Visualizing Personal Data in Context: An On-Calendar Design Strategy for Behaviour Feedback

by

Dandan Huang
B.Sc., University of Electronic Science and Technology of China, 2006
M.Sc., University of Victoria, 2009

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

DOCTOR OF PHILOSOPHY

in the Department of Computer Science

© Dandan Huang, 2016
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.
Visualizing Personal Data in Context: An On-Calendar Design Strategy for Behaviour Feedback

by

Dandan Huang
B.Sc., University of Electronic Science and Technology of China, 2006
M.Sc., University of Victoria, 2009

Supervisory Committee

Dr. Melanie Tory, Supervisor
(Department of Computer Science)

Dr. Lyn Bartram, Co-Supervisor
(School of Interactive Arts and Technology, Simon Fraser University)

Dr. Robert Gifford, Outside Member
(Department of Psychology)
ABSTRACT

Visualization tools are frequently used to help people understand everyday data in their lives. One such example is visualization in behaviour feedback tools. Behaviour feedback tools are used to try to help people improve their health or personal well-being or to carry out sound environmental sustainability practices. However, understanding and reasoning about personal data (e.g., pedometer counts, blood pressure readings or home electricity consumption) or gaining a deeper understanding of one’s current practices and learning how to make a change can be challenging when using data alone. My literature review of this field showed that two of the main challenges in actual practice are providing a context in which to reason about the data and reducing the cost of maintenance to fit those tools into everyday life routines. Thus, I propose to integrate time-varying feedback data within a personal digital calendar. This combination of calendar and feedback data can provide contextual information to interpret data and make the data accessible in an attentionally ambient way that is suitable for maintaining awareness. I propose that the familiarity and common practice of using digital calendars can minimize the cost of learning and maintenance for people and easily fit into one’s daily life routines.

The viability of this approach was confirmed in my quantitative lab experiments. The results showed that visualization of feedback data integrated on a digital calen-
dar is comprehensible, and it does not interfere with regular calendar use with proper visual encodings. After confirming the viability of my proposal, I implemented the on-calendar visualization as a web application that was synchronized with Google Calendar API and a real-time feedback data stream. To further investigate this approach in a real life situation, I deployed the application in the field for longitudinal field studies: two case studies as pilot deployment and an eight-week field study. Results showed that people liked the idea of integrating feedback data into their personal digital calendars. It required a low cost in learning and maintenance. The calendar events provided rich context for people to visualize and reason about their feedback data. The design enabled people to quickly identify and explain repeated patterns and anomalies. Meanwhile, I found that people’s existing information use habits (in this case, how they use digital calendars) can highly influence the effectiveness of the feedback design. Moreover, I derived a feedback model that identifies basic components in feedback design and illustrates the role of feedback tools. With that I articulated possible design barriers that could prevent ongoing use of feedback tools. Reflecting on the effects of the on-calendar design approach, I discussed design implications inspired by this work.

This work introduces a reflective approach in feedback design that can easily fit into people’s existing information ecosystem (specifically, a personal digital calendar in this work). The main contributions of this thesis are: the first systematic literature review of personal visualization design used in everyday life; the design and implementation of an on-calendar design that integrates feedback data on people’s personal digital calendars to provide context for reasoning and support easy access for ongoing use; the extended definition of ambience from spatial location to attentional demand; a viability study to confirm the on-calendar design approach; longitudinal studies to investigate the effects of the on-calendar design approach and the feedback model of design mechanism to inspect ongoing factors in feedback designs.
Contents

Supervisory Committee ii
Abstract iii
Table of Contents v
List of Tables x
List of Figures xi
Acknowledgements xiii
Dedication xv

1 Introduction 1

2 Personal Visualization and Personal Visual Analytics 8
  2.1 Background ........................................ 9
  2.2 Review Method and Process .......................... 10
  2.3 Design Dimensions and Research Interest to Date .......... 12
    2.3.1 Design Dimensions ................................ 12
    2.3.2 Research Interest to Date ......................... 16
  2.4 Design Challenges in PV&PVA ........................ 19
    2.4.1 Fit in Personal Routines and Environments ............ 19
    2.4.2 Recall of Relevant Context for Reasoning ............... 20
    2.4.3 Defining Appropriate Baselines ..................... 21
    2.4.4 Sharing and Privacy ................................ 22
    2.4.5 Evaluation ........................................ 23
  2.5 Limitations ......................................... 24
  2.6 Latest Work in PV&PVA .............................. 25
2.6.1 Increasing User Control ........................................ 25
2.6.2 Include Users in Design ..................................... 26
2.6.3 Variety of Interactions ........................................ 26
2.6.4 Develop Insights with PV&PVA .............................. 27
2.6.5 Fit in Routines and Ecosystems ............................... 27
2.6.6 Evaluation ....................................................... 28
2.7 Contribution ....................................................... 29

3 Related Work ...................................................... 30
3.1 Feedback Design .................................................. 31
  3.1.1 Feedback Applications .................................... 32
3.2 Persuasive Design ................................................. 34
3.3 Ambient Visualization .......................................... 35
3.4 Context Use in Feedback Design ............................... 36
3.5 Evaluation of Personal Visualization in Everyday Context .. 37

4 Visualization Design .............................................. 39

5 Research Methods .................................................. 44
  5.1 Viability Study .................................................... 45
  5.2 Design Study ..................................................... 45
  5.3 Summary .......................................................... 47

6 Viability Study ...................................................... 48
  6.1 Background ....................................................... 48
  6.2 Experiment Design .............................................. 49
    6.2.1 Participants ............................................... 49
    6.2.2 Experiment I: Calendar Tasks ............................ 50
    6.2.3 Experiment II: Visualization Tasks ...................... 50
    6.2.4 Procedure ................................................ 51
    6.2.5 Apparatus ............................................... 51
  6.3 Experiment Results ............................................. 52
    6.3.1 Experiment I: Calendar Tasks ............................ 52
    6.3.2 Experiment II: Graphical Perception .................... 56
    6.3.3 Aesthetics ............................................. 59
  6.4 Discussion of Lab Experiment Results ....................... 60
A  Lab Experiment Tasks in Viability Study  
   A.1  Tasks in Experiment I of Viability Study  
   A.2  Tasks in Experiment II of Viability Study  

B  Post-Experiment Questionnaire in Lab Experiments  
   B.1  Please rate visual distraction of each visualization option (Experiment  
        I only)  
   B.2  Please rate graphical perception of each visualization option (how dif-  
        ficult to perceive the data) (Experiment II only)  
   B.3  Please rate aesthetics of each visualization option (how appealing to  
        perceive the data) (Experiment I and II)  

C  Interview Outlines in Pilot Study  
   C.1  Household Energy Consumption  
   C.2  Personal Fitness  
        C.2.1  Interview One  
        C.2.2  Interview Two  

D  Screenshots of On-Calendar Application (Final Version)  

E  Screenshots of Fitbit Web Application  

F  Background Questionnaire in Field Study  

G  Interview outlines in Field Study  
   G.1  Interview 1  
   G.2  Interview 3  

H  International Physical Activity Questionnaire (IPAQ)  

I  Protocol for IPAQ Long Form  
   I.1  Continuous Score  
   I.2  MET Values and Formula for Computation of MET-minutes  
        I.2.1  Work Domain  
        I.2.2  Active Transportation Domain  
        I.2.3  Domestic and Garden [Yard Work] Domain  
        I.2.4  Leisure-Time Domain
I.2.5 Total Scores for all Walking, Moderate and Vigorous Physical Activities ........................... 155
I.2.6 Total Physical Activity Scores ........................................ 155

Bibliography ................................................................. 156
List of Tables

Table 2.1 Design dimensions, levels with examples from the literature . . . 13
Table 2.2 Summary of surveyed papers . . . . . . . . . . . . . . . . . . . 14
Table 2.3 Summary of evaluation methods showing the number of papers
that included each evaluation method. . . . . . . . . . . . . . . . . . . . 14

Table 6.1 Total accuracy rates (%) with different visual encodings of three
display conditions in Calendar Tasks and Visualization Tasks
(The Visualization Tasks do not include a control condition) . . . 53

Table 9.1 Participants in Fitbit field study . . . . . . . . . . . . . . . . . . 75
Table 9.2 Most frequently used contextual information for reasoning (nor-
malized as frequency per participant). . . . . . . . . . . . . . . . . . . . 89
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>PV&amp;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.</td>
</tr>
<tr>
<td>4.1</td>
<td>Design alternatives displayed overlapped with calendar events (top: line graph; middle: coloured region; bottom: luminance)</td>
</tr>
<tr>
<td>4.2</td>
<td>Design alternatives displayed side by side overlapped with calendar events (top: line graph; middle: coloured region; bottom: luminance)</td>
</tr>
<tr>
<td>6.1</td>
<td>Boxplots showing task time of Experiment I (comparing between two Display Location)</td>
</tr>
<tr>
<td>6.2</td>
<td>Boxplots showing task time of Experiment I</td>
</tr>
<tr>
<td>6.3</td>
<td>Visual distraction reported by participants (-2 is “very distracting”, 2 is “not distracting”)</td>
</tr>
<tr>
<td>6.4</td>
<td>Boxplots showing task time of Experiment II (comparing between two Display Location)</td>
</tr>
<tr>
<td>6.5</td>
<td>Boxplots showing task time of Experiment II</td>
</tr>
<tr>
<td>6.6</td>
<td>Graphical perception reported by participants (-2 is “very difficult” and 2 is “very easy”)</td>
</tr>
<tr>
<td>6.7</td>
<td>Aesthetics reported by participants (-2 represents “very poor” and 2 represents “very good”)</td>
</tr>
<tr>
<td>7.1</td>
<td>Web application of on-calendar visualization using Google Calendar API and displaying household smart meter data</td>
</tr>
<tr>
<td>7.2</td>
<td>Data flow of the on-calendar visualization system</td>
</tr>
<tr>
<td>7.3</td>
<td>Architecture of the on-calendar visualization system</td>
</tr>
<tr>
<td>8.1</td>
<td>Westhouse: eco-friendly smart home for the case study</td>
</tr>
<tr>
<td>8.2</td>
<td>Fitbit trackers used in case study of personal fitness</td>
</tr>
</tbody>
</table>
Figure 8.3 On-calendar feedback application used in the field study (Chapter 9). This is an example of week view with Fitbit data displayed as a line graph overlapped with calendar events. See more screenshots in Appendix D.

Figure 9.1 Study procedure

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 participant’s change in average MET scores.

Figure 9.3 System usage (top: system access versus time of a day; middle: total system usage and bottom: single session duration).

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.

Figure 9.5 Model of behaviour feedback process.

Figure D.1 Coloured region with day view

Figure D.2 Coloured region with week view

Figure D.3 Coloured region with month view

Figure D.4 Luminance with day view

Figure D.5 Luminance with week view

Figure D.6 Luminance with month view

Figure E.1 Dashboard view (aggregated summary data)

Figure E.2 Detail activity log
ACKNOWLEDGEMENTS

My mom always expected me to be a medical doctor. I become a doctor in a different way. Hopefully it would somehow fill in her expectation. I still feel my decision of returning to graduate school was right. It gave me the chance to think and rethink a lot of things, although the path is long and the doctoral work is challenging. With the support of my family and friends, I made it through to the end.

I would like to thank:

• Dr. Melanie Tory: my supervisor, who has given me valuable advice, support and guidance more than I can count, in and out of work. Joining the VisID lab was the turning point of my career. I wouldn’t achieve this today without you. Thank you for all your patience and suggestions to help me handle the difficult situations.

• Dr. Lyn Bartram: my co-supervisor, who provided me the great support and valuable insights from her expertise and experience.

• My husband, Steven Ness: your love and understanding have meant a lot to me, specially in the last year of my doctoral work. Thank you for your patience and sweetness to deal with my moody moments.

• My parents and in-law parents for all your encouragement and support even when you are far away. Particularly, I would thank Fern Ness for helping me edit my thesis. You are so awesome!

• My supervisory committee: you each provided a unique and important perspective for my research.

• My collaborators: Bon Adriel Aseniero, Scott Bateman, Sheelagh Carpendale, Anthony Tang and Robert Woodbury. It was a great pleasure working with you. We made the impossible possible. Writing a paper through collaboration within two weeks is still my record!

• My labmates from both VisID lab and HCSSL lab. Thank you each for the constant feedback on my research and your participation in my pilot studies.

• All the participants for your kindness and patience in my user studies.
• The GRAND research network and NSERC for funding me during my doctoral studies and making this research possible. The GRAND research network was a good memory.
DEDICATION

To my sweet husband, Steven Ness.
Chapter 1

Introduction

Behaviour feedback tools are used to enable people to reflect on their behaviour choices, to ascertain influences that affect their behaviour and to indicate appropriate behaviours, for example, adopting responsible energy consuming behaviours or healthy life choices, etc. This field has been attracting a great deal of attention in recent years as sensing technology has become cheaper and more accessible. Meanwhile, visualization of this increased collection of feedback data is commonly applied in such feedback tools. However, on deeper inspection, these applications are not as perfect as expected. Studies showed that people still have difficulties in understanding their data [119, 21]. It is hoped that people will develop awareness and inferential knowledge by interpreting the data with respect to contextual information 1 in their daily lives. This can help them understand why certain patterns emerge, for example, “why did my home consume even more electricity when I was away for vacation?”,” “what is my fitness level on weekdays versus the weekend?”, etc. However, lack of sufficient context for reasoning is a common shortcoming in many behaviour feedback designs (Chapter 2). Meanwhile, people are already overwhelmed with information and applications, and they are unlikely to frequently use specialized tools. Temporal drop-offs are often reported in previous studies [21, 86, 20, 7, 125]. Thus, the challenge is how to best support interpretation of behaviour feedback data and the development of knowledge without making the user adopt an entirely new suite of tools or spend significant time in assembling context for this data from other sources.

With respect to this challenge, I propose to integrate the visualization of personal

---

1In this thesis context or contextual information is referred to activities and conditions that are related to the feedback data.
feedback data \(^2\) into one’s existing information ecosystem combining it with common information tools that represent the desired context (specifically in this thesis, personal digital calendars). The goals of the thesis are to investigate if the mash-up approach (integrating visualization of personal feedback data on a digital calendar) is a viable design approach that is comprehensible and does not interfere with people’s existing information use habits (i.e., regular calendar use), how people interpret personal feedback data with the context provided by their calendars and how people react to and use the on-calendar visualizations as a feedback tool. That is, a contextual frame of feedback data is needed to minimize the effort required to gather and to integrate the various streams of data: simply adding “yet another app” defeats this goal. Towards this end, a personal digital calendar could be used as such a frame for this purpose.

In this work, the term *context* refers to a broader view than “context-aware” technology [4]. *Context* addressed in this thesis is based on the definition from Activity Theory [118]. Externally, it could be artifacts (e.g., devices or tools), the physical environment or cultural and social situations. Internally, it could be one’s goals, skill sets, preferences or personal experience. These factors might highly influence how people make sense of their personal data and develop insights from that. Specifically in the case of personal behavioural feedback data in this thesis, I define *context* to activities and conditions that are related to one’s behaviour data. For example, daily activities that could provide inferential information about the high and low physical movement, or at-home activities that related to water (or electricity) use.

With respect to this approach, integrating feedback data with one’s personal digital calendar has advantages. Using digital calendars for time management is a common practice for most people. Schedules on a personal digital calendar reflect people’s daily activities that might be helpful in explaining patterns of their feedback data. The familiarity of using a digital calendar could minimize the cost of learning and maintenance, fitting easily into daily routines. As the visualization integrated on a digital calendar the routine use of a personal digital calendar could make people encounter their data quite frequently, enhancing their awareness to an even greater extent.

In this thesis, the work focuses on applying the on-calendar visualization in feedback applications, specifically feedback for personal fitness and household energy conservation. To learn energy consuming behaviours, people collect data about their

\(^2\)In this thesis, feedback data are referred to data about household energy use and personal fitness.
home or workplace, e.g., electricity consumption, water use, temperature, etc. Meanwhile, it is a common practice to collect fitness data about oneself, especially with wearable fitness trackers, e.g., exercise tracking or pedometer counts. It is important to note that this type of data is common in people’s lives, and understanding this data is important for individuals. These tools can provide ongoing feedback of one’s behaviours. Better understanding of these patterns can help individuals make better behavioural choices to influence their lives. Visualization of these types of feedback data typically focuses 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 self-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. For example, Upstream, a water usage indicator in the shower, engages people to make efficient water use behaviours immediately while showering [99]. In some cases, it employs behaviour change models to coerce expected actions, for example, applying social pressure [148]. In contrast, my interest is to investigate how visualization design could enhance the reflective understanding of one’s behaviours in a non-persuasive way. The general philosophy centers around helping people know their behaviour with operational understanding [21] on a ongoing basis rather than digitally lecturing them into behaviour 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.

Calendars are both a planning tool and a record of people’s daily activities, and thus, may be useful in helping people understand data that is 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 [151]. This could provide a general context for comparing data patterns (e.g., comparing weekdays in a week) and aggregating feedback data with respect to the time scale (e.g., switching between week view and month view). The use of a digital calendar is a common practice nowadays. The familiarity of calendar views and the routine use of calendars should require low learning and management cost, and it should easily
fit into one’s life routines. More importantly, calendar events provide rich contextual information about one’s daily life, which could be relevant to understanding personal behaviour data. Evidence shows that people are inclined to check personal calendars for reasoning about their personal data in everyday life [27, 101, 112, 119]. Thus, a personal digital calendar can provide appropriate contextual framing to help people recognize and understand information patterns.

