UNIVERSITY OF SOUTHAMPTON

FACULTY OF PHYSICAL AND APPLIED SCIENCES

Electronics and Computer Science

Understanding and Exploring People's Food Beliefs to Design Healthy Eating Applications

by

Feng Gao

Thesis for the degree of Doctor of Philosophy

January 2014

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF PHYSICAL AND APPLIED SCIENCES
Electronics and Computer Science

Doctor of Philosophy

UNDERSTANDING AND EXPLORING PEOPLE'S FOOD BELIEFS TO DESIGN HEALTHY EATING APPLICATIONS

by Feng Gao

Wellness is a domain of growing interest in computing. Many interventions are designed to modify people's behavior to make them healthy. But often, they ignore people's beliefs and socioeconomic context so no sustained change are achieved. This thesis, therefore, focuses on rethinking how we can leverage people's food beliefs to design healthy eating systems that can achieve sustained change.

The thesis begins with a literature review on why healthy eating is complex and how it depends on various contexts including people's food beliefs. Then we reflect upon current design notion of healthy eating applications and acknowledge the current designs can not address all the complex of healthy eating. Thus, we should explore the alternative design notion, which tries to leverage what people think about food.

A study on diet-related topics and social interaction patterns on forums led us to the idea that crowd's wisdom on food's healthiness could help people to better understand food and potentially encourage them to change eating behavior. Thus, we designed and evaluated one early prototype that allows people to judge each other's food's healthiness by giving ratings and checking 5 predefined heuristics. The outcome of this prototype study confirmed that people like the idea of expressing own thoughts about food and like to have others' thoughts to validate owns. We further explored exactly what kinds of food beliefs people hold via a survey study with real food pictures. And we found 8 categories of food beliefs and also validated that even experts can not reach consensus on food's healthiness and thus there is clearly a gap between what we know from nutrition literature and how interventions are actually designed. To further leverage people's food beliefs, we need to first capture it. We developed three simple UIs to capture such food beliefs, and evaluated the UIs from three different perspectives. We also found two patterns from the data we collect which can help us better visualize the data and identify group of people. In the end of the thesis, we discuss the thesis work and present the future works.

Contents

Declaration of Authorship **ix**

1 Introduction **1**
1.1 Motivation and Objectives . . . 1
1.2 Research Approach . . . 4
1.3 Research Outcomes . . . 5
1.4 Research Publications . . . 5
1.5 Thesis Walkthrough . . . 6

2 Backgrounds and Related Works **9**
2.1 Wellness Application and Its Challenges in Human Computer Interaction 9
2.2 The Healthy Eating Problems and Health Eating Guidelines . . . 12
2.3 The Complexity of Everyday Eating . . . 14
2.4 People's Interpretation of Healthy Eating . . . 16
2.5 Health-Related Behavior Change Theories . . . 17
2.5.1 Individual Behavior Change . . . 17
2.5.1.1 Transtheoretical Model . . . 17
2.5.1.2 Health Belief Model and Theory of Planned Behavior . . 18
2.5.1.3 Summary . . . 19
2.5.2 Interpersonal . . . 19
2.5.2.1 Social Cognitive Theory . . . 19
2.5.2.2 Social Support Theory . . . 20
2.5.2.3 Social Influence Theory . . . 21
2.5.3 Community-based . . . 22
2.6 Behavior Change Technology . . . 23
2.6.1 Self-Monitoring . . . 23
2.6.2 Goal Setting . . . 25
2.6.3 Information Provision . . . 27
2.6.4 Social Support and Social Influence . . . 29
2.7 Discussion: Beyond Persuasion . . . 30
2.7.1 Alternative Design Notions . . . 31
2.7.2 The Approach: Leveraging People's Thoughts to Support Behavior Change . . . 33

3 Initial Study: Investigating What People Talk About on Diet-related Forums **37**
3.1 The Study Design . . . 38
3.1.1 Data Collection Method and Dataset Overview . . . 38

3.1.2 Data Analysis Method . . . 39
3.2 The Findings . . . 39
3.2.1 Contextual Information and Personal Conditions . . . 40
3.2.2 You and Me: We Are Similar . . . 40
3.2.3 Sharing Similar Experiences . . . 40
3.2.4 Emotion: Loving Food . . . 41
3.2.5 Emotion: Enjoying Diet . . . 42
3.2.6 Share Opinions on Products and Information . . . 42
3.3 Discussion . . . 43
3.3.1 Rethink about Food Logging Technology . . . 43
3.3.2 Think about the Social Supports . . . 45
3.4 Conclusion . . . 46

4 Design Prototype 47
4.1 Rationale Behind The Design . . . 47
4.2 The Detail Design and First Prototype . . . 49
4.3 Deployment and User Study . . . 51
4.3.1 First User Study . . . 52
4.3.1.1 The Study Design . . . 52
4.3.1.2 Participants . . . 52
4.3.1.3 Pre-study Interview . . . 52
4.3.1.4 The Deployment . . . 53
4.3.1.5 Lesson Learnt from this study . . . 54
4.3.2 Open Deployment . . . 55
4.4 Similar Applications in the Real World . . . 56

5 Understand How People Evaluate Foods Healthiness and Its Design Implications 59
5.1 Method . . . 60
5.2 Study Results . . . 61
5.2.1 The Relation between Attributes Usages, Ratings and Demographics 66
5.2.1.1 Differences in Ratings between Two Groups . . . 67
5.2.1.2 Differences in Attributes Usages between Two Groups . . 68
5.3 Follow-up Study . . . 69
5.3.1 Participants . . . 69
5.3.2 Results . . . 70
5.4 Discussion . . . 71
5.4.1 Negotiable System for Dietary Change . . . 73
5.4.2 Support Reflection on Healthy Eating Understanding and Decision-Making . . . 75
5.4.3 In Situ Information Support: Shopping and Preparation . . . 76
5.5 Summary . . . 77

6 Capturing and Visualizing What People think about Food's Healthiness 79
6.1 Study: User Interface Design for Capturing Rationale behind Ratings . . 80
6.1.1 UI Design: Data Capture . . . 81
6.1.2 Methods . . . 83

6.1.3 Experiment 1: How Does UI Type Impact Time on Task and Data Quality? . . . 83
6.1.3.1 Participants . . . 84
6.1.3.2 Results . . . 85
6.1.4 Investigating the Potential for Presentation of the Collected Data . 88
6.1.5 Visualization Design: Present Food Knowledge . . . 88
6.1.6 Experiment 2: Does Priming With Collective Crowd Wisdom Influence Ratings? . . . 89
6.1.6.1 Participants . . . 89
6.1.6.2 Results . . . 89
6.1.7 Experiment 3: Does Reflecting On Collective Crowd Wisdom After Rating Change Ones Decisions? . . . 90
6.1.7.1 Participants . . . 90
6.1.7.2 Results . . . 90
6.1.8 Experiment 4: Do 7-Point Ratings Differ If Provided Nutrition Label Instead of Visualization? . . . 91
6.1.8.1 Participants . . . 92
6.1.8.2 Results . . . 92
6.2 Discussion . . . 92
6.2.1 User Interface Design for Capturing Rationales behind Ratings . . 93
6.2.2 New Knowledge Captured by Predefined list: Patterns of Attribute Distribution . . . 95
6.2.3 Towards a data infrastructure for future food-related systems . . . 96
6.3 Summary . . . 98

7 Conclusion **101**
7.1 Thesis Summary . . . 101
7.2 Contribution . . . 104
7.3 Future Work . . . 105
7.4 Closing Remarks . . . 107

Appendix A Larger Figures **109**

References **113**

Declaration of Authorship

I, Feng Gao , declare that the thesis entitled *Understanding and Exploring People's Food Beliefs to Design Healthy Eating Applications* and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- none of this work has been published before submission

Signed:.....

Date:.....2015.2.6.....

Chapter 1

Introduction

1.1 Motivation and Objectives

Health problems such as diabetes and obesity are now considered to be "epidemic" by World Health Organization (WHO) (James, 2008) and become much worse than last decade. One way to prevent those diseases is to engage general public in practicing health behaviors such as eating more vegetables and increasing physical activities. Thus a good intervention for wellness is necessary. Until now, various efforts, from mass media education to local support group, have been made. Those traditional interventions for wellness to help people change behavior have been proven to be effective at certain scale and level but better solutions are still needed. With the growth of computing technology, technology-based wellness interventions (especially mobile interventions) become promising in the problem space of health behavior change and have brought both researchers' and practitioners' eyes.

Persuasive Technology, which aims to help people change behavior through persuasion and social influences (Fogg, 2002), has been widely adopted and applied in both academic and commercial domains (Consolvo et al.; Grimes et al., 2010; Mamykina et al., 2008). One basic idea of persuasive technology is to simplify certain tasks such as food logging and nutrition assessment and convey persuasive messages to users. Thanks for the current success of Ubiquitous Computing technology, which enables computing to be possible everywhere and anywhere, various sensors are available to collect data about human behavior and health status. Such so called self-monitoring or quantified-self notion is popular in today's design and it does offer both researchers and users a good opportunity to collect enough data about an individual's behavior and assess those behavior based on quantitative standards (e.g. daily calories requirement). Besides, those systems also try to offer feedbacks or suggestions based on behavior change theories and guidelines. Social support also has been taken into account as social networking

is so popular today. (More Details on using those different technologies and successful examples are presented in Chapter 2.)

Despite that those systems achieved certain success, some researchers have criticized that systems in the category of persuasive technology have problems in taking away opportunities to practice skills from users (Purpura et al., 2011), not addressing external barriers and resources (Maitland et al., 2009), not addressing people's values and beliefs (Halko and Kientz, 2010), and making people over-rely on technology (Purpura et al., 2011) (Detail arguments regarding these criticisms are discussed in Chapter 2). Those criticism lead to the question whether those persuasion-based interventions can assist users in practising and maintaining healthy behaviors in a sustained way. In fact, most of successful persuasion-based interventions are not evaluated in a long-term study. And even among few long-term studies, participants reported that they find it hard to maintain the behaviors (Halko and Kientz, 2010). Therefore, it is vital to re-think how can we design wellness applications that can help people change their behaviors in long-term, sustained way.

