summaries and detail views to access this information.

Personal **goals** could be short-term or long-term. As an example of a long-term goal, V7 was using Fitbit feedback to motivate himself to build regular gym routines that could fit in his current schedule. On the other hand, V6 used the feedback tool to track short-term daily and weekly step goals. Personal goals strongly influenced what participants expected to see about their state. V2 had to manage a health condition, so he focused most on sleep quality and resting heart rate. C8, who already had a regular exercise routine, mostly used feedback tools to track her exercise plan (two runs and two gym visits per week). In some cases, goals also influenced data collection. V9, hoping to know the impact of depression on his productivity, set up a daily self-report system to track his mood. V8 replaced her Fitbit device with a different model because she wanted to monitor her cardio status while exercising, a feature that was not possible with the first model.

Goals varied widely across our participants. Examples included progress checking (checking daily or weekly progress), in-the-moment monitoring (monitoring heart rate in cardio zone), exploration (exploring what exercise fits better), problem investigation (investigating sleep quality), or medical/physical condition management (managing diabetes). Ones goal may vary with age as well. For example, an older participant stated,

> "Fat burn, you can get how often I am doing, hitting the cardio level ... If I was younger that might be importantI think that probably for older people using the Fitbit, that probably the most important tool is to see that the improvement is there on a daily basis." (V6)

Personal goals motivate people to look at their data to gain **awareness**, and to **reason** about their data to gain **insights**, by posing and answering questions. I categorized three types of questions: (1) *What* ("What is the current status or performance?", "Do the data accurately reflect my situation?", "Have I done 3 runs

this week?", "What are the data patterns in a year/month/week/day?"), (2) *Why* ("Why do I have a trend like this?", "Why is the pattern on Friday night different?", "Why do I always see a spike in my data early in the morning?"), and (3) *How* ("How can I improve?", "How can I fit running into this weeks schedule?", "How can I customize my exercise goals?").

People performed a variety of tasks to seek answers to these questions. These tasks included making comparisons, mapping Fitbit data to calendar events, integrating data, looking up items, changing the timeline scale, counting, identifying regular patterns, identifying anomalies, observing overall trends, searching related domain knowledge, and exploring what-if experiments. (table of examples) During the interviews, participants were asked to reason about their Fitbit data patterns, e.g., peak values, repeated patterns, or anomalies. The tasks they performed most often were *comparisons* (19 out of 19 participants) and *activity mapping* (19 out of 19 participants), in which people related data patterns to activities that happened during the same time slot. Participants also *compared their progress* with their goals (18 out of 19 participants) or *baselines* (3 out of 19 participants), with the *historical performance*(3 out of 19 participants), or with *others* (8 out of 19 participants). By relating their activities to Fitbit data, participants could identify anomalies (10/10 in Visualization group and 5/9 in Control group) and reason about regular data patterns (5/10 in Visualization group and 3/9 in Control group). For better insights, V9 (?) expected to be able to do what-if exploration (specifically, adjust his bed time to investigate sleep qualities), but none of the feedback tools he used supported this task.

Insights derived from **reasoning** helped people to optimize their goals and inspired them to reflect on **behavior choices**. C9 customized her goal based on insights from the historical data,

> "I found 9k is reasonable for me to make it every day but I wanna make the goal everyday for a while before I upgrade it." (C9)

C4 found that keeping the same goal every day was not realistic for her, because it did not take into account the ups and downs of her energy level during the week. Most commonly, insights encouraged immediate action. Many participants reported they took an extra walk to meet their daily step goal (V3, V5, V6, C1, C4, C6, C7, C9).

> "I will find this silly to confess, if my steps are like 7,000, 8,000 at the end

> of the day, Ill jog in front of the television." (V3)

Insights also encouraged participants to adjust their exercise plan according to their progress and their schedule (V2, C3, C5, C8, C9).

> "So say I have done two gyms this week, check. And I have two runs this week, check. If I just did one, I should find some time to fix it." (C8)

Other participants identified motivational strategies that might work for them by reflecting on the historical data, especially when referring to their calendar schedules.

> "...because if I schedule it, like, I don't think an hour – it doesn't really take an hour, it takes an hour to do the workout, it takes ten minutes in the change room and then 20 minutes to get there, 20 minutes to get, like I need to have that full time, and so otherwise I schedule things too close." (C4)

The analysis also helped people select different types of exercises.

> "I don t know why, it s easy to fit in a walk or a run but with the same amount of time I won t do weights or push ups." (C7)

I also observed that reasoning about data could evoke **emotional engagement**. C8 found that her sleep data reflected how her son was doing over night. V6 could view the step data of her sister from the challenge feature (on the Fitbit application), and once identified some anomalies in her sisters data: it turned out her sister was sick. More common examples included the enjoyment of reminiscing and the satisfaction of keeping to an exercise plan (V3, V6, V9, C4, C5, C9). Such emotional engagement makes reasoning and reflecting an enjoyable process, and may itself become a goal for using a feedback tool.

The core of our feedback model is a loop. Acting on **behavior choices** leads to changes in **state**. When the feedback tool is used for ongoing monitoring and reflection, the user is continually **reasoning** about their state data to gain new **insights**, which continually leads to new **behavior choices** and revised personal **goals**. Meanwhile, the feedback loop is embedded in a personal context, reflecting how feedback-tool use varies with individual differences. Due to individual differences (e.g., physical conditions, domain knowledge, data analysis skills), peoples monitoring measures, goals, reasoning strategies, and influences on behaviors vary. According to this model, the design approach in this thesis was aimed to mainly tackle two problems:

- providing easy access and great degree of exposure to state data that are easy to perceive and comprehend
- offering contextual information in reasoing for awareness and insights that might lead to re-evaluate personal goals or/and behavior choices

9.5.4 Effects of the On-Calendar Visualization

In this section we discuss the effects of the on-calendar visualization with respect to the feedback model above.

Revealing state: Participants reported that the on-calendar visualization was good for showing overall trends, consistent repeated patterns and peak values.

> "This is great ... where Im at work its pretty clear peaks for my morning walk, my lunch time and my home break whereas with the kids its just sort of this little ... and when Im on my own there is peaks in intensity." (V3)

It could also be perceived simply with a glance (with line graphs):

> "this is all information I am just gleaning from glance and I like that a lot." (V2).

On the other hand, there were some limitations of the visualization to present information about state. The tool made it difficult to monitor specific measures in detail since values were not labeled on the graph (e.g., total daily steps). In addition, although step data reflected exercise generally, participants found it challenging to segment different activities:

> "It would be really nice if somewhere on here it would show when I played squash and then I could count the days since I had last played and that kind of thing" (V2).

These requirements might relate to participants personal goals, e.g., comparing with daily goals or tracking specific exercises. However, in the on-calendar application time-series fitness data were displayed with the time line aligning with calendar events, aimed to link the inferential informaiton to support the reasoning. In contrast, in default Fitbit application (see Figure **??**) daily goal was displayed as a reference line to compare with. That means, the context used in reasoning might be determined by

personal goals, and relevant context could be helpful to display with data. Although the month view with color coding could show a brief overview of the month (see Figure 4.1), participants barely use it because the lack of labels (e.g., "5,000 steps") or the overwhelming filled color. It brought up the data granularity issue in visualization design (see in Discussion).

Reasoning: A real strength of the approach was that participants could easily relate data to calendar activities to explain data patterns, especially with the week view. We coded the instances when participants were not sure or could not figure out the reason for a data pattern. In the control group this frequency dropped from 23 (first interview) to 15 (last interview), while in the visualization group it dropped further, from 29 to 8, suggesting that the calendar visualization was more helpful for reasoning than the baseline tools.

Participants were excited to tell us their new insights during their use. V1 mostly put work events on her calendar. She noticed she was actually more active when working in the office than working at home.V9 recalled a concert experience (an event on his calendar). He was surprised to see (via the Fitbit data) that almost half of the time was for the intermission. V9 was able to reconstruct non-documented events by viewing the on-calendar visualization:

> "I was just sitting around and then I went to yoga after that, so this probably indicates to say I went along to get some dinner and I ate the dinner and I walked to yoga. And then in the evening, I went for another walk" (V9).

The calendar visualization also helped participants to identify and reason about patterns and anomalies. V1 and V8 identified the intense spikes from the running competition in the city. V8 noticed data spikes during an exam on her calendar, which she attributed to Fitbit capturing the hand movement. V10 identified days he commuted by bike or car.