However, the display of additional visualization layers on a calendar needs to be approached with care. When “mashing up” information sources in a familiar tool, I 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 calendar), the visualization design has to balance visual interference with normal calendar tasks, and attention should be placed on the perceptibility and legibility of the added data stream. The design needs to minimize visual interference while supporting effective data perception. Meanwhile, people have their own habits of using digital calendars (e.g., choice of calendar views, sparse or dense layout, colour coded event boxes). The design might need to provide the flexibility to cope with the varying conditions.

This thesis investigates the viability of on-calendar visualizations for behaviour feedback. The study is divided in multiple phases: a systematic 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 prototype, and a longitudinal field study to deeply investigate how the on-calendar feedback tool is used in everyday life context.

First, I reviewed the literature of visualization designs used in everyday life to identify design dimensions and challenges in personal visualization. I proposed the design approach and alternatives and aimed to tackle two challenges: providing contextual information to enhance understanding and easy access to encourage ongoing use. After that, I conducted a lab study to evaluate design alternatives with respect to the interference and perceptibility. With the selected visualization alternatives based on the study results, I implemented an on-calendar visualization that 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 visualizations used in everyday life?
- How do design characteristics of visual encoding and display location influence the perceptibility of on-calendar visualization?
- How do design options of visual encoding and display location influence the level of interference with primary calendar tasks?
- What are the design options that can balance the interference and perceptibility for the on-calendar visualization?
- To what extent can people use calendar events as context for reasoning about their personal feedback data?
- Could people improve their understanding of their data and behaviours better with the context provided by personal digital calendars?
- How do people react to and use the on-calendar visualization as the feedback tool in everyday life?

This thesis describes the following significant and novel contributions:

1. The first systematic literature review of data visualization used in a personal context. It provides the starting point for researchers and designers to consider design requirements and design dimensions for applying visualization design in an everyday life context. This work introduces the design field of Personal Visualization and Personal Visual Analytics and brings together research that was previously scattered in different disciplines and research fields.

2. A design approach that integrates the visualization of personal quantitative feedback data into a personal digital calendar. In this work, I employed this approach in two examples: personal fitness and home energy conservation. The design approach can provide daily-life context for people to better understand their feedback data. It incorporates people's existing habit of information use, requiring low learning and managing cost and engaging ongoing use.
3. A real-life use example of applying the concept of *attentional ambience*. It extends ambient visualization from spatial location (displayed in peripheral location) to attentional demand [22]. This perspective changes “environmentally appropriate” to “attentionally appropriate” with respect to attentional demand.

4. Quantitative lab experiments that confirmed the viability of the on-calendar design approach. These experiments showed the promise of the proposed mash-up approach where additional visualization layers could be perceived ambiently with proper visualization choices.

5. Qualitative field deployments that demonstrated the effects of the on-calendar design approach. The studies confirmed that personal digital calendars could provide rich contextual information for people to reason about their data. Particularly, they proved helpful to identify and reason about data patterns and anomalies. In addition, the on-calendar visualization made the relevant information easy to access.

6. A new model of the behaviour feedback process and the role of feedback tools based on the study results. The model offers a structure to investigate the roles of design components in feedback design and also indicates 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 behaviour 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, my investigation of on-calendar visualizations provides a foundation for exploring a potentially large design space of attentionally ambient visualizations.

In the following chapters of this thesis, a systematic review across several fields is conducted first to identify design unique requirements, design dimensions and challenges of personal visualization & personal visual analytics (PV&PVA) used in everyday life in Chapter 2. Among all the challenges, this work focuses on how to provide context for people to reason about personal feedback data and how to make
it easily fit into everyday 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 research method in this thesis to investigate the on-calendar design approach in Chapter 5. First, a viability study was used to evaluate the design alternatives with respect to visual interference and perceptibility 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 version of implementation in a real life context, I deployed the application in two case studies (home energy conservation and personal fitness) as pilot studies (Chapter 8), where the application was connected with real time data (household smart meter and Fitbit devices respectively) and the 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 is then presented in Chapter 10 along with conclusions in Chapter 11.

Supplementary information includes the task lists used in the viability study (Appendix A), post-experiment questionnaires for the viability study (Appendix B), interview outlines of pilot studies (Appendix C), screenshots of the on-calendar web application (Appendix D), screenshots of the Fitbit web application (Appendix E), background questionnaire for screening in the field study (Appendix F), interview guide for the field study (Appendix G) and International Physical Activity Questionnaire (IPAQ) used in the field study (Appendix H).
Chapter 2

Personal Visualization and Personal Visual Analytics

Visualizing personally relevant data has attracted a great amount of attention in recent research and practice. However, examples of this work have been scattered across several research communities, for example, environmental psychology, personal informatics, ambient visualization, information visualization and human computer interaction. There has been no previous systematic review of personal visualization work and no work to bridge this research among these communities. Meanwhile, gaining knowledge of the design patterns and gaps in previous practice should be the first step in a study and should precede 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.

This chapter introduces the design field of Personal Visualization and Personal Visual Analytics, where I describe and discuss the first systematic literature review of data visualization used in a personal context (Contribution 1). It provides the starting point for researchers and designers to consider design requirements and design dimensions for applying visualization design in an everyday life context. This work was published in the journal of IEEE Transactions on Visualization and Computer Graphics [82], to which my collaborators (Melanie Tory, Bon Adriel Aseniero, Lyn Bartram, Scott Bateman, Sheelagh Carpendale, Anthony Tang and Robert Woodbury) also contributed.

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 research needs to face.

\section{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 data on our resource usage (e.g., utilities, such as water and 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 ourselves (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 volume of data 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 data into insights. Visualization is a powerful practice in this transformation. However, research into this practice has been mostly focused on professional situations. As the interest of applying visualization in everyday life to explore personal data or personally relevant data is growing, how can designers better design visualizations to meet this fast growing need? What are the new requirements and challenges researchers and designers have to face? Since the relevant research is currently scattered 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.

\textit{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 difference between the two areas is analogous to the difference between Vis and
VA: Personal Visual Analytics involves both visualization and automatic computer assisted analysis, whereas Personal Visualization focuses on visual data representations. Note that in normal conversation and writing I expect that people will use either PV or PVA but not both terms together. However, for the purposes of the current review and summary of the areas, in this work I will refer to the two areas collectively as PV&PVA.

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 those 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 people’s 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 context. 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 provides 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 (Table 2.1). These dimensions provide a basis upon which to identify clusters of work and also challenges in this area.
As part of the iterative process, I conducted an extensive literature survey of PV&PVA 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, one’s 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 analyst’s 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 of an additional seven papers encountered in other domain specific venues (summary in Table 2.1). The development of the taxonomy of PV&PVA 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. Open coding [141, 114] is an often used method for such qualitative development by breaking data into conceptual components. Data sources are broken down and the concepts of the pieces are labeled. Researchers iteratively examine the concepts and develop categories to merge similar concepts based on their properties. 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 my collaborators and I reached consensus and when the dimensions could adequately represent the unique attributes of PV&PVA tools in the collection. The final set of dimensions are in Table 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 refers only to their latest version of the tool. This reduced the set of 66 papers to 59 designs (Table 2.2).

2.3 Design Dimensions and Research Interest to Date

The literature review showed that research attention in PV&PVA has been steadily increasing over the past five years (Table 2.2) and the greatest number of papers has come from the HCI community. The papers cover a variety of applications that address most aspects of daily life, e.g., residential energy consumption, fitness, personal health, social networks, politics, residential environment, life logging, personal finance and recycling. In 56 out of 66 papers, researchers evaluated their designs using one or more methods (details in Table 2.3). With the collection of 59 designs, I coded and mapped them to an interactive parallel sets plot of design dimensions that helped to identify design trends and research interest to date (Figure 2.1). 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 2.1. The interactive tool is available online \(^1\). In this section, I articulate these dimensions in the design space and discuss emerged design patterns that reflect the research interest to date.

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 actionability (34 out of 45 cases, seen by hovering the mouse over “no effort” or “no control” in the interactive visualization. See in Figure 2.1). 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.

\(^1\)http://ieeexplore.ieee.org/xpls/icip.jsp?arnumber=6908006
Table 2.1: Design dimensions, levels with examples from the literature

<table>
<thead>
<tr>
<th>Categories</th>
<th>Dimensions</th>
<th>Definition</th>
<th>Levels</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data</td>
<td>Data Scope</td>
<td>Who the data is about</td>
<td>self</td>
<td>Sleep quality [25]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>family</td>
<td>Internet bandwidth shared at home [36]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>peers</td>
<td>Relationship with friends [73]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>community</td>
<td>Hang-out patterns on campus [135], online conversations [62]</td>
</tr>
<tr>
<td>Data Effort</td>
<td>Amount of effort that is expended in data collection</td>
<td></td>
<td>none</td>
<td>Online search history [24]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>sensor</td>
<td>Nonintrusive sensing devices, e.g., wearable sensors [67]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>manual</td>
<td>Manual logging pictures and annotations [109]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>mixed</td>
<td>Combination of sensor recording and manual input [106]</td>
</tr>
<tr>
<td>Data Agency</td>
<td>The degree of control a person has over what data is collected, and when and how it is collected</td>
<td></td>
<td>no control</td>
<td>Online conversation logs [62]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>partial control</td>
<td>Users have control of whether or not to collect the data but cannot customize what data they would like to collect [67]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>total control</td>
<td>Manually recorded childrens photos and growth progress [88]</td>
</tr>
<tr>
<td>Design Context</td>
<td>Who designed and developed the application</td>
<td></td>
<td>self</td>
<td>Visualization designed by oneself</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>group</td>
<td>Tools designed by a study group to chart their progress</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>participatory</td>
<td>Using an online survey to get feedback on early visual design concepts [67]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>third party</td>
<td>Visualization of music listening history designed by the researcher [27]</td>
</tr>
<tr>
<td>Settings</td>
<td>In what situation the tool is used and how it is used</td>
<td></td>
<td>personal</td>
<td>Personal laptop (mostly non-mobile but used by oneself) [112]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>domestic</td>
<td>Ambient display at home (mostly non-mobile) [69]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>mobile</td>
<td>Used on a mobile phone while on the go [67]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>shared</td>
<td>Visualization of physical activities viewed by co-workers [107]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>public</td>
<td>Visualization to promote energy conservation presented in a public space [77]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>mixed</td>
<td>Combination of above, e.g., visualization of residential energy on a personal computer and a mobile phone [23]</td>
</tr>
<tr>
<td>Influence Context</td>
<td>Who the application is intended to inform</td>
<td></td>
<td>self</td>
<td>My physical condition [46]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>family</td>
<td>Children’s growth progress [88]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>community</td>
<td>Inform public about elections [155]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>mixed</td>
<td>Encourage water drinking for peers and oneself [38]</td>
</tr>
<tr>
<td>Attentional Demand</td>
<td>How much attention is required to interact with the tool</td>
<td></td>
<td>low</td>
<td>Cell phone wallpaper [25], ambient display [69]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>mixed</td>
<td>Pedometer counter on a cell phone and the historical data on a desktop [106]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>high</td>
<td>Exploration of music listening history with focused attention [27]</td>
</tr>
<tr>
<td>Interaction</td>
<td>The ability to explore the data</td>
<td></td>
<td>low</td>
<td>Visual metaphor representing one’s physical activities [67]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>mixed</td>
<td>Data overview on a cell phone and exploration of history data on a client service [23]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>high</td>
<td>Exploring historical data with dynamic queries [27]</td>
</tr>
<tr>
<td>Insight</td>
<td>Degree to which the insight gained from using the tool can guide future actions</td>
<td></td>
<td>low</td>
<td>Does not inform further action, e.g., supports reminiscing about the past [127]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>mixed</td>
<td>Smartphone app that gives reminders about when to exercise and enables reflection on progress via historical records [106]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>high</td>
<td>Engages a certain behaviour, e.g., encourage energy conservation [94]</td>
</tr>
<tr>
<td>Automated Analysis</td>
<td>Data mining or other automated analysis methods are employed</td>
<td></td>
<td>Yes/No</td>
<td>Classify physical activities to types of transportation [67]</td>
</tr>
</tbody>
</table>
Table 2.2: Summary of surveyed papers

<table>
<thead>
<tr>
<th>Before 2008</th>
<th>VIS Community</th>
<th>HCI Community</th>
<th>Other</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008</td>
<td>2</td>
<td>6</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>2009</td>
<td>1</td>
<td>7</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>2010</td>
<td>3</td>
<td>6</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>2011</td>
<td>2</td>
<td>8</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>2012</td>
<td>2</td>
<td>16</td>
<td>1</td>
<td>19</td>
</tr>
<tr>
<td>2013</td>
<td>N/A</td>
<td>6</td>
<td>N/A</td>
<td>6</td>
</tr>
<tr>
<td>Total</td>
<td>10</td>
<td>49</td>
<td>7</td>
<td>66</td>
</tr>
</tbody>
</table>

Table 2.3: Summary of evaluation methods showing the number of papers that included each evaluation method.

<table>
<thead>
<tr>
<th>Evaluation</th>
<th>VIS Community</th>
<th>HCI Community</th>
<th>Other</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>3</td>
<td>6</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>Lab Study</td>
<td>4</td>
<td>11</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>Interview</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Survey</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Field Study</td>
<td>0</td>
<td>33</td>
<td>5</td>
<td>38</td>
</tr>
</tbody>
</table>
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.

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 [118] 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” [85], 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 own data with different goals, backgrounds, and expectations (i.e., internal context), which can highly influence 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 one’s family or oneself. I 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 the 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 the use of mobile devices and shared views aimed to achieve 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 the use of automated computational assistance (e.g., classification algorithms) is common in visual analytics generally, this type of analysis was not common in the tools that I 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 sets 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.

Enabling Exploration for Curiosity

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

The first trend is designs for enabling exploration 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 and were mostly used to understand something 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” [13], “my computer usage” [19], “places I have been to” [84] or “my finance” [137]). These tools enabled user exploration facilitated by typical analytical tasks such as select, reconfigure, encode, elaborate, filter, connect, etc. For example, by interactively exploring (as in traditional InfoVis techniques) with music listening history [27] 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 of these tools.

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 [27]. Spending data can be explained by relevant routine activities, e.g., coffee drinking habits [137]. 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.

Supporting Awareness for Action

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

Another common design trend is to provide in-the-moment or ongoing awareness with respect to personal behaviours. 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 [25]. 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 [46] or disaggregating water consumption based on water use behaviours [69]. Some used graphical metaphors to remind people of the potential impact of their behaviour. For example, to encourage people to exercise and to take green transportation [67], a polar bear on a piece of ice was displayed; the ice began to melt if the user’s behaviour 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 behaviour change (e.g., for drinking more water [38], staying physically active [107] or encouraging recycling [148]). However, social engagement and comparison may also raise other problems: inappropriate social strategies actually made the design less effective or caused undue stress [148], and the viewing of other people’s personal data also raises privacy concerns.

**Taking Care of Family**

*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 to better fit in the home environment [69]; others ran on a personal computer, enabling close exploration and organization of family data to track progress [88]. These applications were designed to monitor or engage behaviours towards 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 different cognitive levels (children versus adults) in the family [69]. Additional contextual knowledge was also provided in visualizations to help people interpret the data, for example, by narratively depicting quantitative measures [69], 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 [36] was designed to educate family members about their shared Internet usage. Family members could view each other’s online activities and bandwidth usage could be prioritized with respect to social roles. Here, some family members noted an unwelcome intrusion on privacy. The challenge is how to balance the diversity of users in a family with respect to cognitive capabilities, skills and social roles.

**Reflecting on Communities**

*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 in. These designs were usually
intended to inform the public or a certain social group, e.g., raising public awareness of elections [155], supporting easy exploration of survey data [54] or revealing topics evolving from social networks [53, 62]. In a few examples, they were also used to encourage behaviours valued within the community, for example, ambient displays deployed in a department lobby to encourage energy conservation and physical activities [77].

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 and usually required 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 [110] and sentiment analysis [53].

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 [138] and employing non-traditional representations to compensate for the limited analytics skills of non-experts [54].

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&PVA 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, designs may wish to support fleeting use of a fitness tracking application without interrupting one’s life routines or to customize a visualization’s appearance so that it matches the aesthetic of a living room where it will be deployed.

Fitting into people’s lives means that designers should consider availability, ease of access and ease of use for long-term adoption. Kim [91] identified two stages of how people adopt everyday technologies: in the early stage, interest is the main motivation; then gradually the tool is adopted into daily routines. In a later stage, people’s 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. People’s 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 person’s 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. The on-calendar approach (Chapter 7) employs this concept that the feedback data visualization is integrated into one’s existing information ecosystem (that is, the routine use of personal digital calendar). The familiarity and common practice of using a digital calendar minimizes the cost of learning and maintenance. People can frequently encounter their behavioural feedback data without changing their information use routines (Chapter 9).

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 people’s 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 [8] has
recognized that people’s understanding and use of information artifacts are strongly influenced by the context (experience, preferences, competencies, values, etc.) Relevant context for interpreting data in a PV&PVA tool might be the knowledge of one’s 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 behaviours. Lee and Dey conducted a study with older people on pill taking [101]. 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. However, 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 people’s 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 could explain the anomalies. Similarly, personal calendars can infer rich context about one’s daily activities and places that might be related to explain patterns or anomalies of the behavioural feedback data (Section 9.5.5). Meanwhile, a typical digital calendar frame provides the flexibility of assembling such context and quantitative time-varying feedback data (Chapter 7).