To potentially address and tackle limitations of persuasion-based interventions, alternative design notions have been proposed by different researchers. One of them is the notion of reflection (Purpura et al., 2011; Mamykina et al., 2008) that advocates technology should supply time and create opportunity for users to think critically and practice necessary skills. Similar notion is the mindful technology (Munson, 2011; Chen, 2011) but it more focuses on examining what people think and feel and try to find a way to let users could express those feeling and thoughts. Maitland's negotiation framework (Maitland, 2009) explores another way to design behavior change technology, in which Maitland highlights the importance of addressing external constrains and resources and suggests the behavior change should be a process of negotiation between stakeholders: experts, patients, policy makers, etc.

The three alternative notions mentioned above have one thing in common that people's knowledge and beliefs can influence their behavior. And this point is supported by the well-known theory of planned behavior (Ajzen, 2002) and Oinas-Kukkonen's work (Oinas-Kukkonen, 2013), those previous works thus lay the general foundation for this thesis: people's behavior are driven by their thoughts and beliefs and only when thoughts and beliefs are changed the behavior might be changed in a sustained way. Based upon this general assumption, we thus believe it is important to understand people's thoughts and beliefs about one particular health behavior so the wellness application can 1) align its intervention with people's thoughts and beliefs so people will feel more comfortable to adopt the new behavior, take baby-steps and keep practicing the healthy behavior and 2) possibly identify problematic thoughts and beliefs so the intervention can help people change them and achieve sustainable behavior change.

In this thesis, we will reflect on the success and challenges of persuasion-based systems and try to explore the question that how can we leverage people's thoughts and beliefs to create wellness applications that enable sustainable behavior change. Particularly, since the author has special interest in food and dietary change, this thesis will explore the question in the context of dietary change. The dietary change mainly involves the change of one's choice of food intakes and we believe that taking advantage of people's thoughts and beliefs around food is one key method to help people eat healthy and improve their food knowledge for better food decision making. The specific question of the thesis, therefore, is **how can we capture and then leverage people's thoughts and beliefs about food to potentially better personalize the intervention and engage people in adopting to the new behavior in a sustained way?** To break it down, the thesis asks following four questions:

- How are people's thoughts about food different from the message conveyed by current healthy eating applications? and how does it create a gap between nutrition theory and real world wellness application design?

  As we briefly discussed above and will elaborate further in Chapter 2, current applications inspired by persuasive method have problems in fully support people's change in a long-run. Thus one important work we did in the thesis is to identify problems and gaps in current application design through digging into people's thoughts and beliefs. And further to understand how those thoughts and beliefs are different from current design so we can leverage them to design better interventions.

- What are people's thoughts and beliefs about food and what kinds of categories/-topics emerge from those thoughts about food?

  To leverage people's thought and belief about food, we need to investigate what exactly people think and believe about food so we have the knowledge to guide our work on collecting and using such information. The categories or topics emerge from those thoughts will help us to narrow focus and possibly design application which can collect those information in a relatively structured way.

- What kinds of user interface can we develop to capture those thoughts about food?

  To leverage thought and belief, we need to have interface that can capture it in first place. However, we are not just going to offer survey to collect such inforamtion but the data collection process will be part of an application. However, the scenario of the application may in turn constrain the approach we take to collect data. That means, we need to figure out the trade-off between richness of information and time if it's an application offering real-time service. Strategies on how to balance different requirements of the UI is needed to be explore.

- How can thought about food be visualized and what is it's effect on influencing people's understanding of food?

We also need to understand that once people's though and belief about food is collected, how can we visualize them so we can help people make better choice. We also concern whether such visualization might help improve their knowledge so they can make better decision not just in one particular case but also in the long run. That requires understanding how such visualization might change people's mind thus indicates its effectiveness on improving understanding of food.

1.2 Research Approach

In this thesis, we take an exploratory rather than hypothesis-driven approach to carry out our investigation. The specific research approach evolved over time as we understand much better about what people think about food. To investigate the alternative notion of leveraging people's thoughts about food, we first needed to understand how people actually use basic food logging solution to help them eat healthy and what do they actually need to fully support their practice so we can identify a gap between what support really is needed by people and what support is offered by current interventions and explore how people's thoughts might be leveraged to fill the gap. We thus investigated two online healthy eating forums, which is a natural place for people to ask questions and express own thoughts, to understand what type of topics on dietary change people discuss online and how they support each other on different topics. The informative answers we got from the qualitative analysis of online forum posts guided us to narrow our focus to people's food beliefs, which means how people understand and assess food's healthiness.

Then we designed a simple food-logging application to help people express their assessment of food's healthiness and use these information as feedback to support each others' dietary change. In the pilot study we conducted, people offered positive feedback on expressing their own beliefs and reflecting upon own thinking based on others' feedbacks. Coincidently, there were another two food logging applications using the similar approach to capture people's assessment and use it as feedback to support dietary change. Thus, we decided to focus on exploring interaction behind such applications rather than developing and studying such application itself.

As reported in Chapter 5, our first attempt was to run a survey study to understand how people construct their assessment of food's healthiness so we can develop better UI to capture such information. This led us to understand what type of attributes of food people take into account when assessing food's healthiness and based upon this understanding, we further studied how can we design efficient user interfaces to capture those thoughts in a relatively structured way to support future re-use of those knowledge. The series of quiz studies we carried out compared three interfaces we proposed from

three different perspectives (more details in Chapter 6) and led us to think about how we might use those knowledge we capture to support future dietary change interventions.

1.3 Research Outcomes

This thesis covers following outcomes of research works we have done:

- Identify the gap between what supports people need for dietary change and what supports dietary change applications offer through literature review and two studies. Our online forum study indicated that current applications fail to address context information that includes physical context and people's belief but also emotion information that reveals how people feel about dietary change and food. We further found that the usual calorie counting method employed by most applications is flawed and that indicates we need to leverage people's belief about food to offer multi-dimensional, context-dependent healthy eating meaning in application design
- Proposed, designed and implemented a new food logging application that facilitates people expressing what they think about the specific food's healthiness and uses such judgements as feedback to help people reflect upon own thoughts and support each other.
- Studied how people assess food's healthiness and Identified 8 categories food-related attributes people take into account when assessing food's healthiness. That study and its follow-up study with experts population indicate that the healthiness of food has no single, correct answer and it is always "it depends".
- Designed, implemented and evaluated three user interfaces to capture people's beliefs about food. Based upon the results, we proposed the strategies on how to combine three user interfaces to better capture food knowledge in different settings and also proposed strategies on how to visualize food knowledge that we captured by three different user interfaces.
- Designed, implemented and evaluated the visualizations for those data captured by three user interfaces

1.4 Research Publications

Results of Chapter 3's initial study on diet-related forums was published at Pervasive Health 2011 conference:

Feng Gao; Schraefel, M., "Social interaction around diet applications: An initial study", 2011 5th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth), pp.262,265, 23-26 May 2011

Results of Chapter 5's work on understanding how people assess food's healthiness was published at Ubicomp 2012 conference:

Feng Gao, Enrico Costanza, and M. C. Schraefel. 2012. "Honey=sugar" means unhealthy: investigating how people apply knowledge to rate food's healthiness. In Proceedings of the 2012 ACM Conference on Ubiquitous Computing (UbiComp 2012). ACM, New York, NY, USA, 71-80.

1.5 Thesis Walkthrough

In Chapter 2, we will first review the wellness application challenges in Human Computer Interaction to lay out the challenges and questions we will address in this thesis. Then we will review problems and issues around healthy eating to provide the research background. Then three level behavior change theories are reviewed to offer readers the basic theoretical knowledge about behavior change. Finally, previous behavior change technologies are reviewed to bring about the discussion on the potential problems in persuasion-based systems and construct the argument that we need to leverage people's thoughts and beliefs to achieve sustained behavior change.

The next four chapters report the three studies we have conducted and one prototype we developed. In Chapter 3, we reported our efforts on understanding what kind of gaps exist between supports offered by existing interventions and supports people need so we can further explore how people's thought and belief about food might be leveraged to fill the gap. Our approach to investigate the question is through studying how and what people talk about food on two online forums. The forums we selected are both communities for food logging application users. We deliberately selected this type of forum with the expectation that those people's conversations around food and dietary change can reflect what kinds of support is missing in current interventions and how people's thoughts and beliefs might be leveraged to offer those supports. The study led us to focus on the topic of food's healthiness and then in Chapter 4 we present our first application prototype and what we learned from the user study.

Chapter 5 presents the result of a survey study to investigate how people evaluate food's healthiness individually. The analysis of the survey responses concludes that people have the ability to evaluate food's healthiness but not based on calorie counting. Attributes of healthy food that are identified in survey responses are reported and mapped into design implications in this work.

Based upon result of this study, in Chapter 6, we designed and evaluate three simple user interfaces to capture such knowledge about food's healthiness from users. The evaluation of the interfaces were conducted by looking at three different aspects: efficiency, data quality and effect of visualizations. We report the result of the study and implications of results in Chapter 6.

Finally, the conclusion of this thesis and directions for future works are discussed in Chapter 7.

Chapter 2

Backgrounds and Related Works

The work in this thesis lies at the intersection of several academic disciplines: computer science, social science, and health/wellness research. The purpose of this chapter is to review related backgrounds and theories to construct the argument that behavior change technology should be beyond persuasion and explore the alternative notions: leveraging what people think and feel about healthy behavior to support behavior change. That is divided into four parts: First, we will overview the Human Computer Interaction challenges and research agenda in the area of wellness. Second, since our research focuses on the dietary change, we will review related theories about dietary change and how people choose food in general to explain why it is so hard to achieve dietary change and how it is different from other behavior changes such as increasing physical activities. Third, we will review general theories and methodologies on how to change behavior and based upon theories, we will review previous technology-based interventions to reflect upon what succeed and what failed. Finally, we will go over those three alternative design notions for design of behavior change technology: slowness, mindfulness and negotiation. Then related theories on how minds could drive actions to change will be explored and related theories on reflection will be discussed to ground our final argument on how can we leverage people's mind to support behavior change.