Six participants also reported that the on-calendar visualization helped their awareness, e.g.,

> "I'm surprised that I'm actually more active on the days where I have to go to work, as opposed to the day when I work from home. I thought I would be more active on those days [at home], but I take less breaks ... I find because I put [things], like on a Google calendar, the day, hours that

> I'm working from home. So, I can remember it too and then it shows that I'm less active ... that was surprise" (V1)
>
> "Whereas before it is kind of, you get so caught up in your life and your schedule and what you're doing, you might not even think that you haven't been out. But it's the awareness that helps." (V2)

Behavior choice: I found that the calendar visualization might not provide direct actionable insights to instruct users'behaviors; however, the influence on actions could be long term. Interestingly, V5 used the calendar visualization as a logging tool. She added calendar events expressly for the purpose of explaining the line graph patterns, e.g., "putting kids to bed", "dinner with family friends", etc. She reported that doing this helped her to recall and reason about her data patterns. The on-calendar application also helped participants to plan their exercises:

> "Well I look at weeks and then I think in terms of, instead of a daily thing ... the calendar has helped me focus on maybe a week or a month in advance and what I have to do". (V6)

C4 also mentioned when she did put an exercise plan in her schedule, she was more likely to follow that plan. When she was introduced the calendar in the end, she was exited to see it was what she needed,

> "Yeah, so then I've the flexibility of determining where it fits and then if I could just come back and say did run or whatever". (C4)

Affective engagement: The on-calendar application helped participants to recall their previous experiences, often evoking pleasant emotions.

> "You know what is this? [on Eastern Sunday] we were hiding eggs in the midnight for the kids. [laugh]"(V3)

With the Fitbit data and events on ones calendar, they may reconstruct their life and go through the past.

> "Yeah, I think so that will be cool, because then I could say like, the exact time when I met the person would be the time that I stop walking to talk to them and then so if I need to know it for like some reason the exact moment that I talk to them or something, maybe its useful for detective work ... I do like the ability to look at my history and this is such a cool feature like, I will be sad when this study ends ..." (V9).

The contextual framing of data (i.e., personal calendars) facilitated serendipitous exploration [108], eliciting emotion in reminiscence.

In contrast, emotion associated with feedback use for participants in the control group was limited to summary data: participants mentioned an emotional response to the smiley face that represented meeting their daily goals, weekly summary of weekly progress, or total steps in social challenges. However, we did not observe similar emotional mentions when they were looking at their raw Fitbit data or recalling related events.

9.5.5 Context for Reasoning

When participants were asked to investigate their data, they usually referred to contextual information. I coded all events when participants were trying to use contextual information to reason. The most frequently used types are shown in Table 9.2.

Most of the frequently used information could be found on the participants personal calendars. All participants in the Visualization group reported that they liked having their Fitbit data and life events aligned together a calendar. It was easy to access and also provided contextual information. Participants in both groups usually spontaneously brought up or referred to their personal calendar (17 out of 19 participants). Participants in the Control group usually brought up their phones or online calendar service (e.g., Google Calendar) while recalling the events. This observation confirmed our intuition that personal calendar events could provide relevant lift context to help people reason about their temporal fitness data. Moreover, the on-calendar visualization makes this contextual information easier to access. The control group had to bring up their calendar as a separate application or on a different device.

Context Information	Frequency (control group)	Frequencys (visualization group)
Schedules and holidays	2.6	3.1
Family events	2.4	1.0
Social activities	1.4	1.6
Location	0.7	1.7
Life routines not on the participants calendar	0.7	0.9

Table 9.2: Most frequently used contextual information for reasoning (normalized as frequency per participant).

In addition, the timeline of a day provided general context, for example the time to get up, to run for bus, to jog during lunch and to exercise in the evening, all of which could be quickly perceived with a glance. Especially with the week view of a calenar, participants could glance routines across the week.

> "I can see that I was active only in the middle of the day and without knowing what the numbers are just comparing this day or this day, I know that like lets say this was roughly five or six p.m., I went straight home, but this is all information Im just gleaning from glancing and I like that a lot." (V2)

Interestingly, C9 had even implemented a similar system (non-digital) 15 years ago. She logged exercise on a paper based wall calendar, and later put it into an Excel spreadsheet that could be visualized using charts.

Meanwhile, other relevant context information could not be found on calendars. Data granularity on activity level was neither available in the calendar visualization nor the native Fitbit application. Even with the time-varying step data, spikes were sometimes still not informative.

> "It's hard to make some days like this day ... kind of hard for me to tell sometimes even if its like kind of spikes and stuff." (V8)

Participants reported that it would be helpful to identify and then compare different activities (V2, V9, C4, C7). For example, C7 reported that she usually ran on a treadmill that could tell her the distance of every single run, which enabled her to compare with the previous runs and see if she improved.

> "I want to be able to look over time and say even though the steps go up and down, but my run has got longer or faster" (C4).

One challenge of the on-calendar visualization was that participants could not compare with baselines (e.g., daily goals, historical average or statistics representing performance in the population). V8, who was trying to investigate her sleep problem, reported that she expected to know how her data statistically fit in a larger population. In addition, participants felt they lacked domain knowledge to understand the measures (V1, V3, V6, V9, C1, C3, C4), for example how calories were calculated, what active minutes meant, and what was meant by heart rate zones.

9.5.6 Encouraging Ongoing Use

A feedback tool might aim to support ongoing tracking. However, ongoing use does not mean infinite use. In this thesis I define ongoing use as the long-term adoption towards reaching ones personal goals. After reaching those goals, ones curiosity or interest may drop, or it may be possible to maintain consistent status without assistance from feedback tools. However, I would like to prevent premature discontinuation of tool use due to other reasons. We explore this matter through our feedback model (Figure 9.5). Any factor that might prevent the loop moving forward is likely to discourage further use of a feedback tool.

The first inhibitor to ongoing use is when the representation of state cannot reflect ones goals. V8 and V2 reported that steps were lower priority for them. Instead, they were more interested in sleep quality and records of playing squash respectively; our current calendar visualization did not support showing these data. In addition, how much effort one needs to dedicate to accessing and managing the application matters. We found that people who had greater access to their phone used feedback tools primarily on the phone, while those who spent more time working with computers used feedback primarily on their browser. The on-calendar visualization was designed to provide easy and smooth access to view one's state information, by integrating the data stream to an existing frequently used tool. The ambient visualization on the backgroup can even minimize the effort of physical view switch, for example, bring up another device or application. The mash-up design approch might make it easily fit in one's already-formed routines. V8 had the habit of keeping all tool open (as browser tabs) while working with the computer, for example the digital calendar, the task management tool, and the email tool, etc. In the study, he added the calendar application to the tabs he always kept open on his browser. The session on-calendar feedback tool was usually last for hours, sometimes for days, possibly could enhance awareness of his fitness data that informed the present state. The study of Kim et al. also shows similar evidences [63]. Frustration arised when the data collection was not inline with people's goal (as they expected it should do) or data reliability was questioned.

Second, a large gap between the state and ones personal goal can lead to frustration. V5 reported that she barely used the feedback applications for a few weeks, because she felt she was much less active than before and she did not want to see the trends going down. If one could not meet ones goal for a long time, frustration might

make her/him drop the tool in the end. This indicate that simply comparing with a pre-set reference (daily goal of total steps) might not work in some cases.

Lack of support in reasoning might also prevent ongoing use. Domain knowledge might be required to interpret the data, or one may pick an inappropriate baseline to compare with (e.g., C4 found it was not realistic to compare herself with friends who were way more active than her.) This may limit meaningful insights and lead to feelings of powerlessness. Meanwhile, the common design strategy of representing the final results (e.g., status to represent active or inactive, or a pre-set heart rate zone) may be inadequate; this black box output can make people feel less involved, preventing ongoing use. Without knowing how calories were calculated, C4 chose to skip related features provided by Fitbit product. V2 could not make sense how heart zone correlated with steps, so he had to search online resources for more knowledge.

Lack of actionable insights for behavior change and planning could also break the loop. For example, V3 owned an exercise bike and knew she should exercise with it, but did not know where to start, leaving the bike in the basement for hanging clothes.

Meanwhile, emotional engagement is an interesting factor in ongoing use. In participants' experience, interest and engagement in using feedback tools like Fitbit wanes over time; there is an initial period of novelty and excitement, followed by some routine use, and an eventual drop-off in interest. The current study design was unable to assess whether deeper emotional engagement might encourage ongoing use, but this might be an interesting topic to investigate in the future. This topic is likely tightly interwoven with social engagement, a strong motivator for behavior change. One may invite a friend to exercise together (V8, C4), or hire a trainer as commitment (V5). In some cases, people kept using a feedback tool just because her/his friends stayed with the same one (C9).

9.5.7 Feeeback Tools for Seniors

It was intereting to include seniors in this study (V6 and C9). Suprisingly, they were early adopters of personal fitness trackers, with experience of more one fitness trackers before. Fitness feedback might be more demanding for them because they felt the increasing needs to manage with physical aging. They did careful pre-research before the purchase, comparing with different products based on their requirements. They were eager and also skilled to investigate the fitbit data. They tried to apply data analysis skills to learn about their data, for example, exporting raw data for statistic

analysis, designing their own analysis tools or visualization, etc. The intrinsic goal of coping with aging might be strong motivation for seniors who are actively use feedback tools for fitness.