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 children’s development to milestones provided by a pediatrician [88], family members could compare their water usage with each other or among different rooms [69] or 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 person’s 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 person’s values. For instance, many people may be interested in leading healthy lives. Yet what constitutes “healthy” may differ - for 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 [127, 149]. One intriguing space is to apply social interactions to enhance motivation or persuade behaviour change, for example, setting group goals [107], comparing your own progress to others [38] or even interfering with social surveillance [148]. 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 other’s 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 personalized data encodings may be an essential design feature. One example is UbiFit, which provided a view of one’s 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 the 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 people’s 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 reconsider 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 one’s daily life, habits and routine? Can one ease into the use of the tool without making effort to breaking from one’s 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 people’s environments, routines, habits and social experiences and how this comfort level evolves and adapts 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 [71, 96]. Dedicated applications have to face the low usage issue [21, 86, 59, 150]. How to encourage ongoing use is a critical factor in PV&PVA
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 [149]. In contrast, the 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.

In this thesis, I evaluated the on-calendar approach in multiple phases with a combination of traditional lab-based visualization evaluation and qualitative field studies. The cognitive metrics (e.g., task completion time and task error that are commonly used in visualization evaluation) in lab experiments help to confirm the viability of my design approach. However, to investigate how people would react to and use the on-calendar application in a real life condition requires a longitudinal field deployment beyond the lab, e.g., if it is easy to learn and fits into people’s existing routines, how their existing information use habits impact on the effectiveness and ongoing use of the design, etc. (Chapter 6, Chapter 8 and Chapter 9).

2.5 Limitations

I am aware that the taxonomy is based on the literature data that have limitations because of the venues and the years covered. The literature search was focused on visualization and HCI (human computer interaction) venues before 2013. It cannot be neglected that research and practice in this area has been substantially growing since then, even beyond these venues. As a starting point to foster PV&PVA research, the taxonomy will evolve and expand as PV&PVA becomes established as a field.

The coding process of the literature data was based on my collaborators’ and my expertise and insights in previous research, mostly in information visualization and human computer interaction. Possibly, the lack of diversity in experience and expertise may bring bias in the qualitative coding and constrain the generality of the taxonomy.
For the literature search criteria, this work only includes academic work, but the industry practice related to PV&PVA has been rapidly growing, e.g., the Quantified Self movement. As well, these industry tools share common ground with academic work. However, for tractability, the comprehensive review focuses on the academic literature.

2.6 Latest Work in PV&PVA

Since this work was published, I have observed the tremendously growing trend in this field. For example, designers and researchers have applied new techniques to support data collection, and investigated tools and new interactions not only to support personal analytics but also social awareness. Research efforts have also been devoted to critical issues in PV&PVA, such as privacy, personalization, emotional influence, evaluation, etc. In this section, I reflect on the latest research and practices in this field and investigate this growing trend with respect to my early work. The work discussed in this section is mostly selected from the research papers in CHI, IEEE VIS and UbiComp in the past two years.

2.6.1 Increasing User Control

My previous literature review showed that manual collection could enable people to control data collection. However, the big drawback is that it usually requires a lot of user effort to cope with the collecting process. Recent research has been exploring this area to minimize user burden in manual collection and meanwhile provide people better control of what, when and how personal data are collected [93]. In many cases (e.g., sleep quality) manual collection might be a good choice to gather the data, e.g., feelings or sleep-related food and activities. Choe et al. [40] made use of the lock screen and home screen widgets to reduce the data collection effort and improve access to information.

Recent research has shown growing interest in privacy. Specially in PV&PVA, people have an increasing need of controlling information they share with others, for example, what, how and with whom to share. One of the most investigated areas is location sharing [18, 124, 50, 9]. Almuhimedi et al. showed the advantage of using privacy nudge, the contextual information about location data that were accessed by mobile applications, and engage the user to reconfigure the permission [9].
Many studies suggested that personalization could help to protect privacy. Kobsa et al. 2014 developed a model of privacy attitudes and suggested to improve client side (e.g., on one’s smart phone) personalization to increase perceptual privacy [97]. In the study of VeilMe, Wang et al. [152] investigated how people configure the privacy preferences of sharing and pointed out the benefit of personalizing the initial privacy settings. In a different setting, Davies et al. addressed the privacy issue when people share content from their personal device (e.g., mobile) to a public display [50]. In this work, they proposed a design model to protect individuals’ identity information while sharing content with a public display.

2.6.2 Include Users in Design

I am glad to see the trend of including end users in the design process [121, 10, 153, 144, 18]. The design of BodyVis had a 15-month iterative design process with the participatory method of Cooperative Inquiry [121]. Amini et al. conducted series of workshops to observe how professional storytellers narratively used video data to understand the creation process of novices with video content [10]. To design FeedFinder [18], a location sharing application for breast feeding, researchers and designers included participants in each of the design phases, gathering requirements with interviews, exploring design options with workshops, evaluating medium-fidelity prototype with cooperative evaluation, and investigating the implementation with field deployment.

2.6.3 Variety of Interactions

More and new personal devices have been used to support PV&PVA. E-textile displays were used to support group awareness in running events [111] and show human anatomy interactively for elementary students [121]. VR devices were used to improve people’s meditation experience [72]. Ambient displays were put in the context to provide in-the-moment awareness, for example, by a laundry machine [29]. This could also go beyond the digital display. Lee et al. gave an example of visualizing physical activities by piercing a wristband [103].

Meanwhile, researchers investigated the cross-device experience because it becomes common that multiple devices are applied in a personal digital ecosystem. Hamilton et al. proposed a framework for constructing cross-device applications that enable connections and interactions to explore data on multiple displays [75].
Chen et al. investigated interactions between a smart phone and a smart watch with motion and touch input [35]. Davies et al. improved the privacy of public display for showing content from personal mobile phones [50]. As technology continues, more and more devices will be deployed in this field. For that, better supporting experience of using PV&PVA across multiple devices in a holistic ecosystem would be a demanding research area.

2.6.4 Develop Insights with PV&PVA

As one of the key dimensions, I observe the growing attention on supporting people to develop insights with their data. Choe et al. categorized the types of insights when people use personal visualization [39]. These categories were based on professional use of visualization systems [134]; however, it could be a good start to investigate the types of insight in personal scope.

Insights of using PV&PVA is highly related to personal context. Many studies showed the increasing need of contextual data for reasoning purposes [18]. Normark et al. integrated contextual information about various groceries (e.g., local news, weather, tweets and organic blogs relating to the product), aimed to help people better understand eco-friendly grocery shopping [120]. The study of Wood et al. showed the importance of contextual data to understand biking routes through visualization [154]. Due to the constraints of the system intelligence, Kendall et al. even suggested to personalize such context to improve the interpretation of one’s blood pressure change according to one’s own condition [87]. This could help people understand the flexibility of behaviour choices. Especially in the area of sustainable HCI, many researchers considered behaviour change was negotiable rather than standardized [120, 29].

Meanwhile, affect factors could also influence insights towards behaviours. The framing of information on the visualization could impact the interpretation and the behavioural outcome. For example, negative framing would engage the behaviour change more effectively than positive framing [93]. Another example is “Walking by Drawing” [133], in which people could directly see how their actions (i.e., walking) generate visualization patterns and be creatively engaged in personal fitness.

However, insights of using PV&PVA are not always related to actions or analytics. PV&PVA have been also used to encourage family connection, social conversation, or personal reminisce. An interactive ambient display enabled family members to
share their tea time by sending personalized messages [31]. Thudt et al. designed a visualization tool to support serendipitous exploration and help people reminisce their location history [150].

### 2.6.5 Fit in Routines and Ecosystems

Designers also tried to take more consideration of the unique PV&PVA context in everyday life, making the design better fit in people’s daily routines and information ecosystems. In the design for searching for a breast feeding location, Balaam et al. suggested that it was necessary to consider routines how women did breast feeding [18]. For example, these women only had one hand free to interact with the mobile phone for the searching task. Lee et al. applied an ambient display to help seniors reflect on their history of taking medication [102]. The study showed that individuals had integrated the feedback use into their routines to support their self-awareness. Also, the feedback application could help seniors develop their own medication schedule that can better fit in their daily life routines. Sørensen et al. suggested interaction designs in “personal ecologies” might need to fit in the real-life digital ecosystem [146], in which collaboration among multiple users and multiple devices needs to be fully considered.

### 2.6.6 Evaluation

Field deployment is commonly used to evaluate PV&PVA tools in recent research, especially in HCI communities. Researchers explored and investigated the evaluation criteria specifically for personal use cases different from traditional evaluation of visualization systems, for example, engagement, adoption or abandonment, withdraw effect, etc. In the study of an exploratory visualization with CO$_2$ pollution data [30], Boy et al. investigated user engagement based on the interaction deepth, measured by the number of interactions during the exploration. Gouveia et al. investigated user engagement with a fitness tracking application, by reflecting the metrics of engagement, the number of access and session time, with respect to user goals [70].

In some cases, PV&PVA is designed to encourage certain behaviours. Lee et al. investigated the withdraw effect of the feedback tool for medication history reflection [102]. Although they found the feedback helped improve the consistency of medication-taking, this improvement did not persist after the feedback was removed.
Recent research has started to focus on evaluating the adoption of these tools for personal use [70, 42, 37]. Clawson et al. investigated the long-term adoption and the abandonment of using fitness tracking applications [42], indicating the gaps between user goals and application capabilities. In an inspiring example, Chilana et al. presented a case study to investigate the research-to-product transition [37] and called a change of design perspective from user centered to adoption centered with a focus on adopters and stakeholders of the product.

Overall, how to evaluate PV&PVA application is still an open question. New research methods and evaluation criteria will be developed as this field is dynamically growing.

2.7 Contribution

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. This literature review has identified new challenges in visualization design used in everyday life and a taxonomy of design dimensions to provide a coherent vocabulary for discussing Personal Visualization and Personal Visual Analytics. It should help designers and researchers better understand the unique characteristics and requirements in this field and bridge work from different communities. Particularly, the on-calendar design approach proposed in this thesis mostly aims to tackle two of the challenges: fitting in personal routines and providing relevant context for reasoning. Evaluating designs in this field is generally difficult because of its unique characteristics. The combination of quantitative and qualitative evaluation could be an example of exploring appropriate evaluation methods on this path with a focus on participants’ qualitative understanding and the design approach’s fit-in than quantitative task-based performance.
Chapter 3

Related Work

The previous literature research of PV&PVA showed the current state of data visualization and design requirements and design dimensions for visualization design in everyday use. Specifically, behavioural feedback design is one of the major application domains of PV&PVA that faces the same set of challenges. Such designs are aimed to help people understand their behaviours (e.g., behaviours towards energy conservation or healthy life choices) and influence their decisions with respect to behavioural choices. Lessons learned in PV&PVA provide the directions and guidelines to investigate this particular field. Moreover, the on-calendar design approach is proposed to tackle two of the challenges: providing contextual information in which to reason about personal feedback data and supporting flexibility to fit in everyday routines.

In the rest of the thesis, I focus on the behavioural feedback design and the investigation of the on-calendar visualization used as a behavioural feedback tool. This chapter provides the background in behavioural feedback design. First, a selection of design examples were analyzed, with the focus on applications in energy conservation and personal fitness. Among the examples, I then reflect on the common design strategies in this area: persuasive design and ambient visualization. Particularly, this work investigated and applied the concept of “attentional ambience” in practice to tackle a real-life problem. Attentional ambience describes a design in which an ambient visualization is extended from spatial location to attentional demand [22]. In the last section, I discuss the evaluation methods that have been typically used in visualization design and HCI research that inspired my study design in this work.
3.1 Feedback Design

The definition of feedback dates back from the study of behaviour science in organization learning and management theory [130, 16, 81]. Ramaprasad’s definition, “information about the gap between the actual level and the reference level” [130], points to the three key components in feedback: current status, reference and gap. Herold and Greller [81] also claim that feedback information would reflect behaviour influence, indicating appropriate behaviours with respect to the reference (e.g., a goal) and how well these behaviours have been executed. Later, the concept of feedback was introduced in mechanical systems where it is used to control and adjust the system behaviour by monitoring the output and feeding it back to the system [64]. Current studies of feedback systems in human computer interaction probably extended from environmental psychology where people are assumed to lack awareness, and the understanding of their behaviours in everyday life could lead them towards sustainable living [68].

In systems designed for influencing behaviours, behaviour related information was typically provided as antecedent or consequence interventions to affect behavioural decisions [5]. For example, goal setting, commitment or public media campaigns are typical antecedent interventions, and feedback or rewards are popular consequence interventions. This thesis focuses on the feedback data of behavioural consequences. That is, data feeds are subsequent to behaviours. Such behavioural feedback systems are designed to inform the behavioural outcome, show people what are the behaviour choices and engage them to adopt certain behaviours. Specifically, my interest here is how to design and apply these behavioural feedback designs in two types of application areas: energy conservation and personal fitness. For example, by showing people their fitness data as feedback from fitness trackers, they could possibly make better behaviour decisions towards healthy life choices.

However, in contrast to environmental psychology, researchers in HCI are more interested in designing and engineering interactive feedback systems and exploring design approaches and how to apply designs to help people [52, 126, 21, 68, 139, 143]. Disolvo et al. reviewed work in sustainable HCI and analyzed the genres and scope of research topics in this emerging field [52], offering a multi-faceted perspective to rethink emerging issues in sustainable HCI. Moreover, Pierce conducted a review focusing on Electricity Consumption Feedback (ECF) [126], in which they outlined the research in an even broader scope beyond interactions. They also suggested
energy-related HCI research should inspect design strategies that increase awareness and engage individuals in practice rather than narrowly focusing on behaviour 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 [66]. For example, the frequency of updating feedback matters: the more frequently feedback is given the more effectively it influences behaviours [5]. Consolvo et al. also summarized different types of feedback and compared the advantages and disadvantages of each type [45].

3.1.1 Feedback Applications

From a data perspective with the taxonomy discussed in previous Chapter (Section 2.3), behavioural feedback tools mostly provide information about individuals (e.g., one’s transportation behaviours [43], sitting postures [74], physical activities [106, 7], water drinking habits [38], family and home data (e.g., electricity use [131, 135], water use [59, 69], air quality [90]), aimed to engage responsible or healthy behaviour change. This approach was taken to benefit communities as well in some cases. For example, co-workers [107] or the general public [15, 77].

As a result, the use situation varies. Kjeldskov [94] developed a mobile phone application, Power Advisor, which enables users to watch the electricity use on one’s smart phone. Froehlich designed and built an ambient feedback display [69] installed in participants’ home to educate them about water usage shared by family members. Schwartz et al.[136] designed the feedback to display on a household’s TV, so users could review their electricity consumption during commercial periods. Similarly, such energy use feedback displays could be used in public buildings for the building occupants and visitors [15, 77]. Meanwhile, sensor-based wearable devices and mobile phones are commonly used in this area, enabling users to collect or show data in the moment and on the go [43, 67, 25].

Behavioural feedback tools could also support different types of interaction and attentional demand (Section 2.3). In-the-moment feedback requires low attentional demand and engages people to take an action in the situation. Upstream [99] indicated water usage during one’s shower. Similarly, waterBot [14] showed users their water usage over the sink. Rogers [131] applied artistic visualization to show electricity intensity when the appliance is in use. A real-time updated visualization can
alert an inappropriate sitting posture when it happens [74]. As smart phones are
getting increasingly more popular, visualization of feedback has been integrated into
the phone use [43, 67, 25], for example, as the phone wallpaper. This increases the
chance people encounter their data with minimal attention and effort. On the other
hand, people also expect to explore historical data to learn patterns, causes and influence of their behaviour decisions. Replacing the paper bill, Dubuque [59] was a web portal where users could review data patterns of water consumption. Li proposed a design that incorporated people’s geolocation and social interaction while displaying pedometer data [106], hoping to improve people’s interpretation of their fitness level. More interestingly, such reflection could also be supported by artistic style of visualization [61].

One of the common design components to develop insights and knowledge through interaction is to support comparisons with peers, families, neighbors or oneself. Comparisons with a similar household in the neighborhood may help people have better understanding of their energy consumption level [59, 116]. Many designs employed social comparison to engage behaviour change, in which one can share and compare progress with friends, co-workers or peers [122, 26, 107, 38]. Lin introduced group comparison to engage people in physical activities [107]. VERA [26] 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 choices.

Computer assisted analysis has been often used to enhance insight development, especially in large amounts of data where patterns are not easily recognizable visually. For example, Ubifit, by classifying one’s movement data, is able to inform users both the amount and types of exercise one has done [43]. Similarly, transportation behaviours could be recognized by programmatic classifiers [67]. Water usage could be automatically categorized and aggregated by location of water use behaviours (e.g., kitchen versus washroom). However, automated techniques inevitably have flaws (discussion in Section 3.2).

Energy conservation and personal fitness might be the most common application domains of feedback systems. One of the characteristics of these feedback tools is that the feedback data are highly contextualized by one’s everyday activities and conditions. For example, geolocation might infer people’s physical activities [106]. Household energy consumption might be heavily related to the occupants’ domestic activities [56], e.g., cooking, entertaining, laundering, etc. Geolocations might infer
one’s exercise routines [106] (see more discussion in Section 3.4).