2.1 Wellness Application and Its Challenges in Human Computer Interaction

In this thesis, our work looks at how we can better help people change their behaivor to stay healthy or become more healthy, and we are particularly focused on the eating behavior. It is, therefore, important to first clarify that our work falls into the area of wellness but not health. The difference between wellness and health is rather complicated to explain in the thesis but in general, it is accepted that wellness is about

promoting or maintaining good health but health is more about curing illness and correcting poor health (Mackey, 2000; Greenberg, 1985). In this sense, wellness applications are different from health applications. The latter one more focuses on how to deal with the relationship and communication between patients and doctors (Wilcox et al.), how to manage related information of patients (Edwards et al.), and how to better deliver treatment or management of disease (Lee et al., 2012). Wellness applications, on the other hand, aim to monitor people's health state and offer feedbacks on how to maintain it (Mamykina et al., 2008), or assist people in changing current behaviors and adopting new lifestyles so they can avoid medical treatments (Consolvo et al.), or informing people how to avoid threats (e.g. unsafe foods) to their own health (Choi, Jaz Hee-jeong, Foth, Marcus, & Hearn, 2013).

In both business and academic, we already seen the explosion of wellness applications. Various new technologies, especially new type of sensors, are employed to create better tools to monitor people's behavoir (Consolvo et al.), to deliver interventions through new channels (e.g game console Wii) (Klasnja et al., 2011), to engage end-users in keeping following the interventions . The rapid growth of the field raises new questions and challenges in various of field including Human Computer Interaction, health, psychology and so on. In Human Computer Interaction, we could further break down challenges and questions into four categories: behavoir theory, wellness intervention practices, technology, and design (Lee et al., 2012).

Behavior Theories: As we will discuss later, one popular trend in computing is to leverage existing behavior change models to develop appropriate persuasion strategy to help people make change and become much healthier than they currently are. The Transtheoretical Model of Behavior Change (Prochaska et al., 1995) is the most cited model and often it is believed that such model can guide the design of wellness interventions and applications by identifying right target group and coming up with corresponding strategies to deliver the intervention. It is, however, less clear that whether other behavior theories could also be employed and provide different perspectives at behavior change. It is also a challenge that how technology can implement such model of behavoir change in the system so it can adjust its strategies and interventions to reflect user's progress.

Wellness Interventions Practice: Wellness is new to HCI but itself has been studied in health science for a long time. The various interventions created to promote healthy eating or encourage physical activity are well discussed and studied without advanced technologies before and now one challenge is how can we, as HCI practitioners and researchers, reflect upon previous practice and integrate those lessons into current technology-based intervention design. HCI researchers not only need to look at what works in the previous intervention design but also reflect upon why certain intervention does not work. More importantly, HCI researchers should note where is the gap between theories and intervention practice and find a way to close such gaps. We will discuss

more about this in part three on reflecting previous interventions in behavior change technologies and report how we find the gap and what approach we take in this thesis work to address and potentially solve it.

Technology: With the rapid development of sensor technology, we are now able to detect people's motion and collect reliable data. And thanks for the increasing availability of smart phones and wireless network access, more and more people can utilize such sensor-based solution to monitor own behavior and ask for personalized feedbacks on how to do proper training to make self healthy. However, there are still many behaviors can not be monitored by sensors. For instance, we still have no sensor that can automatically log what we eat and that's a big problem in monitoring healthy eating behavior. Data collection is just one challenge in technology development. We also need advanced technology to make sense those data. For instance, by collecting an individual's physical activity data along with related contexts such as location, what kind of algorithm we need to develop to mash-up all those data to detect pattern of the person's physical activity and thus provide actionable suggestions on how can the individual increase physical activity. Various presentation technologies such as visualizations, avatars should also be further developed and carefully studied to understand how data and suggestions can be delivered to the end-user efficiently.

Design: Given the fact that each individual is unique, one intervention can not work for everyone. Therefore, it is necessary to understand "how physical, emotional and cognitive abilities, caused by individual learning histories and health states, may impact the usage and acceptance of these systems"(Lee et al., 2012). That also imposes another important challenge: how can we engage people in keeping using the technology so they can continuously make progress on changing behavior or maintain the behavor and avoid relapse. Both of challenges encourage researchers to better understand end-users in a particular context and take into account related behavior theories to think about how to better choose appropriate technology to actually address such complex in behaviour change.

In this thesis, we will thus reflect upon current practices of healthy eating applications and take into account what we learn in behavior theories as well as nutrition science to think about how to better address people's uniqueness by understanding and leveraging people's thoughts about food to potentially better engage them in adopting to new interventions and make changes sustained for life-long. We will demonstrate our reflection upon current practices and what we conclude throughout the rest of this Chapter.

Figure 2.1: EatWell Plate (Food Standards Agency, 2010) developed by the Food Standards Agency, UK

2.2 The Healthy Eating Problems and Health Eating Guidelines

With the development of modern medicine, number of deaths from infectious diseases, maternal and perinatal conditions, and nutritional deficiencies has been reduced. However, more and more people die because of chronic diseases now. It is predicted by World Health Organization that in 2015, 64 million people will die due to various chronic diseases: such as heart diseases, diabetes, obesity and cancer. Three health-related behaviors are reported to be linked to chronic disease, they are: unhealthy eating, physical inactivity and tobacco use. According to the World Health Organization (World Health Organization, 2005): Five out of the 10 leading global disease burden risk factors - high blood pressure, high cholesterol, obesity, physical inactivity and insufficient consumption of fruits and vegetables - are strongly related to diet and physical activity.

Thus, in response to the urgent problems in health, efforts have been made to promote positive change. One particular area is to promote dietary change to solve the unhealthy eating problem. In UK, the Food Standards Agency developed a healthy eating guideline, which is called EatWell plate (Food Standards Agency, 2010), to help people know what is a good meal. As Figure 2.1 shows, the EatWell plate focuses on five food groups:fruits and vegetables, dairy, starchy foods, high protein foods, and high fat/-sugar drinks and foods. The guideline was criticized that it misleads people to believe they should avoid fats and focus on carbs while the scientific results prove fats are not completely bad (Willett, 2002) and too much carbs is not good (Kopp, 2003). Unlike the UK's EatWell plate, the recent updated guideline from USDA seems to provide a

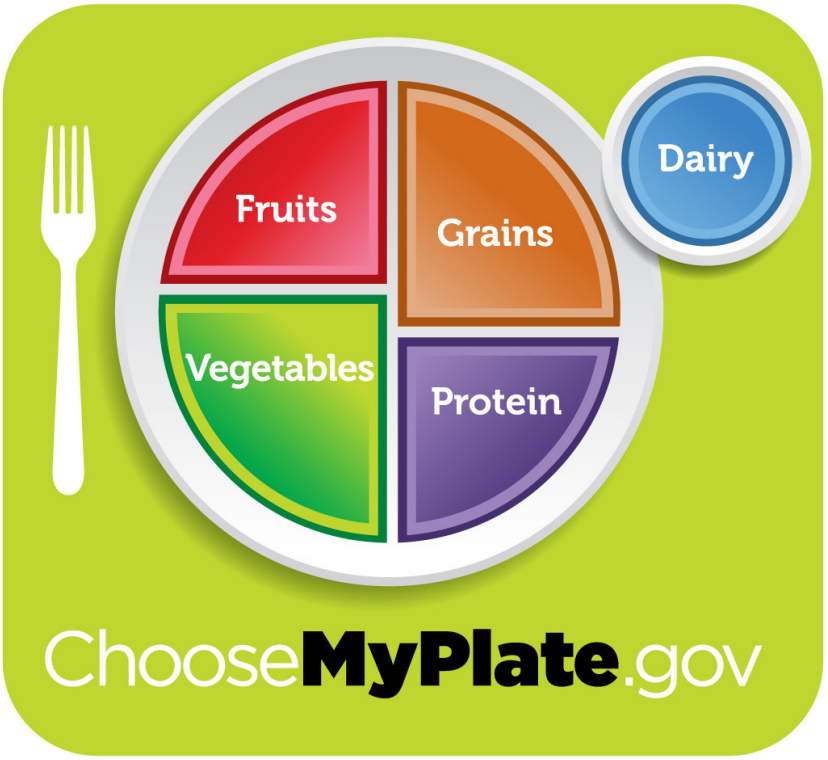

Figure 2.2: MyPlate (ChooseMyPlate.gov, 2011) developed by USDA

much more reasonable dietary plan. The new guideline, MyPlate (ChooseMyPlate.gov, 2011) (see Figure 2.2), also focuses on five food group, but it does separate fruits and vegetables and does not contain high fat foods. However, this diet plan also has been challenged.

Compared with the USDA's MyPlate, the new Harvard Healthy Eating Plate (Harvard School of Public Health, 2011) (see Figure 2.3) has several advantages: First, it clearly states that people should choose whole grains instead of just grains. Second, it also introduces the term "healthy protein" into the diet plan to clarify that not all protein are good for health but only those lean proteins from not heavily processed ingredients such as chicken and salmon could be counted. Therefore, protein from McDonald's hamburgers should not be counted as healthy protein. Third, they recommend people to get water instead of dairy for every meal. Finally, they have one more thing in the plate, the healthy oils. Considering the fact that bad fats are included in the EatWell Plate and MyPlate does not include any fats, It is interesting to see that Harvard Healthy Eating Plate introduces the healthy oils into their plan. Fat, as one essential nutrient, should not be ignored in healthy eating guideline and the right direction is to help people know where to get healthy fat. The Harvard's Healthy Eating Plate, therefore, does a relative good job on this point but it still fails to help people get more healthy fat options.