More importantly, I found that social factors in feedabck tools might play a more important role for seniors than other age groups. They seems to be more engaged in social sharing, with contacts of online social networks, friends and families. They not only competed with peers, but also were eager to sharing their experience with families and friends. For that, feedback back for seniors might also need consider how to better facilitate these sharing needs.

The senior partcipants also mentioned that the needs and physical status of seniors might be different from other age groups. Possibly, providing more context other than that from personal calendars would be helpful, for example, heart rate data together with physical activities, or references of fitness for seniors.

9.6 Discussion

Based on the resutls, I reflect on the on-calendar design approach and feedback model. I also discuss the limitation of the field study.

9.6.1 Reflection on the On-Calendar Visualization

Overall, participants (including those in the control group) were very positive about integrating Fitbit data in a digital calendar; they found it easy to access and understand. The study also confirmed that personal digital calendars could provide rich contextual information for people to reason about their fitness data, and the on-calendar visualization made this information easy to access. I am also aware that digital calendars could not catch all context as demanded in reasoning. However, the study showed the extra context provided by calendars was helpful in data interpretation. Meanwhile, it also showed the frame of a digital calendar could provide general context for the recall.

The results showed that the usage of the on-calendar visualization highly depended on participants' existing information use behaviors, e.g., how often they use digital calendar to manage their schedules and what events were on their calendar. One of the participants was used to keeping browser tab (one of the tabs was Google calendar) open all the time while working. Accessing the on-calendar application (by

adding another brower tab) requires little effort for him, resulting the highest usage among all participants. He continued to use the application even after the study. Meanwhile, he maintained a dense calendar, with evenets about work, appointments, personal activities, and social events. His calendars could provide most of his life activities, so he enjoyed the re-construction process to put pieces of his life back together by refering the movement recorded by Fitbit trackers to calendar events. On the contrary, participants with very sparse calendar seems to use the application less, e.g., V4 only have a few social events marker on her calendar had the lowest usage from the logs. Thus, one's existing information use behaviors might be one of the factors that determines the effectiveness of this design. In this case, it could be how people use digital calendars in daily lives, how often they use these calendar tools, what events are recorded on the calendar, how dense (or sparse) of those events on the calendar, etc. Meanwhile, the characteristcs of the calendar influence attentional ambience of the visualization layers. When the calendar is dense and color coding has been used for calendar events, they try to make the visualization less salient but noticable, e.g., bring down the transparency of grey color. On the other hand, if the calendar is very sparse, ambient is not necessary because the interference is limited. In this case, they even change to a brighter color.

User goals would impact on the effectiveness of the design as well. Li et al. categoried use goals based on the stages of using personal informatics tools: *Discovery* and *Maintenance* [75]. People at each phase would interact with feedback tools differently. In the study, people at *Discovery* phase were trying to figure out the actionable fitness program or plan could fit best. They may be not fully aware of their energy and exercise patterns, or other factors than can influence their behaviors. Li et al. argued that providing and integrating multiple data source are important in this phase. Feedback tools in this case could support them to ask and answer questions based on their data, e.g, what is the difference of physical activities when work at home versus workplace, or what fitness plan could be managable. The context from personal schedules on a calendar is provided for this purpose. Different from that, people at *Maintenance* already had their regular exercise routines, e.g., three runs a week, brisk walks after super, etc. For that, they usually referred to feedback tools for a quick check and see if they finish the routine plan. The periodic nature of a calendar shows advantage of support these check-up tasks, especially in week view and month view. By comparing data patterns in the same time slot across days (or weeks), people could easily perceive the information even only with a glance.

Meanwhile, the on-calendar application was also used in some ways I did not expect. One example was using the calendar for logging: adding calendar events as a log to help recall and explain the Fitbit data at a later time. Interestingly, it also allowed participants to reflect on their past events and experiences by using the fitness data as context for understanding past calendar events and activities not documented on the calendar at all. Similarly, people could use it to plan fitness exercises that could better fit in ones schedule. The mapping between scheduled events on ones calendar and the integrated fitness data could inform the user about their actual performance during the fitness sessions and could make them more accountable to follow their fitness plan. It could also support reminiscing. With easily accessed contextual information, people could re-experience affective responses associated with special moments while recalling the past. For example, in many cases, the emotional reminiscing was related to the experience with family or friends, similar as the findings in [108]. Possibly, feedback tools could be used to enhance social bond or conversation, other than facilitating tracking and awareness. One of the participants was enjoying the on-calendar visualization as it can help him practice memory as linking the data of movement and his schedules. Some of these unanticipated uses turned out to be the most valuable attributes for some of the participants. Possibly, the feedback design need to go beyond single-purpose use and adapt to multi facet in everyday life.

In the field study I used time-varying fitness data as an example to explore the on-calendar design approch. If energy consumption data (used in early pilot studies Chapter 8) are applied here, personal calendar may include less related context to at-home activities. Domestic activities might not be marked on one's schedules because they are mostly daily routines, for example, cooking, watching TV with families. In this case, time diaries [38] would be able to provide more contextual information to reason about energy use at home. Other personal data may be also applicable to apply here as well, e.g., categorical data [105]. In that case, the visualization encoding might be different to make it ambient on a calendar. Participants also suggested to integrate more data source based on the calendar design, e.g., time-varying heart rate measure, calory intake, psychlogical measures etc. For that, appropriate data aggregation and segmentation may need to consider to avoid information overwhiming in visualization.

9.6.2 Understanding versus Coersion

This thesis explored a non-persuasive design approach in feedback design. Possibly, a persuasive design would be more effective to engage immediate behavior taking. Behavior change is usually considered as the goal of feedback design. However, long-term behavior change has been not confirmed in previous HCI research , and the impact of persuasive design on sustaining behavior change has barely investigated [69, 23, 111]. People know the world and themselves by accumulating knowledges. An action taking does not necessarily mean the message received or the action is sustained. Early literature review (Chapter 2) shows that current designs are mostly devised by system designers, who seem to decide what information to present and what metaphor should convey the message without considering the unique perspectives of individuals. However, the varying nature of baseline (Section 2.4.3) indicates that it is different to have the "standard" that can characterize everyone in all situations. Then who and how should define the "expected" behaviors to be coerced into?

According to The Transtheoretical Model (TTM) [93], behavior change is a long-term process through series of stages: precontemplation, contemplation, preparation, action and maintenance. Gathering information and understanding the behviors would be the first step when one is willing and intended to change. As learn about what happens, how it impact on oneself, and why it happens, one could transform one's goal into actionable plan for further action. The insights and knowledge perceived from information would also help to maintain the change. Studies shows understanding the data is still a major barrier when people use feedback tools in everyday practice [16, 103, 87].

Moreover, understanding behaviors is not simply cause-effect analysis. It has to consider the context the behavior is embeded into, physically, socially and culturally. Behavior choices might not always rational, but negotiable [103, 89]. That is, behavior change also need to consider people's "changing expectations and aspirations" [103]. That is, in some circumstances people may not accept the "best" rational behvior choice, but weight other factors more than these "expected" action, e.g., life comfort or convenience. Strengers et al. gave the example of household laundry that people did not adopt the "good" behaviors because of their hygiene standard, even though it was againt the goal of water conservation.

As participants reported in this study, they might not able to fulfill daily goals

every day (e.g., 10,000 steps a day). One may be out of town, or sick for a few days. People may have energy up and down, so a consistant goal might not be reasonable. To better fit into one's schedule, she might choose shorter vigorous excercises rather than longer light excercises. People might put the priority of kids or family events ahead of exercies. Feedback design need to support people to reflect on the flexibility and alternatives. In this case, providing context for the reflection is crutial for people to understand their own situation and choices.

Feedback tools in general are associated with certain numeric or non-numeric measures, and usually expect to bring the measures up (or down) by engaging people change behaviors. This view has been mostly emphasized in previous feedback practice. However, it is not the goal of feedback design. For example, in energy conservation, as people manage to reduce substantial energy use in the first a few months, would this trend keep going in the next a few months? Instead, the design goal of feedback technology should be helping people learn about their behaviors, understand their choices, and sustain the chanage, even without feedback support in the end.

9.6.3 Feedback Model and Related Models

The feedback model (Figure 9.5) characterizes the role that feedback tools can play in evoking behavior change. While this model emerged from our qualitative content analysis, it does bear some resemblance to other models in the literature.

Feedback designs are usually connected with behavior change models. A variety of behavior models have been studied in practice [4, 54, 29], but most of these focus on how to affect peoples motivation and attitude, and consequently influence behavior choices. However, the process of behavior change involves long-term learning and is on-going. Our interest with the feedback model is to explore how information design could facilitate this long-term process, for example, by making information more accessible and comprehensible. Although understanding ones data does not necessarily result in behavior change, we believe the role of feedback tools should also engage people in thinking and reflecting on information they receive; this may help people to set realistic and attainable goals, engaging them in the process.