Meanwhile, ongoing adoption is an often encountered barrier in feedback designs in these application domains (chapter 2). People have lower usage of the feedback tools than expected, although there is a high percentage of people who own at least one feedback tool. Studies showed that people need to feel control of using their tools and that this impact needs to be recognized by oneself or others [7]. Adoption implies multiple phases in ongoing use. Kim [90] identified two stages of how people adopt everyday technologies: in the early stage, interest is the main motivation; then gradually the tool is adopted into daily routines. In a later stage, people’s 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. People’s 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 person’s life.

3.2 Persuasive Design

In feedback designs, persuasion is the most commonly used design strategy. Persuasion technology usually refers to Fogg’s model [63] that described strategies that people would employ to change behaviours or attitudes even without their understanding of the problem or behaviour, e.g., by computer generated suggestion [11]. Meanwhile, behavioural models were brought in from environmental psychology to study behaviour choices in feedback designs. Mostly these methods are based on rational choice or are norm activated [5, 68, 143, 78]. The rational choice approach assumes that people take an action (e.g., towards healthy life or pro-environmental 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 considers the social and culture environment people are situated in. For example, immediate feedback is used to engage behaviour change in the moment (e.g., UpStream [99] provides immediate feedback on water use to encourage people to take shorter showers) even without fully understanding the behaviour itself.

The persuasive technology approach has also been criticized [32, 142, 156]. Strengers et al. questioned the value of in-home display approaches [142]. They found that household behaviour change related to energy consumption cannot be modeled by one or two variables; instead, it is mediated in the context of everyday life, socially, cultur-
ally 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 behaviours are not isolated, rather they are situated in the social-cultural environment [32]. They questioned modernist approaches that human behaviours 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 people’s vision. Thus, they suggested a shift from prescription to reflection. Baumer [26] suggested an approach of “open-ended” awareness because people have varying definitions and assessment of “being healthy”. The findings of the study [143] suggested that feedback research needs to help people to understand technology ethnographically, identifying the origins, courses and influence of their behaviours with “contingent context”. For example, Dubuque is an open-ended system that engages people to reflect about household energy consumption within their own lives [60]. 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 behaviour coercion.

3.3 Ambient Visualization

The goal of the on-calendar design in this thesis is to publish information in an unobtrusive way in one’s existing information use ecosystem (personal digital calendar in this case). It requires that the visualization must be ambient and visually unobtrusive in the calendar frame most of the time, although it is always displayed on the background. Kim et al. discussed design requirements and advantages of ambient displays by comparing with direct information presentation with respect to how they provide feedback of personal activities in engaging sustainable life [92]. Their analysis indicated the essential characteristics of ambient displays in feedback design: subtle and non-interfering with primary tasks (e.g., time management tasks with digital calendar in this work).

With a systematic survey of research systems using ambient visualization, Poussman and Stasko [128] 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 house [23]. As decoration at home, a “tableau machine”, a feedback system to reflect domestic activities, artistically visualized the occupants’ trace at home [127]. Ambient displays also could use physical objects as visualization artifact, e.g., to show energy consumption of a public building [77]. Other than displays embedded in physical environment, designers also publish visualization in the 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 of mobile phones [67]. Another example is ShutEye [25] with visualization of sleep-related data as the wallpaper of one’s phone.

In this thesis, I investigated and applied the concept of “attentional ambience” to tackle a real-life problem; 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 “decentralized” by slightly changing the degree of attention as the task requires [22]. 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 feedback design could help to develop insights and enhance reasoning (Chapter 2). Contextual information is a critical factor in helping people recognize and understand information patterns [21, 51, 143]. Strengers [143] also suggested that energy use was mediated by everyday life context (“social, cultural, technical and institutional dynamics”), with which people can better understand their behaviour and evaluate their alternatives. Specifically, research suggested that technologies designed for behaviour change should consider the context and life routines where and how people use them [123, 98]. The study of Ahtinen [7] showed the importance of “personally relevant” contextual information for people to understand their exercise logs. Without sufficient context to understand the data, it can
be very difficult to make productive behaviour changes. Lack of personally related context for reflecting on one’s data can also prevent further use of the feedback tools [89].

Location has been often used in design as contextual information. Li [106] and Nadalutti [117] used geo location as context to help people understand their fitness exercises. Personal knowledge of water usage could be improved by narrative graphics [69]. Placing data in social context could also convey a better picture, e.g., to compare water consumption among family members [69] or with neighbors [60]. Another way of providing context in visualization is through linking related data, e.g., at-home activities and household energy consumption. Ellegård [56] displayed both occupants’ activities at home and electricity usage together along the timeline to indicate the cause of usage patterns.

Meanwhile, a calendar has its advantage 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 [151] revealed the patterns and the possible causes of energy use in the work place by clustering the consumption data within a calendar that indicated employees’ work patterns through the weeks and holidays. AffectAura [112] 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 their other work-related data (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 [101], seniors often looked up events on their wall calendars when they were asked to reason about situations in which 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 points out, HCI researchers usually have a different methodology compared with those in environmental psychology [68]. Aiming to imply future design, HCI research in behavioural feedback design primarily focuses on qualitative questions (e.g., how does the design work, why it works or does not work
in different conditions, etc.) rather than quantifying the effect on behavioural change. Thus, many studies obtained quick feedback from lab evaluation, and some of them employed field studies but with a small scale and short duration [68]. The natural inclination of HCI evaluation is to inform designs, which makes it unrealistic and unnecessary to apply the method of Randomized Controlled Trial that is commonly used in environmental psychology and medical informatics systems.

Evaluation is one of the major challenges of personal visualization design (Section 2.4.5). Such visualization tools are situated in everyday 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 everyday life: how people react, understand and use the artifact with personally relevant context in one’s life; how easy these artifacts fit into one’s life routines physically, socially and culturally.

In recent feedback design, researchers showed increasing interest in field study as the evaluation method. Many of these studies focused on evaluating specific systems or designs with measuring behaviour change or its outcome in a short duration, for example, emphasizing mostly the influence of decreasing energy usage or increasing exercises [99, 148, 43, 107, 38]. However, measuring behaviour change might be inaccurate and unnecessary [95]. The behaviour 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’ behaviours have truly changed. Instead, it would be more helpful to understand how and why the system or tool works. Such information could inform future design, especially for novel technologies at an early stage [143, 125, 49].

Considering these points, 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 behaviour data and how they react to the idea of integrating feedback data on their personal calendar.
Chapter 4

Visualization Design

What visual encodings are appropriate for representing time-varying quantitative feedback data as an attentionally ambient layer on a calendar? The challenges are to make the visualization salient enough without interfering with the digital calendar use and meanwhile ensuring data patterns are perceptible and comprehensible. For that, the early candidate design alternatives were prototyped with respect to three criteria: display location (overlapped or side-by-side), visual type (line graph, coloured region, luminance) and calendar scale (day, week and month views). Coloured region is line graph with shaded colour, a modification based on line graph. In this chapter, I discuss these alternatives as reflecting on visualization design principles in the literature.

Common representations of time-varying data 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 for visualizing data visualization novices typically prefer visual encodings that they are familiar with and commonly encounter in their lives [71], for example, data visualization encoded by position, length, size or colour. People who use PV&PVA might not have professional experiences or analysis skills of using visualization systems (Chapter 2). I suspected that applying these visual encodings that most people are familiar with would minimize the effort of learning and understanding. Meanwhile, effectiveness of these different visual encodings has been ranked in previous empirical studies and perception theories [108]. Visual encoding of spatial position dominantly reflects the user’s mental model of quantitative data (e.g., line chart or bar chart), but the graph ratio could possibly influence its perceptibility within the limited space of calendar views. Heer et al. evaluated horizon graphs and suggested to optimize the chart size
These known size effects encouraged me to consider different chart scales (day, week and month view) and evaluate the effects in the viability experiments.

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 colour (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) [47]. Based on these criteria, I considered line graph, coloured region (line graph with shaded colour) and luminance (I avoided hue to prevent possible colour interference effects since calendar activity entries are usually coloured) as candidate visual encodings for quantitative data in the on-calendar visualization.

As shown 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 coloured region (a combined case of position and luminance). 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. Coloured region (Figure 4.1 middle) is an alternative version of line graph with filled solid colour in the contour. Luminance (Figure 4.1 bottom) encodes data by grey scale. (Hue is avoided to prevent possible colour interference effects since calendar activity entries are usually coloured.)

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 coloured region always showed continuous data in all calendar views.

Two alternative positions for the visualizations were considered: overlapped and side-by-side. The overlapped option placed the visualization directly in the calendar cells overlapped with the calendar events (Figure 4.1); this is space efficient but may cause interference, especially when the calendar is busy with schedule events. In contrast, side-by-side option introduced an additional band parallel with calendar cells (Figure 4.2). I included it as an alternative, aimed to preserve effective graphical perception while minimizing interference between the two information layers. The width of the visualization band (side-by-side condition) was made smaller than the calendar cell, because the visualizations should not visually dominate the interface.
To enhance the visibility of luminance in the 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 visual interference and perceptibility of these design alternatives were evaluated in later experiments. Their viability of supporting attentional ambience 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 overlapped with calendar events (top: line graph; middle: coloured region; bottom: luminance)
Figure 4.2: Design alternatives displayed side by side overlapped with calendar events (top: line graph; middle: coloured region; bottom: luminance)
Chapter 5

Research Methods

Following the on-calendar design approach, I prototyped a few candidate design alternatives based on the visualization literature. The question remains: could these visualization design alternatives work? This leads to my later empirical investigation to evaluate the design approach. According to the characteristics of behaviour feedback design, what would be the appropriate choices of research method for the evaluation? Thus, the following chapters of this thesis focus on empirical studies in which the on-calendar design approach is evaluated. These studies are designed to answer the following questions:

1. Is it possible to create an on-calendar visualization of quantitative data that is comprehensible but does not interfere with primary calendar tasks? (RQ1)

2. How do people react to and use the on-calendar visualization as a feedback tool in everyday life? (RQ2)

3. To what extent can people use calendar events as context for reasoning about their personal feedback data? (RQ3)

4. Could providing additional context (e.g., related activities or conditions related to feedback data) in visualization improve people’s understanding of their behavioural feedback data? (RQ4)

In this chapter, I discuss the rationale of my method choices in the later empirical studies. Each method has its pros and cons. My choices of these research methods are based on the research questions in each phase of this work. Two main studies are presented in the following parts of the thesis: a viability study with lab experiments
and a formative design study with field deployments. I will discuss research methods used in these studies with their benefits and limitations in the following sections, and the further details are described in subsequent chapters.

5.1 Viability Study

Before the implementation, I designed a viability study (Chapter 6) to confirm my design concept that displaying the feedback data visualization on a digital calendar could support attentional ambience (RQ1) and inform the visualization choices that were later implemented. The viability study included two parts: two lab experiments (within-subject design in each of the experiments) and post-experiment questionnaire.

Lab experiments have been commonly used in evaluating information visualization designs, especially in perception studies [108, 80, 22]. I am also aware that online experiment could be another option to conduct lab experiments and may also meet these requirements [33]. However, it may not be able to provide contextual information about participants, e.g., through observation. Amazon Mechanical Turk \(^1\) was also excluded from being a possible choice due to ethics protocol requirements.

Typical metrics of visualization research are task completion time, task errors, and even insights [134]. Here the metrics I used for my lab experiments are task completion time and errors, assuming these would be correlated with perceptibility and legibility. For example, tasks would be completed with less time and fewer errors if the visualization is easier to perceive or less interfering. The data pattern and the tasks used in the experiments (Chapter 6 and Appendix A) were designed for perception and interference tests, respectively.

Participants were asked to filled in a questionnaire to report their experience with each visualization option, with respect to visual interference, perception and aesthetics (Appendix B). This information was used to cross reference with the quantitative results from lab experiments.

5.2 Design Study

After the design concept was verified, my interest was to investigate how people react to and use the on-calendar design in a real life context (RQ2-4). For that, I designed

\(^1\)https://www.mturk.com/mturk/welcome
field studies complemented with multiple interviews and weekly questionnaires.

A lab experiment was obviously not suitable in this case because the long-term engagement with the visualization tool was hard to study in a short-time lab session. Visualization used in everyday life did not require participants’ continuous focused attention all the time like usability tasks. The interaction and reasoning processes were embedded in an everyday usability task that was impossible to simulate in a lab environment. With contextual interview only, the ongoing nature would be impossible to investigate. Meanwhile, factors of behavioural change are difficult to control in the field, especially in this case where people had access to other tools and their behaviours could be mediated with individual difference or social influence, so controlled field experiments might not be a proper choice. Moreover, I designed a formative design study, aiming to address the open questions (RQ2-4) instead of quantifying the effect of the on-calendar design. For that, I ruled out Randomized Controlled Trial (Section 3.5).

The field study employed a between-subject design with interviews to investigate the experience of before and after deployment. I designed multiple interviews through the field studies to observe how participants used the visualization application, and to investigate how they referred to calendar context to reason about their feedback data. In-depth interviews tend to work well with ethnographic methods, e.g., case study or field study [104]. They were suitable for investigating open-ended questions in research. Specially in this work, the use of behavioural feedback tools was interwoven with culture and social influences (Chapter 2).

I primarily employed a qualitative analysis method in the design study. Applying quantitative methods in this case might limit researchers’ insight into how and why a certain design works, so the data were analyzed with a qualitative approach based on grounded theory through an open coding process [3]. This method was good at analyzing unstructured data, especially for interview transcripts and observations. After the field deployment, transcription and observation data were manually labeled. As patterns and trends of labels were investigated, these labels were merged into higher level categories. As this process was repeated, I developed a model to capture characteristics of behavioural feedback use.

To complement the qualitative method, weekly questionnaires (International Physical Activity Questionnaire [2]) were used to quantitatively evaluate participants’ physical activities. In addition, system logging scripts were used to track how people interact with the on-calendar application through the deployment. These multiple
data sources in the field study were helpful to triangulate the findings from the qualitative analysis.

5.3 Summary

Generally, evaluating visualizations used in everyday life is difficult and challenging (as previously mentioned in Section 2.4.5). The choices of research methods considered in this thesis are based on my research questions (RQ1-4) in each phase of my work. Specifically, this research is mainly aimed at an investigation of a reflective approach of visualization design in behaviour feedback. The research methods used in this work are based on this focus. In the next chapters, I will discuss the two primary studies where these methods applied.
Chapter 6

Viability Study

The on-calendar design alternatives were developed based on visualization principles, but it remains unclear whether or not they are effective. To support attentional ambience when embedding a visualization layer into one’s calendar, designs need to be visually salient but perceptible. This stage of my research evaluated the idea of an on-calendar design approach with proposed design alternatives. This evaluation is aimed to answer the question mentioned in Chapter 5:

- Is it possible to create an on-calendar visualization of quantitative data that is comprehensible but does not interfere with primary calendar tasks? (RQ1)

For that, controlled lab experiments are a good option to evaluate these perception and legibility criteria (Chapter 5). In this chapter, I describe the viability study, where I employed lab experiments to validate my design idea and to evaluate visual interference and perceptibility of design alternatives (Chapter 4). These experiments showed the promise of the proposed mash-up approach where additional visualization layers could be perceived ambiently with proper visualization choices (Contribution 4). It also suggested 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 appointments) (Figures 4.1 and Figure 4.2). The design goal is to present information in a way to support attentional ambience [22], informing people while blending into the environment and requiring low attentional demand [115]. In
other words, I “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 explore the viability of this approach as the first step to investigate the effectiveness of visualization in this design approach. 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 focus on these basic design issues related to interference and perceptibility rather than other factors influencing sustainable behaviours, such as motivation and social interaction.

6.2 Experiment Design

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

- Display location: overlapped (Figure 4.1) or side-by-side (Figure 4.2).
- Visualization type: line graph (Figure 4.1 top), colored region (Figure 4.1 middle) or luminance (Figure 4.1 bottom).
- Calendar scale: week \(^1\) (Figure 4.1 top middle) or month (Figure 4.1 top right).

6.2.1 Participants

Thirty one participants were recruited (14 female, 17 male) including undergrad and graduate students, with a diversity of backgrounds (computer science, engineering, engineering, engineering, engineering).

\(^1\)In the viability study, day view was not considered because it would use an identical visual encoding to the week view.
chemistry, biology, education, social science and political science). Fifteen of them participated in the first experiment and 16 participated in the second one.

6.2.2 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. I 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 A).

Experiment I was to test the following 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.

6.2.3 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 [134]. Furthermore, elementary (local patterns or extreme values) and synoptic tasks (overall estimate or distribution) are typical ways to explore temporal data [12]. Thus, I 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 A.

Experiment II was to test the following 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

- coloured region would be easier to perceive than luminance or line graph (H2.2), because position encoding is a stronger visual cue than luminance encoding and colour difference would enhance visibility of the contour.

6.2.4 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.

6.2.5 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 (Calendar Tasks described in Section 6.2.2) or data interpretation (Visualization Tasks described in Section 6.2.3), 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. To prevent colour effects, all visualizations were in grey (or with luminance varied). The coloured calendar event blocks on the calendar were semi-transparent (with alpha level at 0.6).
6.3 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.

Data of the two experiments were analyzed with respect to three factors: display location (overlapped or side-by-side), visualization type (coloured region, luminance or line graph) and calendar scale (month or week view). Task time was tested with repeated measures Analysis of Variance (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 binominal 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.

6.3.1 Experiment I: Calendar Tasks

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 coloured 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 6.2).