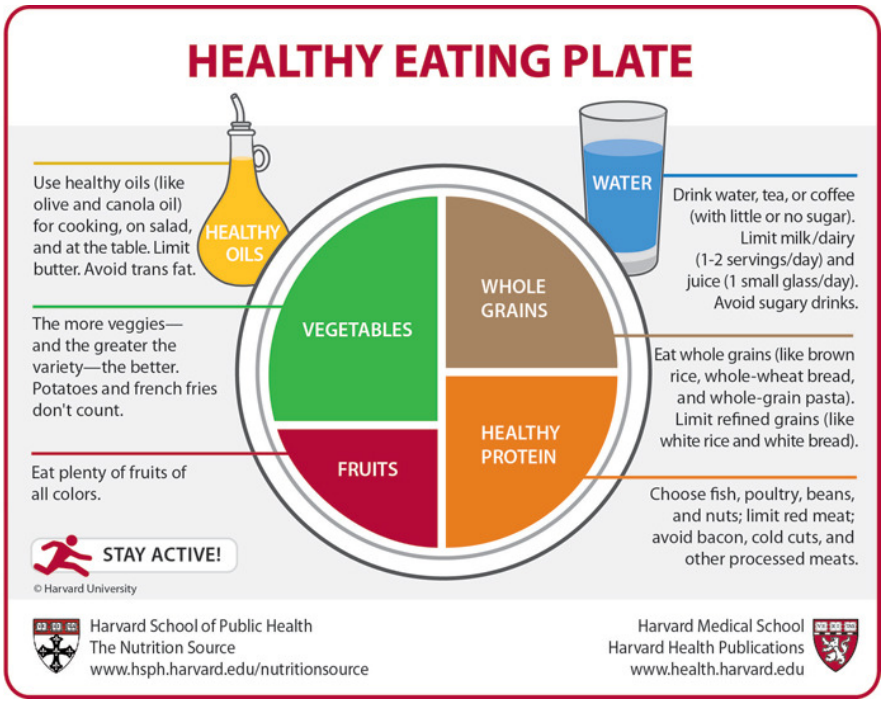

Figure 2.3: The Harvard Healthy Eating Plate (Harvard School of Public Health, 2011)

Besides above diet plans, UK government also promotes a scheme called 5 A Day (NHS Choices, 2011). This scheme aims to increase the public's consumption of fruits and vegetables. It suggests that everyone should get 5 portion of fruits or vegetables per day. However, it is not very clear that what type of fruits or vegetables food item could be counted in 5 A Day and what is the correct portion size. In one internet survey, a majority of the 1,347 participants wrongly believe that a portion of fruit and vegetables could be orange squash, herbs or chips (The Independent, 2008). Therefore, such guidelines are still too vague for general public to understand and really practice well in their day-to-day life.

2.3 The Complexity of Everyday Eating

Dietary change is never easy to be achieved. In one survey conducted in Scotland, only 30.1% of respondents (n=6282) were considered to have a healthy diet (NHS Health Scotland, 2000). There are probably two reasons behind the failure of dietary change: one reason is today's dietary guidelines are quite vague and even conflict each other; the other is that our everyday eating is quite complex. In this section, we do not fully review all theories and models on the complexity of our everyday eating. Instead, we briefly review one major model of food determinants to illustrate how complex our everyday eating could be.

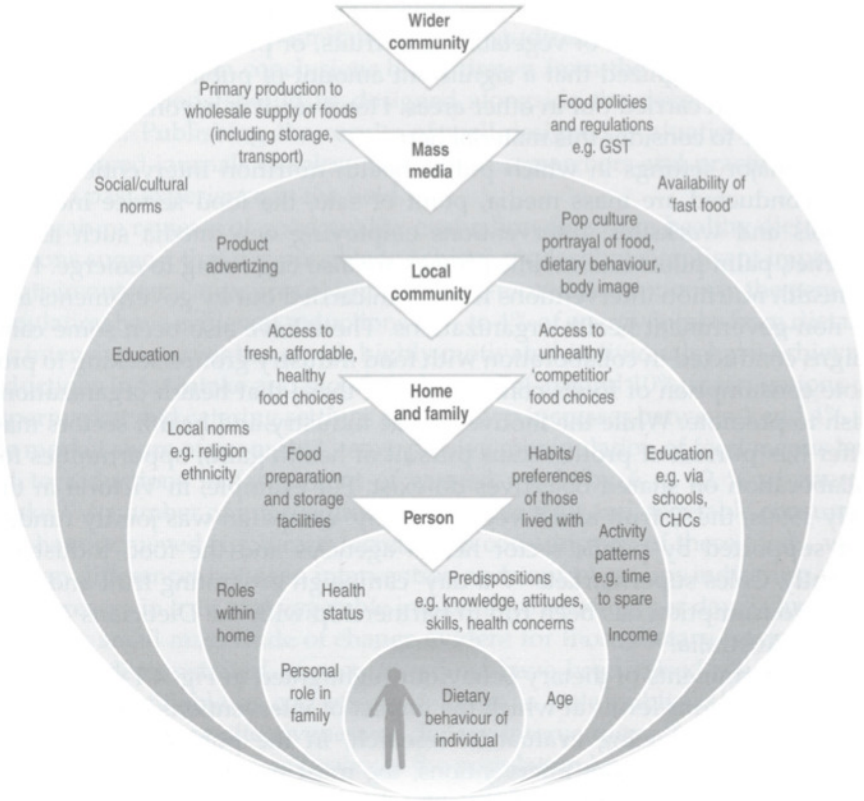

Figure 2.4: The Conceptual Model of the Determinants of Individual Dietary Behavior (Dixon and Paxton, 2005)

The Figure 2.4 shows the conceptual model of the determinants of individual dietary behavior developed by Dixon and Paxton (Dixon and Paxton, 2005). In this conceptual model, five levels determinants of individual everyday food choices are described: person, home and family, local community, mass media and wider community.

Person: At the personal level, individual's beliefs and attitudes to healthy eating is one important determinants (Paquette, 2005). That determines how people will handle different situations in which they need to make decisions for what food they choose and how they will eat them . In addition, individual's own skill and knowledge also influence what the individual will eat. For example, a person who has no experience in cooking probably can not cook one healthy recipe recommended by a friend and has no choice but to choose fast or instant food for himself/herself. More common, personal cultural background and health status also set different constraints in food choice.

Home and Family: Family members' health status and food preference could be constraints to a individual's food choice. For example, if a family has a child who is allergic to gluten, then the family will try to avoid whole wheat product or limit the amount. The family's food choice could also be constrained by the family's daily routine that set time constraints on food choice. Furthermore, the resources a family has like money and space might influence what food the family could buy.

Local Community: One determinant at this level is the social norm. if the social norms at a particular community could form certain discussion, reflection or participation surrounding particular diet behavior, then the social norm could be one positive resource for behavior change. Another important determinant is the availability of food. If a local community does not have certain type of food or does not have convenient market for grocery shopping, then it will be constraints for dietary change.

Mass media: Mass media like television advertisement of food product could be either constraints or resources. If healthy diet messages could receive similar television coverage as fast food products do then people will be motivated to eat more healthy food. Other determinants at this level include education program and pop culture portrayal of food.

Wider Community: There is no doubt that food policies and regulations could influence one's food choice. For example, high tax for certain food type will make the price of this type of food unaffordable by low-income family. In addition, the transportation and production of food is another important determinant at this level. Similar with local community level, the social norm could also influence individual's choice.

2.4 People's Interpretation of Healthy Eating

As we mentioned in last section, a person's beliefs and attitudes to healthy eating could play important role in determining what people eat (Paquette, 2005). Research has shown that people generally categorize food into healthy or unhealthy based upon particular, recurring attributes (Oakes and Slotterback, 2001; Carels et al., 2006). In previous research, large-scale questionnaire and in-depth interviews were conducted to capture those attributes people concern. The literature reports attributes including macro nutrients (protein, fat, carbs), types of food (vegetables, fruit), perceived food quality (fresh, processed) and portion size. For example, some studies (Edwards and Hartwell, 2002; Santich, 1994) highlight that people tend to consider low amounts of fat, sugar, and salt as good for health. Studies involving children, adolescents and adults (De Almeida et al., 2001; Eikenberry and Smith, 2004) reported that vegetables and fruits were most often mentioned as healthy food. Freshness, as opposed to frozen, canned and processed food, tends to be perceived as healthy or healthier (Povey et al., 1998; Sage Research Corporation, 2003). On a similar note, another study (Santich, 1994) indicates that home-made food is considered the most healthy. The concepts of balance and variety are also attributes found in the literature (Falk et al., 2001; Croll et al., 2001; Sage Research Corporation, 2003). However, there also seems to be a general confusion about what a "balanced diet" or meal actually is (Edwards and Hartwell, 2002; Keane and Willetts, 1995). Related to balance is proportion: we generally make poor judgment when it comes to either describe or identify what a good serving size is (Brindal et al., 2012; Huizinga et al., 2009).

Stage of Change	Definition
Precontemplation	the person has no intention to change in next six months
Contemplation	the person feel uncomfortable to current state and think to change in next six months
Preparation	the person start planning change
Action	the person start changing behavior
Maintenance	more than 6 months after changing

Table 2.1: Five Stages of Transtheoretical Model (Prochaska et al., 1995)

2.5 Health-Related Behavior Change Theories

As we discuss before, our everyday eating is so complex but the healthy eating guidelines is not quite clear and strong to provide people with scaffolds to start their dietary change. However, questions such as why people want to change, how they could be motivated to change, and what they need to facilitate the change are still unanswered in this thesis until now to provide readers with solid theoretical foundation based upon that designers, researchers and developers could build technology to influence and support behavior change. In this section, we review major theories and models in health and nutrition fields for three level behavior change: individual, interpersonal and community.