As such, the feedback model enables designers to reflect on a non-persuasive approach. Tools that facilitate the reasoning process rather than enforcing behavior change might be one step forward towards encouraging people to change their be-

haviour on their own in a long-lasting way. Such tools need to present information that can be accessed easily and that reflect one's goal appropriately. Another implication is design for emotional engagement. Rather than being task-oriented, designers could consider designing for serendipitous information exploration.

In relation to more general models, the Technology Acceptance Model [13] is a well-known model of investigating system adoption; however, it is not specific to feedback tools and does not consider the influence of technology on behaviors outside of tool use itself. A closely related but more specific model is the Promoter-Inhibitor Motivation Model (PIMM) [101]. PIMM models factors that promote and inhibit use of casual visualizations that people encounter in everyday life. However, the feedback model in the work captures a specific case: on-going use of a feedback visualization to learn about and influence personal behavior. As such, while many of the influencing factors from PIMM still apply to this context, PIMM does not capture the role of insight in effecting behavior change, nor the subsequent (circular) effects on goals and motivations for using a feedback tool.

Recent studies showed evidences that were inline with the feedback model. In a study of wearable devices, kim et al. categorized the stages that people interact with wearable devices [63]. In the stage of "initiation & experimentation" people would match their goal with what the device could do (i.e., represent their "state"). In "intensifying & integration" stage, the main role of these devices was to accumulate knowledge and develop actionable insights. Any connection between components in the feedback model would prevent the further use of wearable devices, e.g., unreliability of data collection, lack of ways or context for reasoning, lack of actionable insights, etc. The study of prsonal productivity feedback indicated the influence of emotional factors in feedback design [67]

The feedback model is certainly constrained by the scale and nature of our study. However, I believe it provides a starting point for thinking about how design characteristics might influence feedback tool use, including the likelihood of ongoing adoption and behavior change. Nonetheless, I fully expect that it will be revised with more input in future work.

9.6.4 Design implication

Traditionally, success of feedback tools mostly defined based on "behavior change". However, behavior change is a long-term process and interfered by many other as-

pects in one's life. Short-term influence of using feedback tools could be easily over estimated. Instead, researchers and designers may need to view information use as a holistic eco-system that is habitated through one's everday life, considering personal, social and cultrual constraints [103]. When a new tool was introduced, it needs to fit into this eco-system first, and then impact on this system, which involve ongoing interacting, learning and adapting as an ongoing process.

First of all, feedback tools need to present information that is reliable and can be accessed easily. The "novelty" of designs should not go beyond the current information use habits. Instead, designers need to consider the cost of information retrieval and management. Otherwise, the cost of information seeking would quickly dilute the initial curiosity and interest. Meanwhile, the presented information needs to reflect one's goal and impact of behaviors appropriately. Goals might be dynamic even with the same person considering her varying situation (e.g., energy level), and customized for different groups (e.g., age groups).

It is possible that people would take recommendated actions even without undering standing the behavior itself. Behavior change might be persuaded in a short period with proper design strategies. However, being able to reason and understand the cause and impact of behavior choices may be more important to enhance people's feel of control of their life. Designers might need to see behavior taking as negotiable practices (Section 9.6.2), in which people learn about their flexibility and alternatives. Only with that, the knowledge could truely become part of one's own. Especially in feedback design, information and interaction designs needs to facilitate this long-term learning and negotiating process, for example, by providing sources of contexual information for reasoning (Section 2.4.2).

Encouragingly, a recent design, Casalendar [81], adopted the idea of using personal calendar as media to view and contraol smart home environment, e.g., temprature, lighting, appliances, etc. The studies showed that viewing calendar context sid by side with the confiuration and measures help participants easily identify and reason about anomalies.

Another implication is deisgn for emotional engagement. From the study, I found that feedback tools could be a media that engage people emotionally. It offers people oppertunities to reminisce on personal or family trace, encourage social conversation, interaction with friends and family. Rather than task-oriented information deisgn, designers could also consider the design strategy for seredipitious information exploration [108]. The pleasure and satisfaction from that might inspire or enforce interest

and curiosity that inhabit the ongoing use as well.

9.7 Limitation of the study

Since the application was independent (e.g., not implemented within google calendar or iCal) and had limited features (e.g., no custom coloring of calendar events), its use might have been constrained. At least one participant reported trying to manage both Google Calendar and the on-calendar application. Participants expected that the data could be displayed in their own usual calendar.

In the field study, I recruited existing Fitbit users in hopes that they would use feedback tools on a regular basis. These people had previous experience using Fitbits feedback tools, and may therefore react differently to the on-calendar visualizations than a more general population. Additionally, I did not control for use of Fitbits feedback tools; it would have been difficult to constrain or track peoples use of the Fitbit application except through unreliable self-reports. In addition, the field study had a small scale and only focused on physical activity. While I anticipate that on-calendar visualizations could be used to display any sort of personal quantitative feedback data (e.g., heart rate, blood pressure, resource use), future research is needed to understand user needs for other application domains. Meanwhile, the visualization layout was designed for desktop use, so future investigations could examine a version customized for mobile devices.

This study is a starting point for exploring how to integrate personal data within a digital calendar. It suggests that designers may wish to further experiment with this and other non-persuasive approaches, and also give attention to engaging ongoing use. Making contextual information easily accessible and blending feedback into currently used tools are promising design directions.

Chapter 10

Future work

There would be considerable work to be done in the future in regards to feedback design and evaluation. In the study, personal data (i.e., Fitbit data and home energy data) were displayed as continious time-series data. Future design could explore data graularity in visualization as well. For example, aggregated summary to inform the performance compared to daily goal. Such data could also visualized by abstract graphics, for example, color steps or icons. The aggregation could be also based on types of activities or locations. Visualization design could also experiment with multiple data sources, for example, displaying energy use from different rooms or appliances.

Another interesting direction would be visualization designs for social sharing. I did not investigate the design component for emotional engagement according to the feedback model. Participants in the study brought up some interesting use for social sharing and reminiscing. Future designs could explore design features that faciliate social and emotional factors. People may expect to compare with others' data on their display, or share an interesting moment other family members would be interested. It could also be used as a tool to encourage social conversations.

I also would like to see this design could be implemented in native calendar applications, e.g., Google calendar, iCal or Outlook, with customized personal data stream on the addtional visualization layer. In that case, people could get better exposure with their data, with which the imapct could be better observed and studied.

The proposed design in this thesis is to add a visualization layer into one's existing tools, specifically personal digital calendar in my studeis. The design approach could be totally reversed; context from calendar schedules could be added into other data application as well, e.g., integrating calendar events into energy billing report. Studies

showed people currently still have problems to understand their energy bill because of the lack of context **??**. For that, contexual information from one's calendar could be helpful. Designers could explore how to aggregate energy data in the visualization and how to identify and include relevant context from other data sources (e.g., one's calendar events) that can help the data interpretation.

My previous evaluation mainly focused on personal fitness. Future studies could explore the design approach in other aspects of everyday life use, for example, home energy conservation, health management, etc. It may also applied to public education for sustainability [104]. Possible, studies could also investigate its application in perfessional use. For example, building managers could use to map time-series utlity usage with respect to maintenance logs. Meanwhile, new evaluation methods could be explored, e.g., how to capture awareness in the field, how to accurately collect quantitative data from participants, etc.

Chapter 11

Conclusion

Providing context for reasoning and engaging long-term use are current challenges for design of feedback technologies.The goal of this thesis was to exploare a mash-up design approach to visualize personal data, aimed to provide contextual information to support reasoning data patterns and facilitate easy access that might encourage ongoing use.

I systematically reveiw related academia work of personal visualization deisgn, which helped to identify design dimensions and challenges of personal visualization. Considering the design gaps in previous research and practice, I proposed to integrate personal data (specifically, fitness data and household energy consumption data in this thesis) as an addtional visualization layer on a digtital calendar, because personal digital calendar could provide information about one's daily activities and as well it was a commonly used tool.

To further investigate the on-calendar approach, I first conducted a viability lab study to evaluate the interference and perception of design alternatives of on-calendar visualization. The lab study confirmed that the on-calendar visualiztion does not interfere with regular calendar use tasks with proper design choices, and meanwhile it is perceptable and comprehensable. The viability study helped to narrow down design alternatives, with which I implemented the design approach as a web-based application, synchronizing calendar data with Google calendar and fetching personal data either from Fitbit API or data stream of household utility meter. Second, I deployed the early prototype in two case studies as pilot deployments: household energy conservation and personal fitness. Feedback from the pilot studies suggested a revision of the prototype, particularly improving the consistency with exiting calendar application (Google calendar in this case). After that, I conducted an eight-week field

study to investigate how people use the application in everyday life. In the field study, I applied the on-calendar visualization in the area of personal fitness, by linking the data with Fitbit trackers. As mainly focusing on qualitative data from the results, I primarily employed qualitative method to code, categorize and analyze the field study data. Based on the data, I derived a feedback model to identify design components in feedback design with respect to ongoing use.