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.
Table 6.1: Total accuracy rates (%) with different visual encodings of three display conditions in Calendar Tasks and Visualization Tasks (The Visualization Tasks do not include a control condition)

<table>
<thead>
<tr>
<th>Display</th>
<th>Encoding</th>
<th>Experiment I</th>
<th>Experiment II</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Calendar Tasks</td>
<td>Visualization Tasks</td>
<td></td>
</tr>
<tr>
<td>side-by-side</td>
<td>line graph</td>
<td>90</td>
<td>78</td>
</tr>
<tr>
<td></td>
<td>coloured region</td>
<td>80</td>
<td>87</td>
</tr>
<tr>
<td></td>
<td>luminance</td>
<td>87</td>
<td>84</td>
</tr>
<tr>
<td>overlapped</td>
<td>line graph</td>
<td>40*</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td>coloured region</td>
<td>53*</td>
<td>87</td>
</tr>
<tr>
<td></td>
<td>luminance</td>
<td>87</td>
<td>97</td>
</tr>
<tr>
<td>control condition</td>
<td>no encoding</td>
<td>89</td>
<td>N/A</td>
</tr>
</tbody>
</table>

*Significant difference compared with the control condition
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. The figure shows that a lot of variability overshadows most of the differences. Pairwise comparisons between groups of Visualization Type showed that line graph took significantly more task time than coloured region ($p < 0.02$) Figure 6.2. 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 coloured 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%)($Q(1) = 6.82$, $p < 0.02$).
User Report with Questionnaire: Participants were asked to fill in a questionnaire with respect to their experience of visual distraction after the Calendar Tasks. Questions used in the experiments 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.

As shown in Figure 6.3, participants in Experiment I rated the overlapped condition more distracting than the side-by-side condition, and line graph was rated more distracting than coloured region and luminance. I found a significant main effect of Display Location ($F(1,14)=8.90, p = 0.01, \eta^2= 0.39$) but not Visualization Type. The interaction was not significant. Participants also reported that they could not stop being visually distracted by line graph, particularly with the overlapped condition.
Figure 6.3: Visual distraction reported by participants (-2 is “very distracting”, 2 is “not distracting”)

Summary of Experiment I

Coloured-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.

6.3.2 Experiment II: Graphical Perception

Task time: Response times for Experiment II are summarized in Figure 6.5. 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.

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 6.4. Pairwise comparisons showed that luminance took significantly less task time than coloured 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 coloured 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 coloured 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 difference ($Q(1) = 4.17$, $p < 0.41$).

User Report with Questionnaire: Participants were asked to fill in a questionnaire with respect to their experience of graphical perception after the Visualization Tasks. As shown in Figure 6.6, participants in Experiment II rated the overlapped
Figure 6.5: Boxplots showing task time of Experiment II)

condition as easier to perceive than the side-by-side condition. Line graph was rated the most difficult visualization to perceive. I found a significant main effect of Visualization Type ($F(2,30)=.04$, $p = 3.72$, $\eta^2 = .20$) but not Display Location. The interaction was not significant. Pairwise comparisons showed that coloured region was rated significantly better than line graph ($p < 0.01$).

Summary of Experiment II

The overlapped condition facilitated better graphical perception than the side-by-side condition. Luminance was easier to perceive than coloured 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).
6.3.3 Aesthetics

Participants were asked to fill in a questionnaire after the experiments with respect to their experience of visual aesthetics. The Likert scale ratings were analyzed similarly with repeated measures ANOVA.

Side-by-side visualizations were rated more appealing than overlapped visualizations, and line graph was rated the least appealing visualization (Figure 6.7). There were significant main effects of Display Location \( F(1,30)=5.15, \ p =0.03, \ \eta^2 = 0.15 \)
and Visualization Type \((F(2,60)=26.19, p < 0.01, \eta^2 = 0.47)\). The interaction was not significant. Pairwise comparisons showed that line graph was rated significantly lower than coloured region \((p < 0.01)\) and luminance \((p < 0.01)\).

6.4 Discussion of Lab Experiment Results

6.4.1 Interference (Experiment I)

The hypothesis (H1.1) that overlapped visualizations would have greater interference with calendar activities compared to side-by-side was confirmed. In Experiment I, task time was slightly faster with side-by-side 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 overlapped 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.

I hypothesized that 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 coloured region and luminance. I speculate that this interference might be caused by the similar colour (grey) of the line graph and activity text, which could make reading the text more difficult. With coloured region, the filled colour area tones down the interference to some degree. Thus, I eliminated line graph as a visualization option in the later implementation.

6.4.2 Perception (Experiment II)

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

The hypothesis (H2.2) that coloured 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 colour 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 coloured region represented continuous daily data.

6.4.3 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 I 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 I filled the entire cell with a grey 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 grey 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 that, this work offers some design suggestions to optimize visualization of quantitative data on a calendar.

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

This work investigated and applied 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 [22]. It is this capacity of the 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 (coloured 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 I can use a much more subtle level of 0.2 to provide a “just attendable difference” [22] 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’ calendars as well, e.g., the density of calendar events displayed, existing colour 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. I 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 suggested 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 \(^2\)). This is shown in Figure 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 laptop), tablet and mobile phone with a browser, but the layout was designed for desktop browser size (it was not customized for mobile devices). The application was hosted on the server at Simon Fraser University.

![Data flow of the on-calendar visualization system.](image)

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

In the application, personal calendars are placed in the 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 colours). The on-calendar application facilitates 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 enables users to select multiple calendars to display. The interactions of event editing are kept the same as Google Calendar, for example, clicking an event brings up a 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 coloured region, as shown in Figure 7.1, or luminance. To

---

1 https://developers.google.com/Google-apps/calendar/
2 https://dev.fitbit.com/
3 https://d3js.org/
balance the ambience of foreground calendar events and background data stream, the user is also allowed to adjust the settings of transparency, visualization colour (with grey as default) and chart scale. Meanwhile, a script is running at the backend to collect data of user interactions, 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 feedback data in real-life context. Meanwhile, I needed to investigate if the implementation followed the design goal and was inline with the results in the viability study. Thus, I conducted two case studies as pilot studies with the early version of the 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 describe two case studies where I deployed the prototype tool in an experimental eco-friendly town house and with a small group of students, respectively. The cases studies were aimed to test the early version of the implementation in real life situations and identify usability issues, which helped to revise the application for the later field study. As well, I present the implementation revisions at 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 (Figure 8.1) with the data source connected to their electricity meter. This experimental lane house was occupied by tenants and was equipped with different sensors to collect data of various utility consumption sources. The data feeds synchronized with the calendar tool consisted of the overall electricity consumption

¹https://www.sfu.ca/westhouse.html
for the home. Both of the tenants were musicians, had flexible working schedules and practiced at home regularly even on weekdays. I suspected that the calendar tool might be helpful to reveal the pattern of their at-home activities and the relationship between their consumptional behaviours and electricity use.

![Figure 8.1: Westhouse: eco-friendly smart home for the case study.](image)

The study had two phases: a baseline session and a prototype deployment session. The first session was to understand baseline energy consumption, and the participants were not asked 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 in the interviews were mostly about the participants’ understanding of their energy consumption behaviours, their experience of using the calendar tool in everyday life and how they used the contextual information 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 a 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 (she had a 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 a stove took less power than oven because using the oven usually took a longer amount of time. The participants also reported that 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. 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 tool saw 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 inconsistent look 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 made the look 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, I deployed the calendar application to 10 undergraduate students as part of a Psychology seminar course. In contrast to 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 relevant to individuals’ activities and directly linked to one’s personal calendar.

Participants (eight females and two males) in this study were undergraduate students from the Psychology department with age between 21 and 25. They were provided Fitbit devices (all of them were novice Fitbit users), with which they were asked to track their daily physical activities. Figure 8.2 showed two examples of Fitbit devices. However, online Fitbit accounts were not provided, so they could not use the default Fitbit online tool for feedback purposes. Instead, they could only use the
Figure 8.2: Fitbit trackers used in case study of personal fitness

display on the device for feedback information. The study lasted for 4 weeks. In the first two weeks, they could only rely on the device display for feedback. After the first 2 weeks, the on-calendar web application was introduced. In the latter 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, participants were asked about their experience in the first two weeks. I also helped participants set up the device and gave them a tutorial. In the second interview, participants were asked about their experience of using the on-calendar application and how they used the calendar context for reasoning (see the interview guidelines in Appendix C.2).

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, the start of the day, that they were sitting most of the time for classes, that they walked (or ran) during class intervals and that they studied during the weekend. It also helped to easily identify anomalies. For example, a participant found his active activities mostly happened at midnight. It was because he often worked in a restaurant with a night shift. The on-calendar visualization also revealed errors in data collection. One participant 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 the lab results in early viability experiments. They reported that coloured region was easier to perceive than luminance (“I can see the time, where the peaks are, and the intensity”). Visualization display overlapped (“overlapped” option) was regarded as easier to read because the chart was bigger, even though overlapped with calendar events.

However, part of participants’ Fitbit data was often found missing on the visual-
ization. They reported in the interview that they forgot to wear the device or charge the battery. The novice Fitbit users may not be familiar with using the devices or might have been uncomfortable wearing the device all the time. Thus, in the later deployment, experienced Fitbit users were recruited (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 had a busy schedule (because he had to manage a few student organizations) and used a digital calendar very frequently. During the study, he used the on-calendar visualization more than the others. I suspect this might be a factor influencing 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 inconsistent 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 seemed necessary.

8.3 Summary of Pilot Studies

These pilot studies aimed to collect feedback on the design approach and identify usability issues with the early 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, I 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 have been 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 needed 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 coloured region was that the spikes helped them recall things for reasoning. However, luminance displayed overlapped with calendar activities seemed to be most irritating for them. For these reasons I kept the visualization settings user-adjustable.

8.4 Design Revision

The revision of the implementation design primarily focused on the consistency of functions and look & feel with existing calendar tools (specifically, Google Calendar). The final version was implemented with PHP and JavaScript (D3.js for visualization layer), together with Fitbit API \(^2\) and Google Calendar API \(^3\) (Figure 8.3).

First, to make the application easy to manage, previously customized visualization settings are remembered for each individual. People could customize the visualization to fit themselves, and the on-calendar visualization was set with their previous preference the next time they logged in.

The on-calendar application was expected to be used as an everyday calendar tool, so I made the control panel (for adjusting visualization settings) more subtle and put it at the top. This was to reduce the distraction of non-calendar functions while people are performing scheduling tasks.

The calendar layout was redesigned according to the Google Calendar \(^4\) layout.

\(^2\)https://dev.fitbit.com/
\(^3\)https://developers.Google.com/Google-apps/calendar/
\(^4\)https://calendar.Google.com/calendar
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 show in the calendar view. Similar to Google Calendar, full-day events could be shown at the top of the day below the date label (only in day view and week view).

In addition, people were allowed to customize the row height of calendar cells. With smaller height, a longer time period of the day could be seen (only in day view and week view), in order to provide a better overview. In the consideration of privacy, functions of access control were enabled (e.g., automatic sign out of the user from the application).
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 the pilot studies are limited. The early implementations had 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 deployed it in a longitudinal field study. In this chapter, I describe an eight-week field study with the latest implementation connected to Fitbit data. This study primarily focused on these questions (mentioned in Chapter 5):

- How do people react to and use the on-calendar visualization as a feedback tool in everyday life? (RQ2)
- To what extent can people use calendar events as context for reasoning about their personal feedback data? (RQ3)
- Could providing additional context (e.g., related activities or conditions related to feedback data) in visualization improve people’s understanding of their behavioural feedback data? (RQ4)

The field study was designed as a formative design study, aimed to investigate the possibilities, effects and design problems of the on-calendar approach for implying future design, so the focus was not to quantify the effects of the intervention. I was also interested in the advantages and disadvantages of the on-calendar tool as used in everyday practice, compared with the feedback tool people previously used. For that I designed a control group to provide a baseline of people’s current practice of using feedback tools.
This study compared experimental and control groups (who used the calendar prototype and Fitbit’s 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 people’s experiences with the on-calendar visualizations rather than measuring differences in behaviour change, as suggested by early research [143, 125, 49, 95]. Therefore, I employed a qualitative approach with open-ended research questions rather than statistical comparisons between the groups (Contribution 5). With the results, I developed a new model of the behaviour feedback process to investigate the role of feedback tools (Contribution 6).

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 Fitbit’s 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 participated in this study with age ranging from 20 to 60+, 15 female and 6 male (Table 9.1). Two of them (one female and one male) dropped out after the first two weeks. Seven of them had used other fitness trackers before.

9.2 Conditions

Participants were randomly divided into two groups: Control (C1∼C9) and Visualization (V1 ∼ V10). This design allowed me to investigate whether extra context from a personal calendar could improve people’s 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 the web-based calendar visualization after week 2. Visualization group participants were asked to use the calendar application as their primary scheduling and feedback tool; however, they were not prevented from also using their default calendar service (e.g., Google
Table 9.1: Participants in Fitbit field study

<table>
<thead>
<tr>
<th>Participants</th>
<th>Age Group</th>
<th>Gender</th>
<th>Fitbit experience (current tracker)</th>
<th>number of fitness trackers used before current one</th>
</tr>
</thead>
<tbody>
<tr>
<td>p1</td>
<td>30-39</td>
<td>Male</td>
<td>2.5 years</td>
<td>0</td>
</tr>
<tr>
<td>p2</td>
<td>18-29</td>
<td>Female</td>
<td>1 month</td>
<td>0</td>
</tr>
<tr>
<td>p3</td>
<td>30-39</td>
<td>Male</td>
<td>1 year and 1 month</td>
<td>0</td>
</tr>
<tr>
<td>p4</td>
<td>30-39</td>
<td>Female</td>
<td>3 months</td>
<td>0</td>
</tr>
<tr>
<td>p5</td>
<td>18-29</td>
<td>Female</td>
<td>3 months</td>
<td>0</td>
</tr>
<tr>
<td>p6</td>
<td>40-49</td>
<td>Female</td>
<td>11 months</td>
<td>2</td>
</tr>
<tr>
<td>p7</td>
<td>18-29</td>
<td>Male</td>
<td>2 months</td>
<td>0</td>
</tr>
<tr>
<td>p8</td>
<td>30-39</td>
<td>Female</td>
<td>1 year and 1 month</td>
<td>2</td>
</tr>
<tr>
<td>p9</td>
<td>60+</td>
<td>Female</td>
<td>2.5 months</td>
<td>2</td>
</tr>
<tr>
<td>p10</td>
<td>18-29</td>
<td>Female</td>
<td>3 months</td>
<td>0</td>
</tr>
<tr>
<td>p11</td>
<td>40-49</td>
<td>Female</td>
<td>4 months</td>
<td>0</td>
</tr>
<tr>
<td>p12</td>
<td>18-29</td>
<td>Female</td>
<td>1 year</td>
<td>2</td>
</tr>
<tr>
<td>p13</td>
<td>18-29</td>
<td>Female</td>
<td>1 month</td>
<td>2</td>
</tr>
<tr>
<td>p14</td>
<td>30-39</td>
<td>Male</td>
<td>1.5 years</td>
<td>0</td>
</tr>
<tr>
<td>p15</td>
<td>30-39</td>
<td>Female</td>
<td>4 months</td>
<td>0</td>
</tr>
<tr>
<td>p16</td>
<td>60+</td>
<td>Female</td>
<td>1 year</td>
<td>2</td>
</tr>
<tr>
<td>p17</td>
<td>30-39</td>
<td>Female</td>
<td>4 months</td>
<td>0</td>
</tr>
<tr>
<td>p18</td>
<td>30-39</td>
<td>Female</td>
<td>3 months</td>
<td>0</td>
</tr>
<tr>
<td>p19</td>
<td>30-39</td>
<td>Male</td>
<td>3 weeks</td>
<td>2</td>
</tr>
</tbody>
</table>
Calendar or iCal) or Fitbit’s feedback tools. That means participants in visualization group may also use their original Fitbit application (i.e., the baseline application) during the deployment.

![Figure 9.1: Study procedure](image)

### 9.3 Procedure

Before the first week, I 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 (Figure 9.1). I interviewed all participants in week 3, during which Visualization group participants were introduced to the on-calendar visualization. This interview was to investigate how the participants use their feedback tools currently and set up the baseline of using a feedback tool for reasoning to compare before and after the visualization tool was introduced to the Visualization group.

To investigate their initial experience and help the participants on the technical issues of using on-calendar feedback application, I interviewed all participants again in week 5. The interview at this point was mainly for technical support purposes. I provided help if the participant had technique issues of using the on-calendar application, e.g., trouble shooting of the configuration. Participants in the 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 everyday 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 reason about the patterns and anomalies. These tasks were to investigate the general awareness of personal feedback data, how they reason about data patterns and anomalies.
using feedback tools and how they use inferential context in the reflection. Meanwhile, throughout the whole eight week study, participants were asked to fill in a weekly International Physical Activity Questionnaire (IPAQ, also see in Appendix H) [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 G.

The analysis was based on the first and the last interviews. I focused on investigating participants' general awareness before and after the deployment and how they used current feedback tools to reason about their feedback data. In the last interview, I also collected information about their experience and feedback related to using the on-calendar visualization tool.