2.5.1 Individual Behavior Change

Models of individual behavior change include two type of models: social cognition models and stage models. The social cognition models assume people's behaviors are controlled by their own mind. People determine whether change certain behavior based on their perceptions of benefits and costs of health action. On the other hand, the stage models try to categorize people into different stage and assume people will move between different stage to achieve changes. In this section, we will discuss two models in the category of social cognition models: Health Belief model and Theory of Planned behavior. And we will also discuss one stage model: Transtheoretical Model.

2.5.1.1 Transtheoretical Model

Transtheoretical Model(TTM) (Prochaska et al., 1995) is the most widely accepted and adopted model of Health-Related Behavior change. Based on individual's intention to change, the model defines 5 stages: Precontemplation, Contemplation, Preparation, Action and Maintenance. The definition of these five stages are shown in Table 2.1

In addition to the stages of change, the TTM also provides a set of processes of change as tools that could guide and support behavior change throughout different stage of change. The basic idea about how to apply those processes of change is that for each

stage, people who want to change need appropriate but different tools to achieve their goals. According to the TTM, dramatic relief, consciousness raising, self-re-evaluation, environmental re-evaluation, and social liberation are most effective for those in precontemplation and contemplation. While, for those in preparation, action or maintenance, helping relationships, self-liberation, counter-conditioning, stimulus control, and reinforcement management are most effective.

In terms of the application areas of TTM, it was originally developed to understand how people change their behavior to get rid of addictive behaviors like smoking. But until now, TTM has been adopted to develop interventions for other behavior change such as dietary change (Spencer et al., 2007) and physical activity change (Hutchison et al., 2009). It has been proved that the Transtheoretical model could work beyond its original setting. However, when applying the TTM into intervention design, it is not quite clear what it is the best practice to apply the model into a specific behavior change intervention. One problem is that the definition of each stage is not quite clear. In one study conducted by Ni Mhurchu et al, they pointed out that based on TTM individuals could be categorized into multiple stages (Ni Mhurchu et al., 1997). Additionally, the time cutpoints (1 and 6 months) based upon the definition of stage might conflict how individuals plan their own change (Herzog et al., 1999). Therefore, it raises a interesting question that how intervention could correspond to the individual's own thinking to provide essential support.

Additionally, (Adams and White, 2005) pointed out that the contributing factors to behavior change such as gender, income, and lifestyle are ignored by the TTM and this is potentially a flaw because without taking into account those factors, certain behavior change of a particular population can not be explained well by TTM. For example, Resnicow (Resnicow, 2003) conducted a fruit and vegetables consumption change intervention study that recruited non-white, low-income populations. In their findings, they suggested that the interventions themselves and the algorithms used to assess stage of change may not apply equally to all socio-economic and cultural groups.

2.5.1.2 Health Belief Model and Theory of Planned Behavior

The Health Belief Model believes that the individual's perception of the seriousness of a health problem will influence the individual's attitude and intention to health behavior change (Rosenstock et al., 1988). The perception of the seriousness of a health problem might be influenced by the individual's perception of costs and benefits associated with taking actions or it possibly is influenced by personal characteristics such as demographic and personality. While this model provides an explanation for why people do or do not engage in one health actions, it is criticized that it ignores social factors and emotional influence (Browning and Thomas, 2005).

The Theory of Planned Behavior believes that intentions precede actions (Ajzen, 2002). In their theory, the intention could be influenced by the individual's expectation toward what would happen and what would be achieved. It is also mentioned that the intention could be influenced by what other people might think about the outcome of behavior. The Theory of Planned Behavior also take into account the perception of control on resources, opportunities and skills. Interventions based on this model usually try to help people increase their perception of control and change their beliefs. But the model also has been criticized , similar to other two models we reviewed above, that the broader social factors and emotional influence are ignored. Moreover, this theory also is criticized that it focuses on perceived control rather than actual control (Browning and Thomas, 2005).

2.5.1.3 Summary

To summarize, above Individual models focuses on the individual's attitude, ability, and belief that might help the individual achieve the change. Individual models try to convince the individual that he or she has all personal resources to achieve certain change without considering any external contains or resources. In contrast, as we will discuss below, interpersonal model and community model pay more attention to external influences on behavior change.

2.5.2 Interpersonal

Based upon individual level models and theories, the interpersonal level models and theories take environmental and social resources and constraints into account. Especially, interpersonal level models and theories focus on how to utilize the social network to provide people with sufficient supports, either practical or emotional.

2.5.2.1 Social Cognitive Theory

The social cognitive theory (SCT) is a framework to explain the way people perform and maintain changes (Bandura, 1997). Similar with social cognition models described above, the core idea of SCT is "what people think, believe, and feel affects how they behave" (Bandura, 1985). Therefore five cognitive related attributes are highlighted in SCT (see Figure 2.5)

Moreover, unlike individual models, the SCT take the environmental factors into account. As illustrated in Figure 2.6, environmental factors, personal factors and behavior interact and influence each other.Neither the environmental factor alone nor the personal factors alone could predict the outcome of behavior (Glanz et al., 2008). The

Attribute	Description
Symbolisation	Symbols are viewed as mechanisms for thought, and help individuals to remember the consequences of previous actions in order to speculate the results of future behavioural change.
Vicarious Learning	The ability to learn through observation. Observational learning processes include determining the relevance of other's behaviour in relation to .
Forethought	Generating an expected outcome from an as yet unperformed behavioural change.
Self-Regulation	Determines any behavioural change that occurs and any subsequent self-imposed consequences. It relies on internal and external factors such as self-motivation and social and moral standards.
Self-Reflection	When an individual analyses his or her own thought processes, actions, and experienced outcomes, and learns from them.

Figure 2.5: Five Cognitive Attributes Highlighted in Social Cognitive Theory (Maitland, 2009)

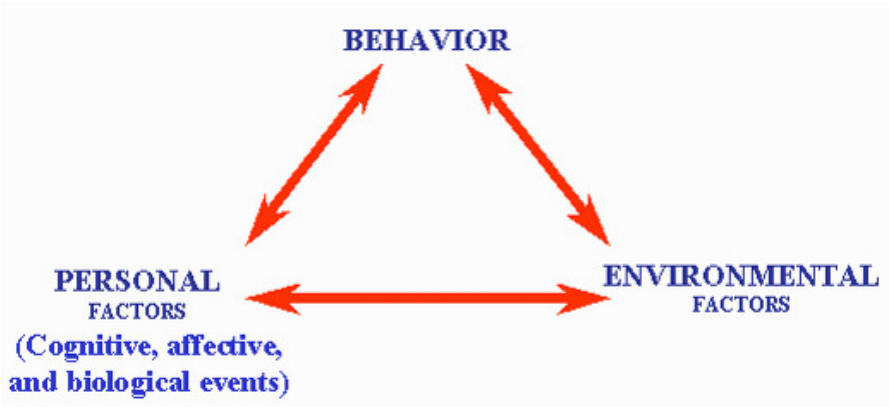

Figure 2.6: The Social Cognitive Conceptual Model. From http://www.emory.edu/EDUCATION/mfp/eff.html

environmental factors here might be physical environmental factors like the availability of certain foods or social environmental factors such as friends, family members and colleagues. In addition, how people perceive the environment (defined as situation) also influences people's behavior. Environment and situation together provide the framework for understanding behavior (Parraga, 1990).

2.5.2.2 Social Support Theory

Social support describes the process that a person changes own behavior with the supports from other people. It has been widely adopted in behavior change intervention and it has demonstrated positive effect on health outcome. However, the underlying mechanisms of social support still remain unclear. In Chronister et al's (Chronister et al., 2006) research, they explain social support from three dimensions: structural, functional and perceptual.

The structural dimension focuses on the quantity of a person's connection with his or her personal network and also the characteristics of the network (Chronister et al., 2006).

This dimension is the easiest one to be measured. The size of network and the frequency of contacts are two common measurements and social networking structure analysis is usually conducted to better understand the structural dimension. The focus of research on structural dimension is on quantity, but critiques point out that quantity does not equate to quality and the influence of quantity is limited. For example, a large network of friends, who do not do any physical activity and love high fat food, will not help a person eat healthier. Therefore, it is not surprising that quantity of a persons' social network connections is the weakest dimension that is correlated with health (Wellman, 1992; Porritt, 1979).

In contrast, the functional dimension focuses on quality of the social support. The functional dimension could be divided into four subtypes: emotional, instrumental, informational and appraisal (Chronister et al., 2006). Examples of functional support include: listening to emotional expression and giving encouragement, providing financial aids, offering advice and giving feedback. It is hard to determine the quality of different functional support and it is unclear that which subtype of functional support is the most efficient. However, it has been widely accepted that the functional support is the most effective indicator of health outcomes (Porritt, 1979).

How people feel about a network is the perceptual dimension of social support (Chronister et al., 2006). For instance, whether a person feels that he or she "belongs" to the network or whether a person feel the support from network is strong or weak are issues perceptual dimension research cares. The perceptual support is strongly correlated with psychological well being (Chronister et al., 2006).

2.5.2.3 Social Influence Theory

The social influence is defined as the process that a individual makes real changes through the interaction with other people. There are four prominent theories about social influence: social facilitation, social comparison, conformity, and social learning. Each of them are briefed below and how they are applied into technology design will be elaborated later:

The social facilitation (Zajonc, 1965) suggests that people could do better in a long run if other people keep being present, participating and observing.The representation of other people could be real or virtual. For example, imaging there is a simple avatar as your buddy in a running game. It always compete with you as "other people", and the presence of this simple avatar potentially could motivate you to keep running since it is fun to compete with another person.

The social comparison (Festinger, 1954) points out that people could do better if they are informed how their performance compares with other people's performance. The theory assumes that people want to use other people's performance as a reality check point to

convince themselves that they are doing good or bad. The effect of social comparison could be strengthened when the performance comparison is conducted between similar people. The similarity between people could be gender, age, culture, health status, health goal and other possible attributes.