The main contribution of this thesis includes: conducting the first systemactic literature review of Personal Visualization and Personal Visual Analytics; proposing the on-calendar design that integrate personal data on people's personal digital calenar to provide context for reasoning and support easy access for ongoing use; extending the definition of ambience from spatial location to attentional demand; conducting quatitative lab experiments to evaluate visual interference and perceptability of design alternatives of on-calendar visualization; implementing and deploying the proposed approch in a longitudinal field study; and developing a new feedback model of design compoments to inspect ongoing factors in feedback designs.

In this research, I investigate the reflective approch of feedback design, integrating personal visualization in people's information eco-system (personal digital calendar in this case). It shows a different respect in feedback design that people's exiting information use behavior need to be considered when tools are designed for ongoing use. I hope this study will help designers reflect on non-persuasive approaches to feedback tool design and considering people's existing information use habits to encourage ongoing use.

Appendix A

Lab Experiemtn Tasks in Viability Study

Appendix B

Post-Experiment Questinnaire in Lab Experiments

Appendix C

Interview Outlines in Pilot Study

Appendix D

Screenshots of On-Calendar Application (Final Version)

Appendix E

Background Questionnaire in Field Study

Appendix F

Interview outlines in Pilot Study

Appendix G

Screenshots of Fitbit Online Tool

Appendix H

International Physical Activity Questionnaire (IPAQ)

INTERNATIONAL PHYSICAL ACTIVITY QUESTIONNAIRE

We are interested in finding out about the kinds of physical activities that people do as part of their everyday lives. The questions will ask you about the time you spent being physically active in the **last 7 days**. Please answer each question even if you do not consider yourself to be an active person. Please think about the activities you do at work, as part of your house and yard work, to get from place to place, and in your spare time for recreation, exercise or sport.

Think about all the **vigorous** and **moderate** activities that you did in the **last 7 days**. **Vigorous** physical activities refer to activities that take hard physical effort and make you breathe much harder than normal. **Moderate** activities refer to activities that take moderate physical effort and make you breathe somewhat harder than normal.

PART 1: JOB-RELATED PHYSICAL ACTIVITY

The first section is about your work. This includes paid jobs, farming, volunteer work, course work, and any other unpaid work that you did outside your home. Do not include unpaid work you might do around your home, like housework, yard work, general maintenance, and caring for your family. These are asked in Part 3.

1. Do you currently have a job or do any unpaid work outside your home?

   ☐ Yes

   ☐ No → ***Skip to PART 2: TRANSPORTATION***

The next questions are about all the physical activity you did in the **last 7 days** as part of your paid or unpaid work. This does not include traveling to and from work.

2. During the **last 7 days**, on how many days did you do **vigorous** physical activities like heavy lifting, digging, heavy construction, or climbing up stairs **as part of your work**? Think about only those physical activities that you did for at least 10 minutes at a time.

   _____ **days per week**

   ☐ No vigorous job-related physical activity → ***Skip to question 4***

3. How much time did you usually spend on one of those days doing **vigorous** physical activities as part of your work?

   _____ **hours per day**
   _____ **minutes per day**

4. Again, think about only those physical activities that you did for at least 10 minutes at a time. During the **last 7 days**, on how many days did you do **moderate** physical activities like carrying light loads **as part of your work**? Please do not include walking.

   _____ **days per week**

   ☐ No moderate job-related physical activity → ***Skip to question 6***

5. How much time did you usually spend on one of those days doing **moderate** physical activities as part of your work?

_____ **hours per day**
_____ **minutes per day**

6. During the **last 7 days**, on how many days did you **walk** for at least 10 minutes at a time **as part of your work**? Please do not count any walking you did to travel to or from work.

_____ **days per week**

☐ No job-related walking → ***Skip to PART 2: TRANSPORTATION***

7. How much time did you usually spend on one of those days **walking** as part of your work?

_____ **hours per day**
_____ **minutes per day**

PART 2: TRANSPORTATION PHYSICAL ACTIVITY

These questions are about how you traveled from place to place, including to places like work, stores, movies, and so on.

8. During the **last 7 days**, on how many days did you **travel in a motor vehicle** like a train, bus, car, or tram?

_____ **days per week**

☐ No traveling in a motor vehicle → ***Skip to question 10***

9. How much time did you usually spend on one of those days **traveling** in a train, bus, car, tram, or other kind of motor vehicle?

_____ **hours per day**
_____ **minutes per day**

Now think only about the **bicycling** and **walking** you might have done to travel to and from work, to do errands, or to go from place to place.

10. During the **last 7 days**, on how many days did you **bicycle** for at least 10 minutes at a time to go **from place to place**?

_____ **days per week**

☐ No bicycling from place to place → ***Skip to question 12***

11. How much time did you usually spend on one of those days to **bicycle** from place to place?

_____ **hours per day**
_____ **minutes per day**

12. During the **last 7 days**, on how many days did you **walk** for at least 10 minutes at a time to go **from place to place**?

_____ **days per week**

☐ No walking from place to place → ***Skip to PART 3: HOUSEWORK, HOUSE MAINTENANCE, AND CARING FOR FAMILY***

13. How much time did you usually spend on one of those days **walking** from place to place?

_____ **hours per day**
_____ **minutes per day**

PART 3: HOUSEWORK, HOUSE MAINTENANCE, AND CARING FOR FAMILY

This section is about some of the physical activities you might have done in the **last 7 days** in and around your home, like housework, gardening, yard work, general maintenance work, and caring for your family.

14. Think about only those physical activities that you did for at least 10 minutes at a time. During the **last 7 days**, on how many days did you do **vigorous** physical activities like heavy lifting, chopping wood, shoveling snow, or digging **in the garden or yard**?

_____ **days per week**

☐ No vigorous activity in garden or yard → ***Skip to question 16***

15. How much time did you usually spend on one of those days doing **vigorous** physical activities in the garden or yard?

_____ **hours per day**
_____ **minutes per day**

16. Again, think about only those physical activities that you did for at least 10 minutes at a time. During the **last 7 days**, on how many days did you do **moderate** activities like carrying light loads, sweeping, washing windows, and raking **in the garden or yard**?

_____ **days per week**

☐ No moderate activity in garden or yard → ***Skip to question 18***

17. How much time did you usually spend on one of those days doing **moderate** physical activities in the garden or yard?

_____ **hours per day**
_____ **minutes per day**

18. Once again, think about only those physical activities that you did for at least 10 minutes at a time. During the **last 7 days**, on how many days did you do **moderate** activities like carrying light loads, washing windows, scrubbing floors and sweeping **inside your home**?

_____ **days per week**

☐ No moderate activity inside home → ***Skip to PART 4: RECREATION, SPORT AND LEISURE-TIME PHYSICAL ACTIVITY***

19. How much time did you usually spend on one of those days doing **moderate** physical activities inside your home?

_____ **hours per day**
_____ **minutes per day**

PART 4: RECREATION, SPORT, AND LEISURE-TIME PHYSICAL ACTIVITY

This section is about all the physical activities that you did in the **last 7 days** solely for recreation, sport, exercise or leisure. Please do not include any activities you have already mentioned.

20. Not counting any walking you have already mentioned, during the **last 7 days**, on how many days did you **walk** for at least 10 minutes at a time **in your leisure time**?

_____ **days per week**

☐ No walking in leisure time → ***Skip to question 22***

21. How much time did you usually spend on one of those days **walking** in your leisure time?

_____ **hours per day**
_____ **minutes per day**

22. Think about only those physical activities that you did for at least 10 minutes at a time. During the **last 7 days**, on how many days did you do **vigorous** physical activities like aerobics, running, fast bicycling, or fast swimming **in your leisure time**?

_____ **days per week**

☐ No vigorous activity in leisure time → ***Skip to question 24***

23. How much time did you usually spend on one of those days doing **vigorous** physical activities in your leisure time?

_____ **hours per day**
_____ **minutes per day**

24. Again, think about only those physical activities that you did for at least 10 minutes at a time. During the **last 7 days**, on how many days did you do **moderate** physical activities like bicycling at a regular pace, swimming at a regular pace, and doubles tennis **in your leisure time**?

_____ **days per week**

☐ No moderate activity in leisure time → ***Skip to PART 5: TIME SPENT SITTING***

25. How much time did you usually spend on one of those days doing **moderate** physical activities in your leisure time?
_____ **hours per day**
_____ **minutes per day**

PART 5: TIME SPENT SITTING

The last questions are about the time you spend sitting while at work, at home, while doing course work and during leisure time. This may include time spent sitting at a desk, visiting friends, reading or sitting or lying down to watch television. Do not include any time spent sitting in a motor vehicle that you have already told me about.