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 used because of its incompleteness. (Fitbit cannot accurately capture activities such as cycling, spin class and swimming. In addition, Fitbit devices occasionally malfunctioned.) Instead, the estimated physical activity (PA) were evaluated by the weekly IPAQ survey, an established method to measure physical activities [7, 44, 28, 76, 34, 55]. V8's 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 [83]. 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 people's 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 weekly questionnaires entries were transformed as a continuous measure as MET minutes [2] (see the calculating protocol in Appendix I). I compared the weekly average physical activities (MET minutes) before and after the deployment. 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 behaviour 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 qualify in behaviour change and measuring behaviour change was not the main goal of this study. Instead, I focused the majority of my analysis on system use, its role in the feedback process and how it influenced people’s 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 participant’s change in average MET scores.](image-url)
9.5.2 System Use

Application logs showed 152 visits (user sessions) and 208 user interactions (setting and view changes) during the study in total among the participants in the Visualization group. The peak usage was in the morning (around 10am) and in the evening (around 9pm). 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 efficiency of the mash-up design approach: which is to use an additional visualization layer to support attentional ambience.

![Graph showing system usage](image)

Figure 9.3: System usage (top: system access versus time of a day; middle: total system usage and bottom: single session duration).

By default the visualization was encoded with coloured region and in grey colour
(with 60% transparency) and displayed with an overlapped layout in the week view (Figure 8.3). When the application was introduced, participants were asked to try and explore all possible settings. The application was implemented so that it could remember customized settings, so the bias of default visualization settings was minimized. Application logs (Figure 9.4) also showed that all participants preferred coloured region (line graph) as the visualization setting. They reported that luminance as the visual encoding required extra cognitive effort to understand, and that colour made the calendar look busy and interfered with calendar events (particularly when the calendar events were colour coded). Grey colour 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 (V3). 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 colour. 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 the 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 Behaviour Feedback Process

I transcribed the interview recordings and conducted content analysis [114]. 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 one’s goals and to changes in behaviour and barriers of current feedback use. These codes were then 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 (people’s understanding of their PA), behaviour choice (choices about when and how to engage in physical activity) and emotion (what emotion could be evoked in the process). I then used the data to build an understanding of relationships between these concepts. This analysis resulted in the behaviour feedback model illustrated in Figure 9.5.

Figure 9.5: Model of behaviour 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 used 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. It was also found in early studies [132, 58] that showed personal data tracking was usually goal driven. V2 had to manage a health condition, so he focused mostly 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 the 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). One’s 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 important. I 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 week’s 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. 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 behaviour 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, I'll 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 is 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 overnight. V6 could view the step data of her sister from the challenge feature (on the Fitbit application), and once identified some anomalies in her sister’s 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 the feedback model is a loop. Acting on behaviour 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 behaviour choices and revised personal goals. Similarly, Fritz et al. pointed out in their long term fitbit study [65] that feedback tools need to support the long term maintenance, for example, the changes of the metrics to reveal the proper “state”.

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), people’s monitoring measures, goals, reasoning strategies and influences on behaviours 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 reasoning for awareness and insights that might lead to re-evaluate personal goals or/and behaviour choices

9.5.4 Effects of the On-Calendar Visualization

In this section I discuss the effects of the on-calendar visualization with respect to the feedback model above. In the first and last interviews, I asked participants to recall and reason about their feedback data with similar processes, so I compared the data between the two interviews. The effects discussed in this section are mainly based on the last interview, compared with the first one.

**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 I’m at work it’s pretty clear peaks for my morning walk, my lunch time and my home break whereas with the kids it’s just sort of this little ... and when I’m 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 information to support the reasoning. In contrast, in the default Fitbit application (Appendix E) the daily goal was displayed as a reference line to compare with. That means that the context used in reasoning might be determined by personal goals, and this context (e.g., baseline to compare with) could be helpful when it is displayed with data. Although the month view with colour coding could show a brief overview of the month (Figure 4.1), participants rarely used it because the lack of labels (e.g., “5,000 steps”) or the overwhelming filled colour. This brings up a data granularity issue in visualization design (see further Discussion in Section 9.6).

**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. I 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 me 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)

**Behaviour choice:** I found that the calendar visualization might not provide direct actionable insights to instruct users’ behaviours; 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 with the on-calendar application 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 one’s 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 [150], 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, I 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 con-
textual 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’ per-
sonal calendars. All participants in the Visualization group reported that they liked
having their Fitbit data and life events aligned together on 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 partic-
ipants). 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 my intuition that personal calendar events could provide relevant 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.

In addition, the timeline of a day provided general context, for example, the time
to get up, to run for the 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 calendar, participants could glance routines across the week.
Table 9.2: Most frequently used contextual information for reasoning (normalized as frequency per participant).

<table>
<thead>
<tr>
<th>Context Information</th>
<th>Frequency (Control group)</th>
<th>Frequency (Visualization group)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schedules and holidays</td>
<td>2.6</td>
<td>3.1</td>
</tr>
<tr>
<td>Family events</td>
<td>2.4</td>
<td>1.0</td>
</tr>
<tr>
<td>Social activities</td>
<td>1.4</td>
<td>1.6</td>
</tr>
<tr>
<td>Location</td>
<td>0.7</td>
<td>1.7</td>
</tr>
<tr>
<td>Life routines not on the participants calendar</td>
<td>0.7</td>
<td>0.9</td>
</tr>
</tbody>
</table>

“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 let’s say this was roughly five or six p.m., I went straight home, but this is all information I’m 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 it’s 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 popula-
tion. 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. Lack of scien-
tific knowledge of investigating personal data might be a common pitfall of personal
feedback tools identified by many studies [41].

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 one’s personal goals. After reaching those goals, one’s 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 discon-
tinuation of tool use due to other reasons. I explore this matter through the 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
one’s 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;
the 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. I 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 background can even minimize the effort of physical view switch, for example,
bringing up another device or application. The mash-up design approach might make
it easily fit in one’s already-formed routines. V8 had the habit of keeping all tools
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. In this way,
the browser tab of the on-calendar feedback tool was usually on for hours, sometimes
for days (as long as the browser is open). This likely made him frequently encounter
his fitness data, possibly enhancing awareness. The study of Kim et al. [89] also shows similar evidence. Frustration arose when the data collection was not inline with people’s goals (doing what they expected it should do) or when data reliability was questioned. Abandonment points would happen in this stage [42].

Second, a large gap between the current state and one’s 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 one’s goal for a long time, frustration might make her/him drop the tool in the end. This indicates 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 the 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 behaviour 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. Previous studies [21, 20, 89, 100] showed that interest and engagement in using feedback tools like Fitbit wanes over time. There might be 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 behaviour 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 Feedback Tools for Seniors

It was interesting to include seniors in this study (V6 and C9). Surprisingly, they were early adopters of personal fitness trackers, with experience of more than one fitness tracker before. As they reported in the interviews, they actively use fitness feedback tools because they felt the increasing needs to manage physical aging. They did careful pre-research before the purchase, comparing different products based on their requirements. They were eager and also skilled to investigate the fitness data. They tried to apply data analysis skills to learn about their data, for example, exporting raw data for statistical analysis, designing their own analysis tools or visualization, etc. The intrinsic goal of coping with aging might be a strong motivation for seniors who are actively using feedback tools for fitness.

More importantly, I found that social factors in feedback tools might play a more important role for seniors than other age groups. They seemed 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 share their experience with families and friends. For that, feedback for seniors might also need to consider how to better facilitate these sharing needs.

The senior participants 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 other seniors.

However, these findings are based on two participants in the study and could not represent the large senior population. Also, these participants have been using fitness tools for a long time, so they may be a distinct subpopulation of the senior group. Although designing behavioural feedback for seniors is out of the scope of this work, I hope such research could get more attention in future.

9.6 Discussion

Based on the results, I reflect on the on-calendar design approach and feedback model. I also discuss the limitations 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 for full reasoning about this data. However, the study showed that the context provided by calendars was helpful in data interpretation. Meanwhile, it also showed that the frame of a digital calendar could provide general context for recall. The system logs and interviews showed that luminance was barely used during the deployment as the visual encoding, which is different from what I found in lab experiments where the participants rated luminance as their favourite. In the lab experiments, the displayed schedule was not participants’ own calendar schedules, and the color coding the schedule events were randomly selected from Google Calendar options, so the visual impact might not able to be revealed in the lab experiments. Moreover, the advantage of luminance in the lab experiments was from the month view. However, in the field, people mostly used the calendar tool with the week view, making the luminance coding less useful. This inconsistancy implied the importance of field deployments to investigate personal visualization tools used in everyday life.

The results showed that the usage of the on-calendar visualization depended strongly on participants’ existing information use behaviours, e.g., how often they use a digital calendar to manage their schedules and what events were on their calendar. One of the participants was used to keeping a browser tab (one of the tabs was Google Calendar) open all the time while working. Accessing the on-calendar application (by adding another browser tab) required little effort for him and resulted in the highest usage among all participants. He continued to use the application even after the study. Meanwhile, he maintained a dense calendar, with events about work, appointments, personal activities, and social events. His calendars could provide most of his life activities, so he can put pieces of his life back together by linking the feedback data and his calendar events. He reported in the interview that he quite enjoyed the process of reconstructing his life in the past, assisted by the on-calendar tool. On the contrary, participants with very sparse calendars seemed to use the application less, e.g., V4 who only had a few social events marked on her calendar had the lowest
usage from the logs. Thus, one’s existing information use behaviours 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 their daily lives, how often they use these calendar tools, what events are recorded on the calendar, how dense (or sparse) those events are on the calendar, etc. Meanwhile, the characteristics of the calendar influence attentional ambience of the visualization layers. When the calendar is dense and colour coding has been used for calendar events, they then try to make the visualization less salient but noticeable, e.g., bring down the transparency of grey colour. On the other hand, if the calendar is very sparse, an ambient mode is not necessary because the interference is limited. In this case, users can even change the visualization to a brighter colour.

User goals would impact on the effectiveness of the design as well. Li et al. categorized use goals based on the stages of using personal informatics tools: Discovery and Maintenance [105]. 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 that could fit best with their life choices. They may be not fully aware of their energy and exercise patterns or other factors that can influence their behaviours. Li et al. argued that providing and integrating multiple data source is 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 working at home versus workplace or what fitness plan could be manageable. The context from personal schedules on a calendar is provided for this purpose. In contrast to this, people at Maintenance already had their regular exercise routines, e.g., three runs a week, brisk walks after supper, etc. For that, they usually referred to feedback tools for a quick check and see if they accomplished their plan. The periodic nature of a calendar has advantages for supporting 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 at 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 past events and experiences that were not documented on the calendar at all. For example, the Fitbit data spikes at midnight invoked the participant’s memories about hiding eggs for the kids on Easter Sun-
day. Similarly, people could use it to plan fitness exercises that could better fit in one’s schedule. The mapping between scheduled events on one’s 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, emotional reminiscing was related to the experience with family or friends, similar as the findings in a recent study [150]. Possibly, feedback tools could be used to enhance social bonds or conversation other than facilitating tracking and awareness. One of the participants was enjoying the on-calendar visualization as it helped him practice memory by mentally 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. This may suggest that feedback design needs to go beyond single-purpose use and adapt to the multi-faceted nature of activities in everyday life.

In the field study I used time-varying fitness data as an example to explore the on-calendar design approach. If energy consumption data (used in early pilot studies Chapter 8) are applied here, the personal calendar may need to include more related context about at-home activities, for example, cooking, watching TV with families, etc. In this case, one solution is to keep daily diaries [56] to provide such contextual information. However, manually recording things with diaries definitely requires a lot of effort in data collection. Other personal feedback data may be also applicable to apply here as well, e.g., categorical data [147]. In that case, the visualization encoding might be different to make it ambient on a calendar. Participants also suggested that I integrate more data sources based on the calendar design, e.g., time-varying heart rate measure, calorie intake, psychological measures, etc. For that, appropriate data aggregation and segmentation may need to be considered to avoid overwhelming clutter.

9.6.2 Understanding Versus Coercion

This thesis explored a non-persuasive design approach in feedback design. Possibly, a persuasive design would be more effective to engage immediate behaviour change. Behaviour change is usually considered as the goal of feedback design. However, long-term behaviour change has been not confirmed in previous HCI research, and
the impact of persuasive design on sustaining behaviour change has barely been investigated [95, 32, 156]. People get to know the world and themselves by accumulating knowledge. Short-term behavioural change does not mean that the behaviour is sustained on a long-term basis. Learning behavioural knowledge and sustaining behaviour change are ongoing processes, and people have their own mental models in learning about how things work. An 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 appropriate baselines (Section 2.4.3) indicates that it is impossible to set a “standard” that can characterize everyone in all situations. Who should define and how should they define the “expected” behaviours to be coerced into?

According to The Transtheoretical Model (TTM) [129], behaviour change is a long-term process through a series of stages: pre-contemplation, contemplation, preparation, action and maintenance. Gathering information and understanding the behaviours would be the first step when one is willing and intends to change. As learning progresses about what happens, how it impacts on oneself and why it happens, one could transform one’s goal into an actionable plan for further action. The insights and knowledge perceived from information would also help to maintain the change. Studies show understanding the data is still a major barrier when people use feedback tools in everyday practice [21, 143, 119].

Moreover, understanding behaviours is not simply cause-effect analysis. One has to consider the context the behaviour is embedded into, physically, socially and culturally. Behaviour choices might not always rational but may be negotiable [143, 125]. That is, behaviour change also needs to consider people’s “changing expectations and aspirations” [143]. That is, in some circumstances people may not accept the “best” rational behaviour choice, but weigh other factors more than these “expected” actions, e.g., life comfort or convenience. Strengers et al. [143] gave the example of household laundry where people did not adopt the “good” behaviours because of their hygiene standard, even though it was against the goal of water conservation. In a recent study, Agapie et al. showed the value of “cheat” behaviours, behaviours as a lapsing allowance, viewing which can help reduce unwanted behaviours [6]. They suggested that designs need to include the collection of these unwanted behaviours and visualize them for lapse management.

As participants reported in this study, they might not be 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 ups and downs, so a consistent goal might not be reasonable. To better fit into one’s schedule, they might choose shorter vigorous exercises rather than longer light exercises. People might put the priority of kids or family events ahead of exercise. Feedback designs need to support people to reflect on a flexible set of alternatives. In this case, providing context for the reflection is crucial 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 to change behaviours. 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 few months, would this trend keep going in the next few months? Instead, the design goal of feedback technology should be helping people learn about their behaviours, understand their choices, and sustaining the change, even without feedback support in the end.

9.6.3 Related Models and Ongoing Use

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

Feedback designs are usually connected with behaviour change models. A variety of behaviour models have been studied in practice [5, 78, 44], but most of these focus on how to affect people’s motivation and attitude and consequently influence behaviour choices. However, the process of behaviour change involves long-term learning and is ongoing. My 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 one’s data does not necessarily result in behaviour change, I 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.

In relation to more general models, the Technology Acceptance Model [17] 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 behaviours outside
of tool use itself. A closely related but more specific model is the Promoter-Inhibitor Motivation Model (PIMM) [140]. 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: ongoing use of a feedback visualization to learn about and influence personal behaviour. 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 behaviour change nor the subsequent (circular) effects on goals and motivations for using a feedback tool.

As such, the feedback model enables designers to reflect on a non-persuasive approach. Tools that facilitate the reasoning process rather than enforcing behaviour change might be one step forward towards encouraging people to change their behaviour 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. A study of smart-device use showed that high cost of maintenance greatly discouraged people from ongoing use, suggesting design of smart devices needs to consider how they can easily fit in people’s routines and habits [100].

The interactions between design components in the feedback model could also help designers investigate barriers with respect to ongoing use. Epstein et al. investigated the reasons that people stop using self-tracking tools [57]. The characteristics of barriers identified in this study can be captured with the feedback model. Current “state” cannot be properly presented if cost of data collection is high or data quality is poor. Feedback tools might not match expectation, indicating gaps between one’s goal and the presentation of “state”. The link between behaviour and updating “state” could also be broken, invoking feelings of frustration that the feedback tool is not capable to show the effect of behaviour change. People might drop the tool if they cannot develop insights out of it to support behaviour choices. Meanwhile, emotional factors could mediate the whole flow as well, e.g., guilt, frustration, etc. However, ongoing use is not nonestop use. Knowledge and skills learned from feedback use may sustain the behaviour even after people abandon the tool. In this case the flow of the feedback model may not be supported by feedback tools but by people themselves internally. For example, people may recall on their weekly exercise routines, compare with their goal and reason about occasions when they break the routine.

Recent studies showed evidence inline with the feedback model. In a study of wearable devices, Kim et al. [89] categorized the stages involved in interacting with wearable devices. In the stage of “initiation & experimentation” people would match
their goal with what the device could do (i.e., represent their “state”). In the “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 feedback model is certainly constrained by the scale and nature of the 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 behaviour change. Nonetheless, I fully expect that it will be revised with more input in future work.

9.6.4 Design Implications

Traditionally, success of feedback tools has been mostly defined based on “behaviour change”. However, behaviour change is a long-term process and is influenced by many other aspects in one’s life. The short-term influence of using feedback tools could be easily overestimated. Instead, researchers and designers may need to view information use as a holistic ecosystem that is habituated through one’s everyday life, considering personal, social and cultural constraints [143]. When a new tool is introduced, it needs to fit into this ecosystem first and then impact on this system. This involves an ongoing process of interacting, learning and adapting.

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 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 participants’ initial curiosity and interest. Meanwhile, the presented information needs to reflect one’s goal and impact of behaviours appropriately. Goals might be dynamic even with the same person considering her/his varying situations (e.g., energy level) and customized for different groups (e.g., age groups).