Unlike social comparison which uses peer's information to motivate people and influence the decision-making process, the conformity (Turner, 1991) theory proposes to leverage peer pressure to push people adopt or avoid a target behavior. Research has shown that people tend to change their attitudes or behaviors to match the group's expectation, which is set by the performance of the majority (Moe, 2002). Sharing information about one individual's progress with other people might be the simplest way to apply this theory into design.

The last theory is social learning (Bandura, 1985). It suggests that a person could be motivated to do a target behavior if he or she could observe what others do and how they are rewarded. The power of modeling is therefore valued in this theory and it suggests that intervention design could consider highlighting some top performance people in a community as the model to the whole community to motivate people observe what they do and learn how they do.

2.5.3 Community-based

Compared with above two models, the community-based model is less frequently discussed. But just as Davis et al (Davis et al., 2005) suggested, there are so many barriers and constraints in our social, cultural and physical environment and it needs to leverage the power of community to bring about change.The general focus of community-based intervention for health promotion is on changing systems, rules, social norms, or laws in order ultimately to change the social acceptability of certain behaviors. The community-based model therefore is developed to address issues related to those social and environmental constraints and resources. Those issues could be raised by the community and solved by the community.Five key concepts of the community-organization model therefore could be introduced: community empowerment, critical consciousness, community capacity, participation and relevance, and issue selection (Byass, 2006).

The five concepts describe the process of how people in a community could work together to achieve their change. First of all, the critical consciousness, and issue selection point out that people in a community could work together to find out what are the problems they feel most strongly about and what are the causes of those problems in their environment. Then the community empowerment suggests people could apply their skills and use the resources they own to solve those problems they identified. In the context of behavior change, it means people need to work together on changing their

social and physical environment, in which they could apply their skills and utilize available resources, in order to achieve their own changes. Finally, the community capacity concerns whether people have the confidence and skills to work together effectively and could proactively becoming involved in the whole process.

2.6 Behavior Change Technology

Thus far, we have discussed background and theoretical foundations for behavior change. In this section, we shift the topic from theory to practice to discuss what technology could be used to support behavior change and what have been done. However, before we dive into the review of systems, we need to clarify what term we might use to refer to the technology to support behavior change. In HCI literature, when researchers mention the systems designed to support behavior change, usually the term "Persuasive Technology" are used to name those systems. However, as we briefly discussed in Chapter 1 that we favour the alternative notions to persuasion in this thesis, using the term of persuasive technology to refer any technology for behavior change might be confusing. Therefore, in this thesis we will use the term "behavior change technology" to refer to any interactive computing systems designed to change behavior. And we will use the term "persuasion-based system/technology" to refer to systems inspired by the notion of persuasion that is proposed by (Fogg, 2002).

In this section, we review previous technologies based upon four major behavior change intervention features, which are summarized in the systematic review of 11 interventions to prevent weight gain by (Hardeman et al., 2000) :self-monitoring, goal setting, information provision, and social support and social influence. These features are usually used not alone in one intervention, but as a combination of several features. The review of how previous technologies implement each features will provide insights into what problems still exist and what type of improvements future systems might make.

2.6.1 Self-Monitoring

Self-monitoring is one commonly used feature in behavior change technology, whose basic idea is to provide people with the behavioral data to look back and reflect on own performance of certain behavior. Self-monitoring is achieved because of the rapid growth of the Ubiquitous Computing, which is a notion first proposed by Mark Weiser in his visionary article about invisible computing (Weiser, 1991):

> The most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it

Thanks to rapid development of sensor technology, this vision has been partially fulfilled. And in recent years, sensor technology has been applied into health domain to monitor different data about human. For example, one common sensor people use now is pedometer to count steps . Beyond simply recording data, self-monitoring also could do inference based upon data recorded. It is possible to make inference on what type of activity (e.g. cycling or running) based upon the input of sensors (Froehlich et al., 2009). However, unlike monitoring physical activity, there is less mature technology available to automatically capture diet-related data. There are currently three common ways to capture diet-related data. The first one is to let people manually type into what people eat, and match the food with database to get calorie information. The flaw of this method is it requires too much effort and it is really time-consuming. The second one is using picture of food and utilize either algorithms to automatically determine what is the food or human computation to let group of people recognize what is the food (Noronha et al., 2011). However, neither of them now is fully reliable. The third way is to scan barcode on the food package (Fooducate [1] as an example). This is very easy for user to do but usually it confuse users because most databases of such systems only include information about major brands food products and also possibly lack accurate data. None of these three methods now could completely eliminate the tedium of monitoring behavior.

The limitation of the self-monitoring technology in the context of diet, therefore, is a big challenge in design for dietary change related technology. Particularly, this limitation affects how people perceive the barrier to starting using the technology. As we will discuss in Chapter 4, our participants felt that it was hard to keep using the photo-based food logging application because the self-monitoring requires them to deliberately take photos of food before every meal. This deliberate effort of self-monitoring might be against the notion of invisible as proposed by Weiser (1991), which implies the self-monitoring could occur without attracting people's attention and requiring their efforts. Interestingly, the ultimate goal – completely automation of the self-monitoring – is also criticized. As pointed out by (Purpura et al., 2011), the completely automation of the self-monitoring might take away the opportunity to proactively think and reflect on how well the individual did. In Chapter 4, while participants thought it was hard to keep logging food, they still appreciated the opportunity to proactivly think about their own eating behavior by taking pictures and rating food. Therefore, the tradeoff between automated self-monitoring and human-involved self-monitoring should be further studied.

Before we move on to the next feature, we need to firstly clarify what type of systems will be used as examples in following sections. Considering the fact that most behavior change technologies employ self-monitoring and the fact that self-monitoring on diet is not quite mature as self-monitoring on physical activity is, we will mainly use systems

[1] http://www.fooducate.com/

supporting physical activity change as examples to illustrate how they employ different features, but our discussion will still try to focus on how dietary change related technology could employ those features.

2.6.2 Goal Setting

Goal setting is another common feature in behavior change intervention. Usually, goal setting is used along with self-monitoring, which is utilized to monitor whether the user make any progress towards the goal. The first issue related to goal setting is what is the goal setting strategy the behavior change technology should employ. Three sub-questions are asked: (1) Who sets the goal? (goal source) , (2) Whether the goal is static or dynamic (i.e. could be changed)? and (3) How the goal was determined?

In the work of Locke and Latham (Locke and Latham, 2002), three types of goal sources are defined: (1) self-set,the user sets the goal (2) assigned, system assigns the goal to users,and (3) Participatory, the goal is negotiated by system and the user. In addition, two more goal sources are defined by Shilts, Horowitz, and Townsend (Shilts et al., 2004): (4) guided, the user choose one goal from multiple goals determined by systems ,and (5) group-set, a group of people set the shared goal for all people in the same group. In the Table 2.2, we summarize examples of how previous physical activity change systems employ different goal setting strategies to answer our three questions above.

Unfortunately, when we conducted literature search, we did not find examples of guided goals or group-set goals. And just as the Table 2.2 shows, most systems choose to implement either assigned goal or self-set goal. The self-set goal is possibly the easiest one to be implemented. Potentially, using self-set goal could help people fully engage in determining what they believe they could achieve. But handing over all control to users also takes the risk that unrealistic goals, which are very hard to achieve in short-term or even long-term, are set during the process of change. Moreover, the goal for physical activity change usually is a value of step counts (e.g. I want to walk 5000 steps today) while the goal for dietary change could be much complicated, which is from eating five fruits a day (quantitative goal) to eating only whole food (qualitative goal). Thus a challenge in implementation of self-set goal for systems supporting dietary change is how the system could manage, understand and track those complicated dietary change goals set by users. In contrast, the assigned goals allow the system to predetermine a set of goals and corresponding strategies to manage and track those goals. But assigned goals also possibly take the risk that goals are to be too hard or easy for particular users. Perhaps, the best way to set goal is the participatory goal method that allows people to negotiate with the system to set the final goal. However, the details on how to implement negotiation between the user and the system are not clear. It is a big challenge that what strategies the system should employ to negotiate with user to tweak goals. In other words, when the user wants to set either a higher or a lower goal than the one proposed

System	Who sets the goal	Static or Dynamic	How the goal is determined
Chick-Clique (Toscos et al., 2008)	self-set	static	Users are required to set one goal based on their baseline steps
LifeStyle Coaching Applicaiton (Gasser et al., 2006)	assigned	static	App sets generic goal: earn 7 point a day
Houston (Consolvo et al., 2006)	assigned	static	App sets the goal based on user's baseline steps and Walking Works program guideline
Fish 'n' Steps (Lin et al., 2006)	assigned	dynamic	A realistic step-count increase is identified in literature and a new goal will be set every six weeks
Ubifit (Consolvo et al.)	self-set	dynamic	The user sets weekly goal
FitTrack (Bickmore et al., 2005)	Participatory	dynamic	A long-term goal is firstly determined by system based on baseline step. The user is allowed to negotiate the daily goal with virtual agent based on current day's steps, step count history and essential increase

Table 2.2: Example of Goal Setting Strategies in Previous Physical Activity Change Systems

by the system, how does the system know it is reasonable to do? What information should the system get from the user to make sense why the user wants to re-configure the goal and what strategies the system should employ to determine whether it should compromise or further convince user to better not change the goal?