26. During the **last 7 days**, how much time did you usually spend **sitting** on a **weekday**?

_____ **hours per day**
_____ **minutes per day**

27. During the **last 7 days**, how much time did you usually spend **sitting** on a **weekend day**?

_____ **hours per day**
_____ **minutes per day**

This is the end of the questionnaire, thank you for participating.

Bibliography

[1] Analysis of qualitative data (aquad 7). Accessed: 2015-09-20.

[2] International Physical Activity Questionnaire. Accessed: 2015-09-17.

[3] *BELIV '08: Proceedings of the 2008 Workshop on BEyond Time and Errors: Novel evaLuation Methods for Information Visualization*. ACM, New York, NY, USA, 2008.

[4] Wokje Abrahamse, Linda Steg, Charles Vlek, and Talib Rothengatter. A review of intervention studies aimed at household energy conservation. *Journal of Environmental Psychology*, 25(3):273–291, September 2005.

[5] A. Ahtinen, E. Mattila, A. Vaatanen, L. Hynninen, J. Salminen, E. Koskinen, and K. Laine. User experiences of mobile wellness applications in health promotion: User study of Wellness Diary, Mobile Coach and SelfRelax. In *3rd International Conference on Pervasive Computing Technologies for Healthcare, 2009. PervasiveHealth 2009*, pages 1–8, 2009.

[6] David Allen, Stan Karanasios, and Mira Slavova. Working with activity theory: Context, technology, and information behavior. *Journal of the American Society for Information Science and Technology*, 62(4):776–788, 2011.

[7] Adrienne Andrew, Gaetano Borriello, and James Fogarty. Toward a Systematic Understanding of Suggestion Tactics in Persuasive Technologies. In Yvonne de Kort, Wijnand IJsselsteijn, Cees Midden, Berry Eggen, and B. J. Fogg, editors, *Persuasive Technology*, number 4744 in Lecture Notes in Computer Science, pages 259–270. Springer Berlin Heidelberg, April 2007.

[8] Natalia Andrienko. *Exploratory analysis of spatial and temporal data*. Springer, Berlin; New York, 2006.

[9] Bruno Antunes, Tiago Guerreiro, and Daniel Gonalves. This Just In! Your Life in the Newspaper. In Tom Gross, Jan Gulliksen, Paula Kotz, Lars Oestreicher, Philippe Palanque, Raquel Prates, and Marco Winckler, editors, *Human-Computer Interaction INTERACT 2009*, volume 5727 of *Lecture Notes in Computer Science*, pages 279–292. Springer Berlin / Heidelberg, 2009.

[10] Ernesto Arroyo, Leonardo Bonanni, and Ted Selker. Waterbot: exploring feedback and persuasive techniques at the sink. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '05, pages 631–639, New York, NY, USA, 2005. ACM.

[11] Ernesto Arroyo, Leonardo Bonanni, and Nina Valkanova. Embedded interaction in a water fountain for motivating behavior change in public space. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 685–688, New York, NY, USA, 2012. ACM.

[12] Susan J. Ashford and L.L. Cummings. Feedback as an individual resource: Personal strategies of creating information. *Organizational Behavior and Human Performance*, 32(3):370 – 398, 1983.

[13] Akhilesh Bajaj and Sarma R Nidumolu. A feedback model to understand information system usage. *Information & Management*, 33(4):213–224, March 1998.

[14] Gabriel Barata, Hugo Nicolau, and Daniel Gon\ccalves. AppInsight: what have I been doing? In *Proceedings of the International Working Conference on Advanced Visual Interfaces*, AVI '12, pages 465–472, New York, NY, USA, 2012. ACM.

[15] Mary Barreto, Evangelos Karapanos, and Nuno Nunes. Why Don'T Families Get Along with Eco-feedback Technologies?: A Longitudinal Inquiry. In *Proceedings of the Biannual Conference of the Italian Chapter of SIGCHI*, CHItaly '13, pages 16:1–16:4, New York, NY, USA, 2013. ACM.

[16] L. Bartram. Design Challenges and Opportunities for Eco-Feedback in the Home. *IEEE Computer Graphics and Applications*, 35(4):52–62, July 2015.

[17] Lyn Bartram, Billy Cheung, and Maureen Stone. The Effect of Colour and Transparency on the Perception of Overlaid Grids. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):1942–1948, December 2011.

[18] Lyn Bartram, Johnny Rodgers, and Rob Woodbury. Smart homes or smart occupants? supporting aware living in the home. In *Proceedings of the 13th IFIP TC 13 international conference on Human-computer interaction - Volume Part II*, INTERACT'11, pages 52–64, Berlin, Heidelberg, 2011. Springer-Verlag.

[19] 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, New York, NY, USA, 2012. ACM.

[20] Eric P.S. Baumer, Sherri Jean Katz, Jill E. Freeman, Phil Adams, Amy L. Gonzales, John Pollak, Daniela Retelny, Jeff Niederdeppe, Christine M. Olson, and Geri K. Gay. Prescriptive Persuasion and Open-ended Social Awareness: Expanding the Design Space of Mobile Health. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*, CSCW '12, pages 475–484, New York, NY, USA, 2012. ACM.

[21] D. Baur, F. Seiffert, M. Sedlmair, and S. Boring. The Streams of Our Lives: Visualizing Listening Histories in Context. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1119 –1128, December 2010.

[22] Elzbieta Biernat, Romuald Stupnicki, Bartlomiej Lebiedzinski, and Lidia Janczewska. Assessment of Physical Activity by Applying IPAQ Questionnaire. *Physical Education and Sport*, 52:46–52, 2008.

[23] Hronn Brynjarsdottir, Maria H\aakansson, James Pierce, Eric Baumer, Carl DiSalvo, and Phoebe Sengers. Sustainably unpersuaded: how persuasion narrows our vision of sustainability. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems*, CHI '12, pages 947–956, New York, NY, USA, 2012. ACM.

[24] N. Cawthon and A. V. Moere. The Effect of Aesthetic on the Usability of Data Visualization. In *Information Visualization, 2007. IV '07. 11th International Conference*, pages 637–648, July 2007.

[25] Emiko Charbonneau, Andrew Miller, and Joseph J. LaViola, Jr. Teach Me to Dance: Exploring Player Experience and Performance in Full Body Dance Games. In *Proceedings of the 8th International Conference on Advances in Computer Entertainment Technology*, ACE '11, pages 43:1–43:8, New York, NY, USA, 2011. ACM.

[26] Marshini Chetty, David Haslem, Andrew Baird, Ugochi Ofoha, Bethany Sumner, and Rebecca Grinter. Why is my internet slow?: making network speeds visible. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 1889–1898, New York, NY, USA, 2011. ACM.

[27] Meng-Chieh Chiu, Shih-Ping Chang, Yu-Chen Chang, Hao-Hua Chu, Cheryl Chia-Hui Chen, Fei-Hsiu Hsiao, and Ju-Chun Ko. Playful bottle: a mobile social persuasion system to motivate healthy water intake. In *Proceedings of the 11th international conference on Ubiquitous computing*, Ubicomp '09, pages 185–194, New York, NY, USA, 2009. ACM.

[28] S. Consolvo, P. Klasnja, D. W McDonald, D. Avrahami, J. Froehlich, L. LeGrand, R. Libby, K. Mosher, and J. A Landay. Flowers or a robot army?: encouraging awareness & activity with personal, mobile displays. In *Proceedings of the 10th international conference on Ubiquitous computing*, pages 54–63, 2008.

[29] Sunny Consolvo, Katherine Everitt, Ian Smith, and James A. Landay. Design requirements for technologies that encourage physical activity. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '06, pages 457–466, New York, NY, USA, 2006. ACM.

[30] 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, New York, NY, USA, 2008. ACM.

[31] Michael Correll, Danielle Albers, Steven Franconeri, and Michael Gleicher. Comparing averages in time series data. In *Proceedings of the SIGCHI Con-*

ference on Human Factors in Computing Systems, CHI '12, pages 1095–1104, New York, NY, USA, 2012. ACM.

[32] Enrico Costanza, Sarvapali D. Ramchurn, and Nicholas R. Jennings. Understanding domestic energy consumption through interactive visualisation: a field study. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, UbiComp '12, pages 216–225, New York, NY, USA, 2012. ACM.

[33] Ron Samuel Dembo and Lyn Bartram. System and Method for Generating, Processing and Displaying Data Relating to Consumption Data with an Application, March 2013. Patent WO/2013/037051.

[34] Carl DiSalvo, Phoebe Sengers, and Hrnn Brynjarsdttir. Mapping the landscape of sustainable HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 1975–1984, New York, NY, USA, 2010. ACM.

[35] M. Dork, D. Gruen, C. Williamson, and S. Carpendale. A Visual Backchannel for Large-Scale Events. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1129 –1138, December 2010.