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

Encouragingly, a recent design, Casalendar [113], adopted the idea of using a personal calendar as media to view and control a smart home environment, e.g., temperature, lighting, appliances, etc. The studies showed that viewing a calendar context side by side with the configuration and measurements helped participants easily identify and reason about anomalies. In a similar case, temperature variation in the workplace was displayed winthin a digital calendar, which helped employees reflect on the public policy of energy consumption [48]. These examples showed the premises of the design concept discussed in this thesis that a digital calendar could be an appropriate contextual framing for displaying feedback data. These studies also provided further evidences of the potential of the on-calendar visualizatoin.

Another implication is design for emotional engagement. From the study, I found that feedback tools could be a media that engaged people emotionally. It offered people opportunities to reminisce on personal or family trace, encouraging social conversation and interaction with friends and family. Rather than task-oriented information design, designers could also consider the design strategy for serendipitous information exploration [150]. The pleasure and satisfaction from this might inspire or enforce interest and curiosity that encourage ongoing use as well.

9.7 Limitations of the Study

As a formative design study, the field deployment provoked some open questions that I might not be able to conclude only with this study. For example, if the deployment could be applied in the household energy case, could the personal calendar provide enough context about the energy use behaviours? If multiple data sources are displayed, should the visualization be revised to reveal them all? How to make the design meet the needs of different level of data granularity? Is it possible to apply the design concept to similar information ecologies other than digital calendars? Nonetheless, this study showed promises for the integrated design approach and would be a good start for further exploration.

Since the application was independent (e.g., not implemented within Google Cal-
endar or iCal) and had limited features (e.g., no custom colouring 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 Fitbit’s 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 Fitbit’s feedback tools; it would have been difficult to constrain or track people’s use of the Fitbit application except through unreliable self-reports.

In addition, the field study was of a small scale and only focused on physical activity. While I anticipate that on-calendar visualizations could be used to display other 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.

9.8 Conclusion of the Field Study

The eight-week field deployment provided a chance to investigate the on-calendar visualization as a behavioural feedback tool in a real life context. The results showed the premises of this design approach. Especially it was helpful to identify and reason about feedback data patterns and anomalies with context provided by one’s personal calendar. The findings indicate that the effectiveness of this design approach could be highly influenced by how people use their personal digital calendar in everyday life. This suggests that behavioural feedback designs might need to consider one’s existing information use habits.

This study is a starting point for exploring how to integrate personal feedback 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 is considerable work to be done in the future in regards to feedback design and evaluation. In this study, personal feedback data (i.e., Fitbit data and home energy data) were displayed as continuous time-series data. Future design could explore data granularity in the visualizations as well. For example, one might wish to show an aggregated summary to reveal the performance compared to a daily goal. Such data could also be visualized by abstract graphics, for example, colour steps or icons. The aggregation could also be 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 facilitate social and emotional factors. People may expect to compare with others’ data on their display or share an interesting moment with family members or friends. It could also be used as a tool to encourage social conversations.

I also would like to see this design implemented in native calendar applications, e.g., Google Calendar, iCal or Outlook, with a customized personal feedback data stream on the additional visualization layer. In that case, people could get better exposure to their data, with which the impact 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 studies. The design approach could be totally reversed; context from calendar schedules could be added into other data applications as well, e.g., integrating calendar events into an energy billing report.
Studies showed people currently still have problems understanding their energy bill because of the lack of context [119]. For that, contextual 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 with data interpretation.

My evaluation mainly focused on personal fitness. Future studies could explore the design approach in other aspects of everyday life, for example, home energy conservation, health management, etc. It may also be applied to public education for sustainability [145]. Possibly, studies could also investigate on-calendar visualizations in professional use. For example, building managers could use it to map time-series utility 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 the user with a context for understanding and analyzing data and encouraging long-term use of these tools are current challenges of behavioural feedback designs. The goal of this thesis was to investigate the integration design approach by visualizing behavioural feedback data on a personal digital calendar. This approach is aimed to provide contextual information to support reasoning and understanding about data patterns and to support easy access to large amounts of personal data and thereby encourage ongoing use.

I systematically reviewed related academic work of personal visualization used in everyday life, with which I identified design dimensions and challenges in this field. This work introduces the design field of Personal Visualization and Personal Visual Analytics, and more importantly, it brings together research that was previously scattered in different disciplines and research fields.

Considering the design gaps in previous research and practice, I proposed to integrate personal feedback data (specifically, fitness data and household energy consumption data in this thesis) as an additional visualization layer on a personal digital calendar. This approach mainly tackles two challenges in behavioural feedback design: providing contextual information to reason about personal feedback data and supporting flexibility to fit in with everyday routines. People’s daily activities on a personal digital calendar could provide such context, and the familiarity of using a digital calendar could lower the cost of learning and adoption.

To further investigate the on-calendar approach, I first conducted a viability lab study to evaluate the interference and perception of candidate design alternatives of the on-calendar visualization. The lab study confirmed that the on-calendar visualization does not interfere with regular calendar use tasks with proper design choices,
and meanwhile it can be perceptible and comprehensible. The viability study also helped to narrow down design alternatives in implementation. The final implementation was a web-based application, synchronizing calendar data with Google Calendar and fetching personal feedback data either from Fitbit API or the data stream of a 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 the existing calendar application (Google Calendar in this case). After that, I conducted an eight-week field study to investigate how people react to and use the application in everyday life. In the field study, I applied the on-calendar visualization as a feedback tool for personal fitness by linking the data source with Fitbit trackers. Meanwhile, I primarily employed a qualitative method in data analysis for the field study. Based on the results, I derived a feedback model to illustrate the mechanism of feedback use and the role of feedback tools. Results from the field study showed that people like the idea of integrating behavioural feedback data on their personal digital calendars. The activities on their calendar could help them reason about the feedback data patterns and anomalies. People also used the tool creatively, e.g., planning exercises, logging related notes to explain data patterns, reconstructing their the life in the past by linking the feedback data with events on the calendar, etc.

The contributions of this thesis include: conducting the first systematic literature review of Personal Visualization and Personal Visual Analytics; proposing the on-calendar design that integrates personal feedback data on people’s personal digital calendar to provide context for reasoning and support easy access for ongoing use; investigating and applying the concept of attentional ambience in a real-life problem; conducting qualitative lab experiments to evaluate visual interference and perceptibility of design alternatives of on-calendar visualization; implementing and deploying the proposed approach in a longitudinal field study and developing a new feedback model of design components to inspect ongoing factors in feedback designs. The lab experiments confirmed the viability of the on-calendar design, and my field deployment showed the premises of this design concept. These results also suggest that behavioural feedback designers might need to consider people’s existing information use habits to encourage ongoing use.

This work demonstrates a reflective design approach (i.e., the on-calendar visualization) of visualizing behavioural feedback data. With this approach, feedback data
are displayed in a commonly used contextual frame (specifically, personal digital calendar in this work), which could help people better understand their feedback data and easily fit in with everyday routines. Encouragingly, I am glad to see the recent studies have started adopting this design approach and their results are inline with my findings in this work. This evidence showed the great potential of on-calendar visualization and its design concept in practical use.
Appendix A

Lab Experiment Tasks in Viability Study

A.1 Tasks in Experiment I of Viability Study
Q: What date is the tea house?

Q: How many TA meetings do you have in Feb?

Q: What day do you have dept. research workshop?
Q: How many meetings do you have with Derek this week?

Q: What date is the tea house?

Q: How many TA meetings do you have in Feb?
Q: What day do you have meeting with Neil?

Q: How many workshops do you have this week?

Q: What date is the dept. meeting?
Q: How many workshops do you have in Feb?

Q: What day is the committee meeting?

Q: How many seminars do you have this week?
Q: What date is the dept. meeting?

Q: How many tea houses do you have in Feb?

Q: What day is the TA meeting of CSC115?
Q: How many lectures of SENG310 this week?

Q: What date is the dept. retreat?

Q: How many workshops do you have in Feb?
Q: What day is the workshop@library?

Q: How many skype meetings this week?

Q: What date is the group lunch?
### Q: How many lab meetings do you have in Feb?

<table>
<thead>
<tr>
<th>Sun 2/19</th>
<th>Mon 2/20</th>
<th>Tue 2/21</th>
<th>Wed 2/22</th>
<th>Thu 2/23</th>
<th>Fri 2/24</th>
<th>Sat 2/25</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>8</td>
<td>2p</td>
<td>10</td>
<td>12</td>
<td>2p</td>
<td>4p</td>
</tr>
<tr>
<td>7</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>12p</td>
<td>1p</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>12p</td>
<td>1p</td>
</tr>
<tr>
<td>9</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>13p</td>
<td>1p</td>
</tr>
<tr>
<td>10</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>14p</td>
<td>1p</td>
</tr>
<tr>
<td>11</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>15p</td>
<td>1p</td>
</tr>
<tr>
<td>12</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>16p</td>
<td>1p</td>
</tr>
</tbody>
</table>

**Answer:** There are 5 lab meetings in February.

### Q: What day is the dept. meeting?

<table>
<thead>
<tr>
<th>Sun 2/19</th>
<th>Mon 2/20</th>
<th>Tue 2/21</th>
<th>Wed 2/22</th>
<th>Thu 2/23</th>
<th>Fri 2/24</th>
<th>Sat 2/25</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>8</td>
<td>2p</td>
<td>10</td>
<td>12</td>
<td>2p</td>
<td>4p</td>
</tr>
<tr>
<td>7</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>12p</td>
<td>1p</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>12p</td>
<td>1p</td>
</tr>
<tr>
<td>9</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>13p</td>
<td>1p</td>
</tr>
<tr>
<td>10</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>14p</td>
<td>1p</td>
</tr>
<tr>
<td>11</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>15p</td>
<td>1p</td>
</tr>
<tr>
<td>12</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>16p</td>
<td>1p</td>
</tr>
</tbody>
</table>

**Answer:** The dept. meeting is on Tuesday, Feb 21.

### Q: How many dept. meetings this week?

<table>
<thead>
<tr>
<th>Sun 2/19</th>
<th>Mon 2/20</th>
<th>Tue 2/21</th>
<th>Wed 2/22</th>
<th>Thu 2/23</th>
<th>Fri 2/24</th>
<th>Sat 2/25</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>8</td>
<td>2p</td>
<td>10</td>
<td>12</td>
<td>2p</td>
<td>4p</td>
</tr>
<tr>
<td>7</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>12p</td>
<td>1p</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>12p</td>
<td>1p</td>
</tr>
<tr>
<td>9</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>13p</td>
<td>1p</td>
</tr>
<tr>
<td>10</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>14p</td>
<td>1p</td>
</tr>
<tr>
<td>11</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>15p</td>
<td>1p</td>
</tr>
<tr>
<td>12</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>16p</td>
<td>1p</td>
</tr>
</tbody>
</table>

**Answer:** There are 3 dept. meetings this week.
Q: What date is the dept. meeting?

Q: How many workshops do you have in Feb?

Q: What day is the committee meeting?
Q: How many guest lectures do you have this week?
A.2 Tasks in Experiment II of Viability Study
Q: Do you consume more energy on Mar 3 than Feb 11?

Q: Which Friday do you consume the most energy?

Q: Do you consume more energy on Wednesday than Thursday?
**Q:** Which evening do you consume the most energy (7pm-9pm)?

**Q:** Do you consume more energy on Feb 10 than Feb 24?

**Q:** Which Friday do you consume the least energy?
1. Do you consume more energy on Feb 18 than Feb 19?

2. Do you consume more energy on Sunday than Tuesday during 12p-1p?

3. Which weekday do you (or your family) leave home the earliest?

---

Q: Do you consume more energy on Sunday than Tuesday during 12p-1p?

Q: Which weekday do you (or your family) leave home the earliest?

Q: Do you consume more energy on Feb 18 than Feb 19?
<table>
<thead>
<tr>
<th>Sun 2/19</th>
<th>Mon 2/20</th>
<th>Tue 2/21</th>
<th>Wed 2/22</th>
<th>Thu 2/23</th>
<th>Fri 2/24</th>
<th>Sat 2/25</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 am</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 am</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 am</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 am</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 am</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 am</td>
<td>10:30-12p Ted Meeting</td>
<td>11:00-12p Ted Meeting</td>
<td>11:00-12p Ted Meeting</td>
<td>11:00-12p Ted Meeting</td>
<td>11:00-12p Ted Meeting</td>
<td>11:00-12p Ted Meeting</td>
</tr>
<tr>
<td>11 am</td>
<td>11:30-12p VisID meeting</td>
<td>11:30-12p VisID meeting</td>
<td>11:30-12p VisID meeting</td>
<td>11:30-12p VisID meeting</td>
<td>11:30-12p VisID meeting</td>
<td>11:30-12p VisID meeting</td>
</tr>
<tr>
<td>12 pm</td>
<td>12:30-2:30p</td>
<td>12:30-2:30p</td>
<td>12:30-2:30p</td>
<td>12:30-2:30p</td>
<td>12:30-2:30p</td>
<td>12:30-2:30p</td>
</tr>
<tr>
<td>1 pm</td>
<td>2:30-3:30p</td>
<td>2:30-3:30p</td>
<td>2:30-3:30p</td>
<td>2:30-3:30p</td>
<td>2:30-3:30p</td>
<td>2:30-3:30p</td>
</tr>
<tr>
<td>2 pm</td>
<td>3:30-4:30p</td>
<td>3:30-4:30p</td>
<td>3:30-4:30p</td>
<td>3:30-4:30p</td>
<td>3:30-4:30p</td>
<td>3:30-4:30p</td>
</tr>
<tr>
<td>3 pm</td>
<td>4:30-5:30p</td>
<td>4:30-5:30p</td>
<td>4:30-5:30p</td>
<td>4:30-5:30p</td>
<td>4:30-5:30p</td>
<td>4:30-5:30p</td>
</tr>
<tr>
<td>4 pm</td>
<td>5:30-6:30p</td>
<td>5:30-6:30p</td>
<td>5:30-6:30p</td>
<td>5:30-6:30p</td>
<td>5:30-6:30p</td>
<td>5:30-6:30p</td>
</tr>
<tr>
<td>5 pm</td>
<td>6:30-7:30p</td>
<td>6:30-7:30p</td>
<td>6:30-7:30p</td>
<td>6:30-7:30p</td>
<td>6:30-7:30p</td>
<td>6:30-7:30p</td>
</tr>
<tr>
<td>6 pm</td>
<td>7:30-8:30p</td>
<td>7:30-8:30p</td>
<td>7:30-8:30p</td>
<td>7:30-8:30p</td>
<td>7:30-8:30p</td>
<td>7:30-8:30p</td>
</tr>
<tr>
<td>7 pm</td>
<td>8:30-9:30p</td>
<td>8:30-9:30p</td>
<td>8:30-9:30p</td>
<td>8:30-9:30p</td>
<td>8:30-9:30p</td>
<td>8:30-9:30p</td>
</tr>
<tr>
<td>8 pm</td>
<td>9:30-10:30p</td>
<td>9:30-10:30p</td>
<td>9:30-10:30p</td>
<td>9:30-10:30p</td>
<td>9:30-10:30p</td>
<td>9:30-10:30p</td>
</tr>
</tbody>
</table>

**Q:** Which Wednesday do you consume the least energy?

**Q:** Do you consume more energy on Monday than Thursday during 6-7p?

**Q:** Which evening do you consume the most energy (6pm-9pm)?
Q: Do you consume more energy on Feb 21 than Feb 25?

Q: Which Sunday do you consume the most energy?

Q: Do you consume more energy on Monday than Wednesday during 7-9p?
Q: Which weekday do you (or your family) get back home the latest in the afternoon?

<table>
<thead>
<tr>
<th>Sun</th>
<th>Mon</th>
<th>Tue</th>
<th>Wed</th>
<th>Thu</th>
<th>Fri</th>
<th>Sat</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
</tr>
<tr>
<td>2p</td>
<td>running</td>
<td>30</td>
<td>5pm</td>
<td>4pm</td>
<td>6pm</td>
<td>3:30pm</td>
</tr>
<tr>
<td>2p</td>
<td>tennis</td>
<td>5pm</td>
<td>4pm</td>
<td>6pm</td>
<td>3:30pm</td>
<td>3:30pm</td>
</tr>
<tr>
<td>2p</td>
<td>tennis</td>
<td>5pm</td>
<td>4pm</td>
<td>6pm</td>
<td>3:30pm</td>
<td>3:30pm</td>
</tr>
<tr>
<td>2p</td>
<td>tennis</td>
<td>5pm</td>
<td>4pm</td>
<td>6pm</td>
<td>3:30pm</td>
<td>3:30pm</td>
</tr>
<tr>
<td>2p</td>
<td>tennis</td>
<td>5pm</td>
<td>4pm</td>
<td>6pm</td>
<td>3:30pm</td>
<td>3:30pm</td>
</tr>
</tbody>
</table>

Q: Do you consume more energy on Feb 2 than Feb 15?

Q: Which Saturday do you consume the least energy?
Q: Do you consume more energy on Friday than Monday?

Q: Which day do you (or your family) get back home the latest in the afternoon?

Q: Do you consume more energy on Jan 30 than Feb 3?
Q: Which Sunday do you consume the least energy?

Q: Do you consume more energy in the evening (after 6pm) on Tuesday than Friday?

Q: What day do things start the latest in the morning?
Appendix B

Post-Experiment Questionnaire in Lab Experiments

B.1 Please rate visual distraction of each visualization option (Experiment I only)

Line chart displayed overlapped?