Another issue related to goal setting is how to provide feedbacks on the individual's progress towards the goal. Different technologies, which vary from text messages to simple data visualization to complicated animated characters, are employed to deliver feedbacks. Houston (Consolvo et al., 2006), for instance, uses a simple asterisk annotation to indicate whether the user meets a specific goal on a specific day. In contrast, Ubifit (Consolvo et al.) develops a complicated animated garden to indicate different level of progress: flowers mean any activity people did (small steps), and butterflies mean goals have been met (milestones). As Figure 2.7 shows, the Houston's asterisk is

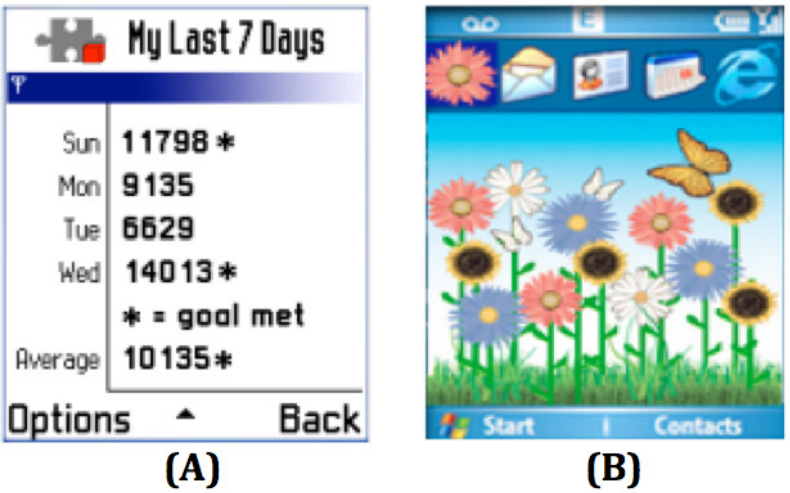

Figure 2.7: (A) is the simple asterisk visualization in Houston, and (B) shows the animated garden in UbiFit

displayed on a simple temporal visualization that could inform users the progress trending information. While the Ubifit's garden is more vivid and attractive, such trending information is missing (even though people could perceive progress trend by looking at what's new element in the garden everyday). However, whether such difference might influence people's performance on behavior change is not clear now.

2.6.3 Information Provision

Providing information based on the monitored behavior is one useful method to help people understand what they could do and assess how they are doing now. The most common information provided in dietary change systems is nutritional information. The common work flow for those systems is that the system first requires user to type into what they eat or scan a product, and then the system could look up corresponding nutritional information for the food item in their database. The limitation of this method is that usually the data in database could be either inaccurate or incomplete. Some applications now allow users to contribute data to solve this problem. For example, one application called TweetWhatYouEat [2] encourages people to contribute missing data or to correct data that they think are incorrect. By utilizing human computation, the data in database could be potentially complete and correct. Alternatively, systems might utilize human computation just in the process of analysis of food to get right information. For example, the system PlateMate (Noronha et al., 2011) assigns a food picture taken by the user to multiple turkers hired on the Mechanical Turk [3]. Turkers are required to analyze the food picture and try to determine what ingredients are in the

[2] http://www.tweetwhatyoueat.com
[3] https://www.mturk.com/mturk/welcome

picture and their quantities. Then the system could automatically calculate the total caloric value for this food by averaging out the totals contributed by each Turker.

Systems also could provide nutritional information for food preparation. (Chi et al., 2008) hypothesized that given nutritional information during cooking, people could be aware of the calorie content of a meal and make necessary change. They developed a smart kitchen to monitor user's each cooking step and provide in-real-time nutritional information about the ingredient user are currently using. In their evaluation, they had three participants who were interested in healthy eating to cook their meals in this smart kitchen. The final result shows that participants' calorie consumption were reduced after using this smart kitchen.

In addition to simply providing nutritional information, some systems go one step further to offer more clear and actionable suggestions based on the nutritional value. In (Mankoff et al., 2002)'s research, they developed a shopping assistant to analyze grocery shopping receipts. The system could analyze the receipts and offer nutritional information alongside suggestion on what people need to change to get a healthier and balanced diet. Another recent released mobile system called Fooducate [4] not only provides detail nutritional information of one food product but also provides information on possible substitutions for the food item and information on what good nutrition in the food will have good health effect on certain type of people.

Alternatively, some systems try to offer more advice beyond just nutrition. MAHI (Mamykina et al., 2008) is such a system developed for patients suffering from diabetes to upload food pictures and own audio/text notes and receive valuable feedback from experts. The motivation behind this work is to create a space for both patients and experts to reflect on what they did and interact with each other to analyze and solve potential problems. Although they did not see significant change in behavior, they did observe that people's locus of control and result of knowledge assessment all were improved after this intervention

Providing information around food and diet might help people change behavior but this feature also has been criticized. One argument is too much information might be distracting because people are not capable to process all those information (Norman and Bobrow, 1975). Another argument is information is not enough because people need real tangible help on changing their diet (Maitland, 2009). For example, people might already know the benefit of eating fruits and vegetables and know where to buy, but they might not have enough money to afford those fruits and vegetables. This is a problem that can not be solved by information and require more tangible help. Finally, as evidenced by (Viscusi, 1990), smokers who access to many information regarding the risk of smoking will overestimate the health risks. The same problem might occur in

[4] http://www.fooducate.com/

System	Wth whom the data is shared?	Competition?	Feedback?
MAHI (Mamykina et al., 2008)	Experts	No	Yes, guided questions and suggestion are given by experts
Chick-Clique (Toscos et al., 2008)	Friends	No	Yes, teammate could send encouragement messages to each other
LifeStyle Coaching Applicaiton (Gasser et al., 2006)	Strangers	Yes, competition between teams	No
Houston (Consolvo et al., 2006)	Friends	Yes, competition between teams	Yes,users are allowed to comment on other people's data
Fish 'n' Steps (Lin et al., 2006)	Friends	Yes, competition between teams	Yes,chats are allowed between team members
EatWell (Grimes et al., 2008)	Local Community members	No	Yes, the system allows people to do so

Table 2.3: Example of Social Features in Previous Systems

the context of dietary change that if people are given too much information, they might overestimate the problem.

2.6.4 Social Support and Social Influence

As we described in behavior change model, social support and social influence are two effective ways to motivate and facilitate behavior change. The basic social features, as described within HCI literature, are simply sharing the data about a user's behavior with other people. For example, the Houston (Consolvo et al., 2006) allows the user to simply share own step counts with other peers. Based on simply sharing data, competition and feedback are expected to be implemented. Below, we summarize how social features were implemented in several example systems

In MAHI (Mamykina et al., 2008), patients' notes and food logs are shared with an expert instead of peers. Experts will look at patients' notes and food logs to figure out what's problem a particular patient might have and provide the patient with guided questions and actionable suggestions. In contrast, the other example systems try to form a "patient" community in which people who want to change behavior could help each other. However, it is not very clear why some studies chose to form real friends as a group while others chose to form strangers as a group. While one participant in Chick-Clique (Toscos et al., 2008) said "If I just compared my steps to everyone in class

I don't think it would really much affect me at all, I don't know why I just don't think it would. I liked that I was really close friends with them because then I could be open with my steps and all and I felt more comfortable talking about it with them.", it is still not very clear whether there is any difference in interaction patterns and intervention effect between sharing with friends and sharing with strangers.

In the two examples that implemented competition feature, it shows that competition is not suitable for every single users. The mixed reactions to competition, as described in Fish 'n' Steps (Lin et al., 2006), suggest that competition could have positive effect on people but such effect could not be persistent. On the other hand, some people feel competition is meaningless and even negative. For example, one user in the Chick-Clique (Toscos et al., 2008) said "I can see that if we were really competitive and stuff that after a while if you just got really sick of everyone doing better than you or if someone got really braggy about it. But I don't think that would happen with us". One way to solve the problem is to change the way people compete with each other. Instead of letting people directly compete with each other, the system could design individualized competition for each user to achieve and share each individual's achievements with other people. In this way people perceive the competition less directly and the negative feeling could possibly be reduced.

Feedbacks like comments or encouragement messages could help people keep changing their behavior. However, within HCI literature, there is nearly no analysis of those feedbacks generated by users. Open questions like what type of messages could be abstracted and applied into user interface? what type of suggestions people want to get from other people? and whether suggestions and encouragement from different people (friends v.s. experts v.s. strangers) have difference effect on people's performance? are worthy to be explored in the future work.

2.7 Discussion: Beyond Persuasion

In previous section, we reviewed four key features of behavior change technology and how previous systems implemented those features. Most of those example systems are mainly inspired by the notion of persuasion and one main focus of those systems is providing sufficient support to simplify behavior change. However, as (Purpura et al., 2011) pointed out, while those systems make certain actions like monitoring behavior much simpler, they also take the opportunity that people could proactively engage in understanding, practicing and maintaining behavior change away from users. In other words, current behavior change system design follows the design philosophy that aims to make people more efficient when carrying out tasks. The design philosophy was developed in the context of office and business, in which people expect that systems could improve their productive and reduce cognitive loads (Hallnäs and Redström, 2001).

But in behavior change, even people dream of a efficient, fast, not-painful change, the reality is quite different. Thus, in this section, we explore alternative design notions for behavior change technology.

2.7.1 Alternative Design Notions

As we discussed before, behavior change is a complicated process and requires people to proactively engage in the whole process of change for a long term. The slowness in learning, practicing and maintaining a new behavior is the nature of behavior change and it is not bad at all. In fact, such slowness gives people the opportunity to know themselves, the environment and other people much better. It expects people could slow down their pace to reflect on what they did and what they could achieve. The slowness, therefore, should be acknowledged by system design and be supported in technology-based intervention. Slow technology (Hallnäs and Redström, 2001) thus could be introduced to help us re-consider how we design behavior change technology. Unlike fast technology (e.g. the persuasive technology we described above) that compresses time to achieve given tasks, the slow technology aims to supply time for people to reflect on what they do, what resources they have and what they could achieve. For example, instead of simplifying how people evaluate food by providing auto or semi-auto calorie counting, the system could create opportunity for people to proactively engage in the evaluation of food to help people reflect on their food choice and improve their understandings on what is healthy food.

In addition to the slow technology, the mindful technology acknowledges the importance of what people think and feel about behavior change. As defined by Shapiro (Shapiro et al., 2006), the mindfulness is the development of an awareness of awareness and paying attention to intention with an open, non-judgemental attitude. In a blog article discussing mindful technology (Chen, 2011), Chen pointed out that:

> Mindful technologies focus on engagement with one's mental and emotional state as a medium for behavior change and well-being. This focus on technology-mediated reflection is contrasted by many health technologies that are reactive in nature: technologies that prescribe behavior change.