[36] G. Draper and R. Riesenfeld. Who Votes For What? A Visual Query Language for Opinion Data. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1197 –1204, December 2008.

[37] Honglu Du, G. Michael Youngblood, and Peter Pirolli. Efficacy of a Smartphone System to Support Groups in Behavior Change Programs. In *Proceedings of the Wireless Health 2014 on National Institutes of Health*, WH '14, pages 3:1–3:8, New York, NY, USA, 2014. ACM.

[38] Kajsa Ellegrd and Jenny Palm. Visualizing energy consumption activities as a tool for making everyday life more sustainable. *Applied Energy*, 88(5):1920–1926, May 2011.

[39] Thomas Erickson, Ming Li, Younghun Kim, Ajay Deshpande, Sambit Sahu, Tian Chao, Piyawadee Sukaviriya, and Milind Naphade. The dubuque electricity portal: evaluation of a city-scale residential electricity consumption feedback system. In *Proceedings of the 2013 ACM annual conference on Human factors*

in computing systems, CHI '13, pages 1203–1212, New York, NY, USA, 2013. ACM.

[40] Thomas Erickson, Mark Podlaseck, Sambit Sahu, Jing D. Dai, Tian Chao, and Milind Naphade. The dubuque water portal: evaluation of the uptake, use and impact of residential water consumption feedback. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 675–684, New York, NY, USA, 2012. ACM.

[41] Chloe Fan, Jodi Forlizzi, and Anind K. Dey. A spark of activity: exploring informative art as visualization for physical activity. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, UbiComp '12, pages 81–84, New York, NY, USA, 2012. ACM.

[42] Siamak Faridani, Ephrat Bitton, Kimiko Ryokai, and Ken Goldberg. Opinion space: a scalable tool for browsing online comments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 1175–1184, New York, NY, USA, 2010. ACM.

[43] BJ Fogg. A behavior model for persuasive design. In *Proceedings of the 4th International Conference on Persuasive Technology*, pages 1–7, Claremont, California, 2009. ACM.

[44] Thomas Fritz, Elaine M. Huang, Gail C. Murphy, and Thomas Zimmermann. Persuasive Technology in the Real World: A Study of Long-term Use of Activity Sensing Devices for Fitness. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, pages 487–496, New York, NY, USA, 2014. ACM.

[45] Jon Froehlich. Promoting Energy Efficient Behaviors in the Home through Feedback: The Role of HumanComputer Interaction. In *HCIC 2009 Winter Workshop*, 2009.

[46] Jon Froehlich, Tawanna Dillahunt, Predrag Klasnja, Jennifer Mankoff, Sunny Consolvo, Beverly Harrison, and James A. Landay. UbiGreen: investigating a mobile tool for tracking and supporting green transportation habits. In *Proceedings of the 27th international conference on Human factors in computing systems*, CHI '09, pages 1043–1052, New York, NY, USA, 2009. ACM.

[47] Jon Froehlich, Leah Findlater, and James Landay. The design of eco-feedback technology. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 1999–2008, New York, NY, USA, 2010. ACM.

[48] Jon Froehlich, Leah Findlater, Marilyn Ostergren, Solai Ramanathan, Josh Peterson, Inness Wragg, Eric Larson, Fabia Fu, Mazhengmin Bai, Shwetak Patel, and James A. Landay. The design and evaluation of prototype eco-feedback displays for fixture-level water usage data. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 2367–2376, New York, NY, USA, 2012. ACM.

[49] L. Grammel, M. Tory, and M. Storey. How Information Visualization Novices Construct Visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):943 –952, December 2010.

[50] Michael Haller, Christoph Richter, Peter Brandl, Sabine Gross, Gerold Schossleitner, Andreas Schrempf, Hideaki Nii, Maki Sugimoto, and Masahiko Inami. Finding the Right Way for Interrupting People Improving Their Sitting Posture. In Pedro Campos, Nicholas Graham, Joaquim Jorge, Nuno Nunes, Philippe Palanque, and Marco Winckler, editors, *Human-Computer Interaction INTERACT 2011*, volume 6947 of *Lecture Notes in Computer Science*, pages 1–17. Springer Berlin / Heidelberg, 2011.

[51] Martin Hand, Elizabeth Shove, and Dale Southerton. Explaining Showering: a Discussion of the Material, Conventional, and Temporal Dimensions of Practice. *Sociological Research Online*, 10(2), 2005.

[52] Daniel Harrison, Paul Marshall, Nadia Bianchi-Berthouze, and Jon Bird. Activity Tracking: Barriers, Workarounds and Customisation. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, UbiComp '15, pages 617–621, New York, NY, USA, 2015. ACM.

[53] William R. Hazlewood, Erik Stolterman, and Kay Connelly. Issues in evaluating ambient displays in the wild: two case studies. In *Proceedings of the 2011 annual conference on Human factors in computing systems*, CHI '11, pages 877–886, New York, NY, USA, 2011. ACM.

[54] Helen Ai He, Saul Greenberg, and Elaine M. Huang. One size does not fit all: applying the transtheoretical model to energy feedback technology design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 927–936, New York, NY, USA, 2010. ACM.

[55] Jeffrey Heer, Nicholas Kong, and Maneesh Agrawala. Sizing the horizon: the effects of chart size and layering on the graphical perception of time series visualizations. In *Proceedings of the 27th international conference on Human factors in computing systems*, CHI '09, pages 1303–1312, New York, NY, USA, 2009. ACM.

[56] David M. Herold and Martin M. Greller. Feedback: The Definition of a Construct. *The Academy of Management Journal*, 20(1):142–147, 1977.

[57] Stephen S. Intille, Emmanuel Munguia Tapia, John Rondoni, Jennifer Beaudin, Chuck Kukla, Sitij Agarwal, Ling Bao, and Kent Larson. Tools for Studying Behavior and Technology in Natural Settings. In Anind K. Dey, Albrecht Schmidt, and Joseph F. McCarthy, editors, *UbiComp 2003: Ubiquitous Computing*, number 2864 in Lecture Notes in Computer Science, pages 157–174. Springer Berlin Heidelberg, October 2003. DOI: 10.1007/978-3-540-39653-6_13.

[58] M. Jett, K. Sidney, and G. Blmchen. Metabolic equivalents (METS) in exercise testing, exercise prescription, and evaluation of functional capacity. *Clinical Cardiology*, 13(8):555–565, August 1990.

[59] Vaiva Kalnikaite, Abigail Sellen, Steve Whittaker, and David Kirk. Now let me see where i was: understanding how lifelogs mediate memory. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 2045–2054, New York, NY, USA, 2010. ACM.

[60] Victor Kaptelinin and Bonnie A. Nardi. *Acting with Technology: Activity Theory and Interaction Design*. The MIT Press, August 2009.

[61] Evangelos Karapanos, John Zimmerman, Jodi Forlizzi, and Jean-Bernard Martens. User Experience over Time: An Initial Framework. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pages 729–738, New York, NY, USA, 2009. ACM.

[62] Julie A. Kientz, Rosa I. Arriaga, and Gregory D. Abowd. Baby steps: evaluation of a system to support record-keeping for parents of young children. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pages 1713–1722, New York, NY, USA, 2009. ACM.

[63] Da-jung Kim, Yeoreum Lee, Saeyoung Rho, and Youn-kyung Lim. Design Opportunities in Three Stages of Relationship Development Between Users and Self-Tracking Devices. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 699–703, New York, NY, USA, 2016. ACM.

[64] Sunyoung Kim and Eric Paulos. InAir: sharing indoor air quality measurements and visualizations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 1861–1870, New York, NY, USA, 2010. ACM.

[65] Sunyoung Kim, Eric Paulos, and Jennifer Mankoff. inAir: a longitudinal study of indoor air quality measurements and visualizations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 2745–2754, New York, NY, USA, 2013. ACM.

[66] Tanyoung Kim, Hwajung Hong, and Brian Magerko. Designing for Persuasion: Toward Ambient Eco-Visualization for Awareness. In Thomas Ploug, Per Hasle, and Harri Oinas-Kukkonen, editors, *Persuasive Technology*, number 6137 in Lecture Notes in Computer Science, pages 106–116. Springer Berlin Heidelberg, 2010.

[67] 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, New York, NY, USA, 2016. ACM.

[68] Jesper Kjeldskov, Mikael B. Skov, Jeni Paay, and Rahuvaran Pathmanathan. Using mobile phones to support sustainability: a field study of residential electricity consumption. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems*, CHI '12, pages 2347–2356, New York, NY, USA, 2012. ACM.

[69] Predrag Klasnja, Sunny Consolvo, and Wanda Pratt. How to evaluate technologies for health behavior change in HCI research. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 3063–3072, New York, NY, USA, 2011. ACM.

[70] Andrew J. Ko, Brad A. Myers, and Htet Htet Aung. Six Learning Barriers in End-User Programming Systems. In *Proceedings of the 2004 IEEE Symposium on Visual Languages - Human Centric Computing*, VLHCC '04, pages 199–206, Washington, DC, USA, 2004. IEEE Computer Society.