<table>
<thead>
<tr>
<th>very distracting</th>
<th>not distracting</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2</td>
<td>1</td>
</tr>
<tr>
<td>-1</td>
<td>2</td>
</tr>
</tbody>
</table>

Line chart displayed side by side?

<table>
<thead>
<tr>
<th>very distracting</th>
<th>not distracting</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2</td>
<td>1</td>
</tr>
<tr>
<td>-1</td>
<td>2</td>
</tr>
</tbody>
</table>

Coloured region displayed overlapped?

<table>
<thead>
<tr>
<th>very distracting</th>
<th>not distracting</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2</td>
<td>1</td>
</tr>
<tr>
<td>-1</td>
<td>2</td>
</tr>
</tbody>
</table>

Coloured region displayed side by side?

<table>
<thead>
<tr>
<th>very distracting</th>
<th>not distracting</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2</td>
<td>1</td>
</tr>
<tr>
<td>-1</td>
<td>2</td>
</tr>
</tbody>
</table>
Luminance displayed overlapped?
very distracting not distracting
-2 -1 0 1 2

Luminance displayed side by side?
very distracting not distracting
-2 -1 0 1 2

B.2 Please rate graphical perception of each visualization option (how difficult to perceive the data) (Experiment II only)

Line chart displayed overlapped?
very difficult not difficult
-2 -1 0 1 2

Line chart displayed side by side?
very difficult not difficult
-2 -1 0 1 2

coloured region displayed overlapped?
very difficult not difficult
-2 -1 0 1 2

coloured region displayed side by side?
very difficult not difficult
-2 -1 0 1 2

Luminance displayed overlapped?
very difficult not difficult
-2 -1 0 1 2
B.3 Please rate aesthetics of each visualization option (how appealing to perceive the data) (Experiment I and II)

Line chart displayed overlapped?

very appealing not appealing
-2 -1 0 1 2

Line chart displayed side by side?

very appealing not appealing
-2 -1 0 1 2

coloured region displayed overlapped?

very appealing not appealing
-2 -1 0 1 2

coloured region displayed side by side?

very appealing not appealing
-2 -1 0 1 2

Luminance displayed overlapped?

very appealing not appealing
-2 -1 0 1 2

Luminance displayed side by side?

very appealing not appealing
-2 -1 0 1 2
Appendix C

Interview Outlines in Pilot Study

C.1 Household Energy Consumption

- Which of the following best describes your home? ¹ (house/townhouse/apartment/other)

- How many bedrooms does it have? ¹

- Do you currently have air conditioning in your home? ¹

- What kind of heat do you currently use in your home? (Natural gas, Heating oil, Electric, heat pump, Other) Do you have full or partial central heating installed in your home? ¹ (full central, part central, no central)

- What is the energy source for cooking?¹ (Electric, gas, other)

- What is the energy source for hot water? ¹(Electric, gas, other)

- How much is the electrical bill per month (or per year)? What is the variance cross seasons?¹

- Which of the following household activities do you think uses most energy? (Heating water, Heating the home, Lighting the home, Use of electrical appliances) Please rank these from 1 to 4 where 1 = most energy and 4 = least energy.

¹interview1 only
- What appliances you have in your home? Do you leave on stand-by or power-off? Which you unplug when not in use?

- On your central heating system, what are temperature controls are in use? (Room Thermostat Thermostatic Radiator Valves Programed Thermostat Other) Do you know what they do? Do you know how to use them? (give an example)

- How would you rank the following energy using activities in terms of the least energy use to the most energy use. Please rank these from 1 to 5 where 1 = most energy and 5 = least energy. (Lighting Refrigeration Cooking Washing Working and Entertaining)

- what is the household electricity use pattern during a weekday/weekend?

- what is the household electricity use pattern during a week?

- How do you think your energy efficiency at home? (Very Efficient / Efficient / Neither /Inefficient /Very Inefficient)

- How do you usually keep track of your energy consumption? How often do you attend to this information? (If you have feedback device at home, where do you put it and how they use it? If not, what is the reason you did not use that?)

- How do you think of your current feedback devices or applications? Is it effective? is it convenient? why?

- What is the electricity use patterns for your household last month? why? (probe regular patterns and anomalies and ask participants to reason about that)

- What efforts did you make to reduce energy consumption?

- How did you use the on-calendar application? How often did you use it? Did you usually use it at work or home? Why? 

- Any interesting things have you found from the on-calendar application? (give an example)

- What visual settings did you usually use to customize the on-calendar visualization?

- What features do you like and dislike? why? Any barriers for you to use it?

\(^2\text{interview2 only}\)
C.2 Personal Fitness

C.2.1 Interview One

- Where do you usually spend most time of the day? what is your job/work like? need sit a lot? moving around? Are you living on campus? if not, how do you commute?

- How active do you think you are? What is your physical-activity level like in a weekday/weekend/week? Are those moderate or vigorous exercises?

- Have you ever tried to improve your activity level for health purpose? what are the pain points for you? what did you or will you do according to these barriers?

- Have you used any feedback tools? What did you track about yourself? Did it work? How and why (not)?

- Review Fitbit data of previous two weeks with Fitbit Website. (probe regular patterns and anomalies and ask participants to reason about that)

- How do you use Google Calendar? What events are marked on that?

C.2.2 Interview Two

- What were your physical-activity patterns in the past two weeks? Was anything different compared to the first two weeks? What was that?

- Review Fitbit data of previous two weeks with on-calendar application. (probe regular patterns and anomalies and ask participants to reason about that)

- How did you use the on-calendar application? How often did you use it? Did you usually use it at work or home? Why?

- Any interesting things have you found from the on-calendar application?

- What visual settings did you usually use to customize the on-calendar visualization?
- What features do you like and dislike? why? Any barriers for you to use it?
- Do you have any comment or suggestion?
Appendix D

Screenshots of On-Calendar Application (Final Version)
Figure D.1: Coloured region with day view

Figure D.2: Coloured region with week view

Figure D.3: Coloured region with month view
Figure D.4: Luminance with day view

Figure D.5: Luminance with week view

Figure D.6: Luminance with month view
Appendix E

Screenshots of Fitbit Web Application
Figure E.1: Dashboard view (aggregated summary data)
Figure E.2: Detail activity log
Appendix F

Background Questionnaire in Field Study

- Do you have a Google account?
- Do you use a Google Calendar?
- How often do you use a Google Calendar?
  - more than once a day
  - occasionally in a week
  - less frequent than once a week
- How do you use Google Calendar?
  - with browser on laptop or PC
  - with desktop app (e.g., iCal)
  - with mobile app
  - other
- Which way do you use Google Calendar most often?
  - with browser on laptop or PC
  - with desktop app (e.g., iCal)
  - with mobile app
  - other
- If you use Google Calendar with a browser, do you usually keep the tab (or window) open while doing other things?
  - I never use Google Calendar with a browser
- No, I close the tab (or window) right away after I finish.
- Yes, sometimes
- Yes, always
- Other

• What types of events are in your Google Calendar?
  - work events
  - personal life events
  - social events
  - other

• Do you share calendar with any other people?

• Who do you share your calendar with?
  - I don’t have a shared calendar
  - family
  - friends
  - coworkers
  - other

• Do you have a Fitbit device?

• What is your goal or motivation of using the Fitbit device?

• Do you use it recently?

• How do you usually check (or view) the Fitbit data?
  - website
  - mobile app
  - never checked the data
  - other

• Which one do you use most often to check (or view) the Fitbit data?
  - website
  - mobile app
  - never checked the data
  - other
• How often do you review your Fitbit data?
  - more than once a day
  - once a day
  - occasionally in a week
  - once a week
  - less than once a week

• What information do you care about from your Fitbit data?  - if I reach the goal
  - daily summary
  - weekly summary
  - monthly summary
  - granular temporal patterns
  - Other

• What is your age group?
  - 18-29
  - 30-39
  - 40-49
  - 50-59
  - 60+

• What is your gender
  - male
  - female
  - other

• How long have you been using Fitbit device?

• Is this your first Fitbit device? If not, how many have you been used?
Appendix G

Interview outlines in Field Study

G.1 Interview 1
Part A
- How long have you used the Fitbit device? Is it the first one? What was your motivation? Any health issue to manage?

- What is your typical day like related to physical activities? (ask participants about weekday, weekend and commuting)

- What recreational exercise do you do? (gym, sports, fitness training, etc.)

- During the week, what are your most active days? And what are your least active days during a week (or month)?

Part B
- Rate level of activeness (1-5). Do you think you are an active person? Why? (ask participants to give example of most active/inactive time) Are you happy with that? What is your goal of using the Fitbit device?

Part C
- How do you usually use your Fitbit? What application do you use for feedback (e.g., mobile app or website provided by Fitbit)? (ask participants to give examples and demo)

- which views of the application do you use most (and least), why? (show with examples)

- Ask participants to reflect on Fitbit data of the past two weeks (First without any feedback tools and then on Fitbit website)
  - See if they identify local and global patterns, anomalies.
  - What context do they use for making sense of the patterns and the anomalies (e.g., comparison, daily schedule, etc.)
  - If the Fitbit data are different from the recall (without the feedback tool) early, ask the participant to explain.
  - If the Fitbit data are different from the weekly survey data, ask the participant to explain)

- What are the barriers of using Fitbit device? (ask participants to give examples)

Part D
- Does Fitbit encourage you to more PA? Why? If not, what do you expect?

- What are barriers to be more active? Any strategies to cope with that in the past or in the future? Anything could be done to improve?
G.2 Interview 3
Part A:
- (without any feedback tool) what do you think of your physical activities in the past four weeks? Anything different compared with week 1-4? (ask participants to give examples)

- Ask participants to review Fitbit data (control group uses Fitbit website and visualization group uses on-calendar application)
  - See if they identify local and global patterns, anomalies.
  - What context do they use for making sense of the patterns and the anomalies (e.g., comparison, daily schedule, etc.)
  - If the fitbit data are different from the recall (without the feedback tool) early, ask the participant to explain.
  - If the fitbit data are different from the weekly survey data, ask the participant to explain)

B:
- Visualization group only:
  - how often do you use the tool? in what situation (e.g., at work, at home, etc.)? with what device?
  - how do you usually use it? Please give some examples.
  - In what circumstances do you bring up the app, managing calendar or viewing data? Why?
  - Do you often keep the tap open? Why
  - What visualization settings have you tried (colour, display location, visual encoding, saturation, scale, view, etc.)
  - Do you find anything interesting would like to share with me? (show with examples)
  - How do you interpret Fitbit data with calendar schedule? (show examples)

   b. Control group only:
   - How did you use Fitbit application in the past 4 weeks? How often do you use it?
     In what situation (e.g., at work, at home, etc.)? with what device?

   - What views do you most (and least)? Why? (show with examples)

   - Do you find anything interesting would like to share with me? (show with examples)

C:
- What requirements do you think the current tool cannot meet? Why?
- What features do you like? What features you don’t like? Why? (show with examples)
- What are barriers of using it?
- Do they share the experience with others (family, friends)? How? Why?
- Is there any other feedback tools have you used? What are they? (show with examples)
- (Control group only) introduce the on-calendar feedback application
- Any other comments or suggestions?
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? Please do not include walking.
   - [ ] ___ 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.

LONG LAST 7 DAYS SELF-ADMINISTERED version of the IPAQ. Revised October 2002.
Appendix I

Protocol for IPAQ Long Form

The long form of IPAQ asks in detail about walking, moderate-intensity and vigorous-intensity physical activity in each of the four domains.¹

I.1 Continuous Score

Data collected with the IPAQ long form can be reported as a continuous measure and reported as median MET-minutes. Median values and interquartile ranges can be computed for walking (W), moderate-intensity activities (M), and vigorous-intensity activities (V) within each domain using the formulas below. Total scores may also be calculated for walking (W), moderate-intensity activities (M), and vigorous-intensity activities (V); for each domain (work, transport, domestic and garden, and leisure) and for an overall grand total.

I.2 MET Values and Formula for Computation of MET-minutes

I.2.1 Work Domain

- Walking MET-minutes/week at work = 3.3 * walking minutes * walking days at work

¹accessible at https://sites.google.com/site/theipaq/scoring-protocol
• Moderate MET-minutes/week at work = 4.0 * moderate-intensity activity minutes * moderate-intensity days at work

• Vigorous MET-minutes/week at work = 8.0 * vigorous-intensity activity minutes * vigorous-intensity days at work

• Total Work MET-minutes/week = sum of Walking + Moderate + Vigorous MET-minutes/week scores at work

### I.2.2 Active Transportation Domain

• Walking MET-minutes/week for transport = 3.3 * walking minutes * walking days for transportation

• Cycle MET-minutes/week for transport = 6.0 * cycling minutes * cycle days for transportation

• Total Transport MET-minutes/week = sum of Walking + Cycling MET-minutes/week scores for transportation

### I.2.3 Domestic and Garden [Yard Work] Domain

• Vigorous MET-minutes/week yard chores = 5.5 * vigorous-intensity activity minutes * vigorous-intensity days doing yard work (Note: the MET value of 5.5 indicates that vigorous garden/yard work should be considered a moderate-intensity activity for scoring and computing total moderate intensity activities.)

• Moderate MET-minutes/week yard chores = 4.0 * moderate-intensity activity minutes * moderate-intensity days doing yard work

• Moderate MET-minutes/week inside chores = 3.0 * moderate-intensity activity minutes * moderate-intensity days doing inside chores

• Total Domestic and Garden MET-minutes/week = sum of Vigorous yard + Moderate yard + Moderate inside chores MET-minutes/week scores

### I.2.4 Leisure-Time Domain

• Walking MET-minutes/week leisure = 3.3 * walking minutes * walking days in leisure
• Moderate MET-minutes/week leisure = 4.0 * moderate-intensity activity minutes * moderate-intensity days in leisure

• Vigorous MET-minutes/week leisure = 8.0 * vigorous-intensity activity minutes * vigorous-intensity days in leisure

• Total Leisure-Time MET-minutes/week = sum of Walking + Moderate + Vigorous MET-minutes/week scores in leisure.

I.2.5 Total Scores for all Walking, Moderate and Vigorous Physical Activities

• Total Walking MET-minutes/week = Walking MET-minutes/week (at Work + for Transport + in Leisure)

• Total Moderate MET-minutes/week total = Moderate MET-minutes/week (at Work + Yard chores + inside chores + in Leisure time) + Cycling Met-minutes/week for Transport + Vigorous Yard chores MET-minutes/week

• Total Vigorous MET-minutes/week = Vigorous MET-minutes/week (at Work + in Leisure)

Note: Cycling MET value and Vigorous garden/yard work MET value fall within the coding range of moderate-intensity activities.

I.2.6 Total Physical Activity Scores

An overall total physical activity MET-minutes/week score can be computed as:

• Total physical activity MET-minutes/week = sum of Total (Walking + Moderate + Vigorous) MET-minutes/week scores.

This is equivalent to computing:

• Total physical activity MET-minutes/week = sum of Total Work + Total Transport + Total Domestic and Garden + Total Leisure-Time MET-minutes/week scores.

As there are no established thresholds for presenting MET-minutes, the IPAQ Research Committee proposes that these data are reported as comparisons of median values and interquartile ranges for different populations.
Bibliography


[23] Lyn Bartram, Johnny Rodgers, and Rob Woodbury. Smart homes or smart occupants? supporting aware living in the home. In *Proceedings of the 13th*


Parmit K. Chilana, Andrew J. Ko, and Jacob Wobbrock. From User-Centered to Adoption-Centered Design: A Case Study of an HCI Research Innovation Becoming a Product. In *Proceedings of the 33rd Annual ACM Conference on


[59] Thomas Erickson, Ming Li, Younghun Kim, Ajay Deshpande, Sambit Sahu, Tian Chao, Piyawadee Sukaviriya, and Milind Naphade. The dubuque electric-


[74] Michael Haller, Christoph Richter, Peter Brandl, Sabine Gross, Gerold Schossleitner, Andreas Schrempf, Hideaki Nii, Maki Sugimoto, and Masahiko


[82] Dandan Huang, Melanie Tory, Bon Adriel Aseniero, Lyn Bartram, Scott Bate-
man, Sheelagh Carpendale, Anthony Tang, and Robert Woodbury. Personal
Visualization and Personal Visual Analytics. *IEEE Transactions on Visualiza-

testing, exercise prescription, and evaluation of functional capacity. *Clinical

[84] 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.

[85] Victor Kaptelinin and Bonnie A. Nardi. *Acting with Technology: Activity The-

[86] Evangelos Karapanos, John Zimmerman, Jodi Forlizzi, and Jean-Bernard
of the SIGCHI Conference on Human Factors in Computing Systems*, CHI ’09,
pages 729–738, New York, NY, USA, 2009. ACM.

[87] Logan Kendall, Dan Morris, and Desney Tan. Blood Pressure Beyond the
Clinic: Rethinking a Health Metric for Everyone. In *Proceedings of the 33rd
Annual ACM Conference on Human Factors in Computing Systems*, CHI ’15,
pages 1679–1688, New York, NY, USA, 2015. ACM.

of a system to support record-keeping for parents of young children. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Sys-
tems*, CHI ’09, pages 1713–1722, New York, NY, USA, 2009. ACM.

[89] Da-jung Kim, Yeoreum Lee, Saeyoungh Rho, and Youn-kyung Lim. Design
Opportunities in Three Stages of Relationship Development Between Users and
Factors in Computing Systems*, CHI ’16, pages 699–703, New York, NY, USA,
2016. ACM.