[71] Stacey Kuznetsov and Eric Paulos. UpStream: motivating water conservation with low-cost water flow sensing and persuasive displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 1851–1860, New York, NY, USA, 2010. ACM.

[72] Matthew L. Lee and Anind K. Dey. Reflecting on pills and phone use: supporting awareness of functional abilities for older adults. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 2095–2104, New York, NY, USA, 2011. ACM.

[73] Paul D. Leedy and Jeanne Ellis Ormrod. *Practical Research: Planning and Design*. Pearson, Boston, 10 edition edition, January 2012.

[74] Ian Li, Anind Dey, and Jodi Forlizzi. A stage-based model of personal informatics systems. In *Proceedings of the 28th international conference on Human factors in computing systems*, CHI '10, pages 557–566, New York, NY, USA, 2010. ACM.

[75] Ian Li, Anind K. Dey, and Jodi Forlizzi. Understanding my data, myself: supporting self-reflection with ubicomp technologies. In *Proceedings of the 13th international conference on Ubiquitous computing*, UbiComp '11, pages 405–414, New York, NY, USA, 2011. ACM.

[76] Ian Li, Anind K. Dey, and Jodi Forlizzi. Using context to reveal factors that affect physical activity. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 19(1):1–21, 2012.

[77] James J. Lin, Lena Mamykina, Silvia Lindtner, Gregory Delajoux, and Henry B. Strub. FishnSteps: Encouraging Physical Activity with an Interactive Com-

puter Game. In Paul Dourish and Adrian Friday, editors, *UbiComp 2006: Ubiquitous Computing*, number 4206 in Lecture Notes in Computer Science, pages 261–278. Springer Berlin Heidelberg, 2006.

[78] Jock Mackinlay. Automating the design of graphical presentations of relational information. *ACM Transactions on Graphics*, 5(2):110–141, April 1986.

[79] Adam Marcus, Michael S. Bernstein, Osama Badar, David R. Karger, Samuel Madden, and Robert C. Miller. Twitinfo: aggregating and visualizing microblogs for event exploration. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 227–236, New York, NY, USA, 2011. ACM.

[80] Daniel McDuff, Amy Karlson, Ashish Kapoor, Asta Roseway, and Mary Czerwinski. AffectAura: an intelligent system for emotional memory. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 849–858, New York, NY, USA, 2012. ACM.

[81] Sarah Mennicken, David Kim, and Elaine May Huang. Integrating the Smart Home into the Digital Calendar. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 5958–5969, New York, NY, USA, 2016. ACM.

[82] Matthew B. Miles and Michael Huberman. *Qualitative Data Analysis*. Sage Publications, Thousand Oaks, Califorinia, 3 edition edition, April 2013.

[83] Andrew Vande Moere. Towards Designing Persuasive Ambient Visualization. In *Issues in the Design and Evaluation of Ambient Information Systems Workshop*, Toronto, Canada, 2007.

[84] Andrew Vande Moere, Martin Tomitsch, Monika Hoinkis, Elmar Trefz, Silje Johansen, and Allison Jones. Comparative Feedback in the Street: Exposing Residential Energy Consumption on House Faades. In Pedro Campos, Nicholas Graham, Joaquim Jorge, Nuno Nunes, Philippe Palanque, and Marco Winckler, editors, *Human-Computer Interaction INTERACT 2011*, number 6946 in Lecture Notes in Computer Science, pages 470–488. Springer Berlin Heidelberg, 2011.

[85] Daniele Nadalutti and Luca Chittaro. Visual analysis of users performance data in fitness activities. *Computers & Graphics*, 31(3):429–439, June 2007.

[86] Bonnie A. Nardi, editor. *Context and Consciousness: Activity Theory and Human-Computer Interaction*. The MIT Press, November 1995.

[87] Carman Neustaedter, Lyn Bartram, and Aaron Mah. Everyday activities and energy consumption: how families understand the relationship. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 1183–1192, New York, NY, USA, 2013. ACM.

[88] Shannon O'Brien and Florian "Floyd" Mueller. Jogging the distance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '07, pages 523–526, New York, NY, USA, 2007. ACM.

[89] James Pierce, Chloe Fan, Derek Lomas, Gabriela Marcu, and Eric Paulos. Some Consideration on the (in)Effectiveness of Residential Energy Feedback Systems. In *Proceedings of the 8th ACM Conference on Designing Interactive Systems*, DIS '10, pages 244–247, New York, NY, USA, 2010. ACM.

[90] James Pierce and Eric Paulos. Beyond energy monitors: interaction, energy, and emerging energy systems. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems*, CHI '12, pages 665–674, New York, NY, USA, 2012. ACM.

[91] Zachary Pousman, Mario Romero, Adam Smith, and Michael Mateas. Living with tableau machine: a longitudinal investigation of a curious domestic intelligence. In *In Proc. UbiComp '08, ACM*, pages 370–379, 2008.

[92] 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, New York, NY, USA, 2006. ACM.

[93] J. O. Prochaska and W. F. Velicer. The transtheoretical model of health behavior change. *American journal of health promotion: AJHP*, 12(1):38–48, October 1997.

[94] Arkalgud Ramaprasad. On the definition of feedback. *Behavioral Science*, 28(1):4–13, 1983.

[95] Yvonne Rogers, William R. Hazlewood, Paul Marshall, Nick Dalton, and Susanna Hertrich. Ambient influence: can twinkly lights lure and abstract representations trigger behavioral change? In *Proceedings of the 12th ACM international conference on Ubiquitous computing*, Ubicomp '10, pages 261–270, New York, NY, USA, 2010. ACM.

[96] Purvi Saraiya, Chris North, and Karen Duca. An evaluation of microarray visualization tools for biological insight. pages 1–8, 2004.

[97] Tobias Schwartz, Sebastian Denef, Gunnar Stevens, Leonardo Ramirez, and Volker Wulf. Cultivating energy literacy: results from a longitudinal living lab study of a home energy management system. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 1193–1202, New York, NY, USA, 2013. ACM.

[98] Julia Schwarz, Jennifer Mankoff, and H. Scott Matthews. Reflections of everyday activities in spending data. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pages 1737–1740, New York, NY, USA, 2009. ACM.

[99] Zeqian Shen and Kwan-Liu Ma. MobiVis: A Visualization System for Exploring Mobile Data. In *Visualization Symposium, 2008. PacificVIS '08. IEEE Pacific*, pages 175 –182, March 2008.

[100] M. Six Silberman and Bill Tomlinson. Toward an ecological sensibility: tools for evaluating sustainable HCI. In *Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems*, CHI EA '10, pages 3469–3474, New York, NY, USA, 2010. ACM.

[101] David Sprague and Melanie Tory. Exploring How and Why People Use Visualizations in Casual Contexts: Modeling User Goals and Regulated Motivations. *Information Visualization*, 11(2):106–123, April 2012.

[102] Yolande Strengers. Peak electricity demand and social practice theories: Reframing the role of change agents in the energy sector. *Energy Policy*, 44:226–234, May 2012.

[103] Yolande A.A. Strengers. Designing eco-feedback systems for everyday life. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 2135–2144, New York, NY, USA, 2011. ACM.

[104] Mengting Sun. *Exploring Aesthetic Visualization for Promoting Consumer Energy Conservation*. PhD thesis, Simon Fraser University, BC, Canada, March 2014.

[105] Sanaz Tavakkol. *Personal Analytical Calendar*. Master thesis, University of Victoria, 2014.

[106] Anja Thieme, Rob Comber, Julia Miebach, Jack Weeden, Nicole Kraemer, Shaun Lawson, and Patrick Olivier. "We've bin watching you": designing for reflection and social persuasion to promote sustainable lifestyles. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 2337–2346, New York, NY, USA, 2012. ACM.

[107] Elizabeth Thiry, Sin Lindley, Richard Banks, and Tim Regan. Authoring personal histories: exploring the timeline as a framework for meaning making. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 1619–1628, New York, NY, USA, 2013. ACM.

[108] A. Thudt, D. Baur, S. Huron, and S. Carpendale. Visual mementos: Reflecting memories with personal data. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):369–378, Jan 2016.

[109] J.J. Van Wijk and E.R. Van Selow. Cluster and calendar based visualization of time series data. In *1999 IEEE Symposium on Information Visualization, 1999. (Info Vis '99) Proceedings*, pages 4 –9, 140, 1999.

[110] J. Wood, D. Badawood, J. Dykes, and A. Slingsby. BallotMaps: Detecting Name Bias in Alphabetically Ordered Ballot Papers. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2384 –2391, December 2011.

[111] Fahri Yetim. Critical Perspective on Persuasive Technology Reconsidered. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 3327–3330, New York, NY, USA, 2013. ACM.