Indeed, those persuasion-based technologies try to prescribe behavior change in a way that technologies encapsulate existing behavior change theories or guidelines into a digital agent. But those existing behavior change theories are usually static and inflexible (Riley et al., 2011), and thus persuasion-based technologies can not make frequent iterative adjustment on the intervention to meet new requirements. Moreover, those persuasion-based technologies do not consider the real constraints people face in their everyday life. For example, people who want to change fruit consumption might face

external constraints such as no fresh fruit supply at local community or personal constraints such as no enough money to afford weekly fruit consumption. Without addressing those constrains, systems will possibly define unreasonable and not actionable plan for users. Consequently,similar with the view held by Don Norman (Norman, 1988) that users usually blame themselves when computer interface can not work, users of such persuasion-based technology will attribute blame to themselves when they fail in the process of behavior change. The mindful technology, therefore, proposes the alternative design for behavior change that instead of steering users towards a goal designers have in mind, designers should create tools to support people's reflection in pursuit of goals set by those people themselves (Munson, 2011). Furthermore, several open questions are inspired by mindful technology: what people think and feel when they interact with the behavior change technology? Do they really think the prescription given by technology is reasonable and actionable? Do they feel guilty when they fail? and How the system possibly react to relieve such guilty and properly help people re-gain confidence in changing?

The above discussions further remind us of the difference between personal trainers and current persuasion-based technology. Unlike persuasion-based technologies, personal trainers, even though they also design interventions based on those static behavior change theories, have the ability to constantly interact with clients to offer personalized intervention that is agreed and satisfied by clients. The interaction between them usually involves dialogues around the progress the client makes, the understanding and beliefs the client holds, constraints the individual faces and resources the individual has. Such interaction gives both of them , the personal trainer and the client, enough time to reflect on the whole process of behavior change and allows them to make sufficient change on the intervention. In fact, the interaction could be viewed as the process of negotiation between the client and the personal trainer; it means both of them need to compromise a bit based upon what they think and work together on the behavior change plan. In Maitland's PhD thesis (Maitland, 2009), the notion of negotiation is firstly proposed as the alternative design framework for behavior change. Maitland first argues we need to shift our focus from behavior change to health promotion, which is defined as:

> Health promotion represents a comprehensive social and political process, it not only embraces actions directed at strengthening the skills and capabilities of individuals, but also action directed towards changing social, environmental and economic conditions so as to alleviate their impact on public and individual health. Health promotion is the process of enabling people to increase control over the determinants of health and thereby improve their health. (World Health Organization, 1998)

As we noticed, the term health promotion highlights that the possible change dose not just require the person who wants to change to make efforts but also require all stakeholders such as government, neighborhood community, and local stores to make changes in order to provide sufficient resources and overcome constraints. The interaction between all of those parties, as called by Maintland, is a process of negotiation. Particularly, Maitland highlights that technology could be designed in a new way that help people to exploit, highlight or develop resources or acknowledge constraints. In terms of acknowledging constrains, Maitland suggests technology should acknowledge that people have different constraints, which are unavoidable and hard to be overcame by individuals, and appraise small progresses users make under those constraints. Besides, while it is impossible for technology interventions to give people more money to buy healthy food or build a new store to provide more healthy food, technology interventions do have the ability to help people reflect on how they could organize and use available resources. For example, a shopping planner could be developed to help people make reasonable shopping plan on a tight budget and tight schedule by highlighting cheap stores or farm markets and planning convenient route for shopping. For more details about the notion of negotiation, we refer readers to (Maitland, 2009).

In a summary, above alternative notions for behavior change technology: slowness, mindfulness, and negotiation do challenge the persuasion-based technology mainly from three aspects: (1) taking too much control and not give people enough opportunity to proactively participate in the behavior change (2) steering users based upon designers' mind instead of letting people argue what they could achieve and what they need and (3) not addressing resources and constraints people have in everyday life.

2.7.2 The Approach: Leveraging People's Thoughts to Support Behavior Change

The above design notions all mention the importance of reflection , which could be viewed as the process to bring unconscious piece of thoughts, beliefs and attitudes to conscious awareness, thereby one can leverage those to make better decisions (Sengers et al., 2005). In other words, as we discussed above, people choose food based upon various factors they may or may not be aware of, and the great effect of reflection is to help people constantly monitor and examine own experience in order to improve or gain knowledge from that and thus become aware of things they may not consciously take into account to take their first step toward possible change (Boud et al., 1985; Sengers et al., 2005). This reflection based approach to motivate and support reflection might be further illustrated by McAlpine et als reflection model(McAlpine et al., 1999) (see Figure 2.8).

In their reflection model (McAlpine et al., 1999), it involves the interaction between six components: goal, knowledge, action, monitoring, decision making and corridor of

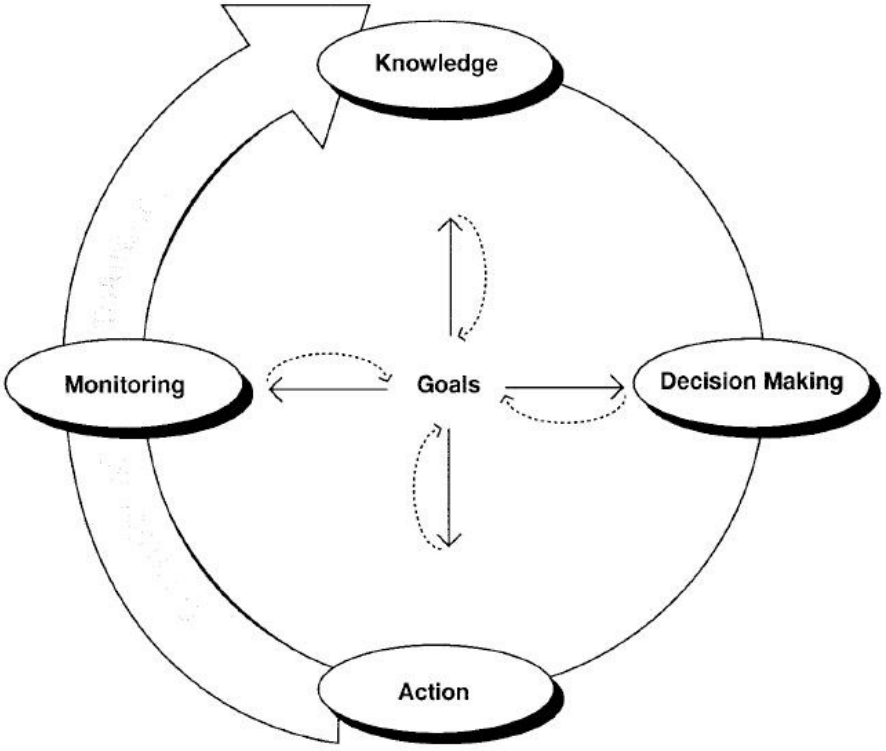

Figure 2.8: How Reflection Interacts with Other Components (McAlpine et al., 1999)

tolerance. As such an interactive system, reflection is driven by goals. In this thesis, such goal could be achieving one specific healthy eating behavior, for example, "Eating 5 Portion of Fruits or Vegetabels A Day" (goal). Then through (1) keeping monitoring how much fruits and vegetables have been consumed per day (monitoring), (2) recording what constraints and contexts are (monitoring), and (3) assessing how well the goal possibly could be achieved based upon the expectation (the corridor of tolerance), the individual could apply own understanding (knowledge) to make appropriate decisions to adjust daily diet plan (decision making) . And by keeping practicing "Eating 5 Portion of Fruits or Vegetabels A Day" (action) everyday, the person could learn from previous experience to build new knowledge, understand what actions should be done, and possibly adjust the goal to make it more reasonable. Indeed, this model actually describes such a iterative process that "the reflection is driven by goals, resulting in plans drawn from knowledge, leading to actions that are constantly being revised and updated as feedback is monitored through the corridor of tolerance and decisions lead to adjustments in actions(McAlpine et al., 1999). Therefore, the reflection could be viewed as an ongoing conversation between present action, past experience, and intentions for the future (Yinger, 1990). And the behavior change technology should play a important role in the process to serve as a platform to help people capture past experience, improve knowledge and refine actions, and then to achieve a intended goal in the future.

In this thesis, we thus argue that we could design such a platform for people by exploring

an alternative approach to behavior change technology: leveraging what people think about the particular health behavior to acknowledge their uniqueness, difference, and constrains in thoughts on behavior change and support their changes. This approach actually is grounded in the theory of planned behavior (Ajzen, 2002) and social cognitive theory (Bandura, 1985), which both point that people's thoughts and beliefs affect how people behave. We here, therefore, try to combine above three notions we briefly reviewed here: By slowing people down to express what they think, we could capture their thoughts and use that in the intervention to help people become aware of their unconscious thoughts and improve their decision making, then further support their negotiation with coach(system/application) to achieve their behavior change goals.

However, it is worthy to note that this thesis work does not aim to provide a complete solution; Instead, our aim is to explore how can we achieve this approach by mainly investigating the very first step: how can we create opportunity to allow people express what they think and believe about food so we can understand and capture those thoughts. As we will later mention, this thesis first explores what kind of thoughts on food and diet could be leveraged and we determined to narrow our focus on people's thoughts on food's healthiness to start the conversation between the user and the coach (i.e. the application/system). And then we aim to develop technology that is capable of capturing those thoughts and find out how those data captured could be leveraged to help people make better decisions. In addition, we are also interested in how those thoughts we captured from people could lead us to further reflect upon the design of behavior change technology. Especially, we expect those thoughts we could investigate in this thesis could help us address the gap between what people think about food and diet and what current design of food(diet)-related application try to "push" people to think.

Chapter 3

Initial Study: Investigating What People Talk About on Diet-related Forums

In the previous chapter, we first reviewed the general background of health behavior change (especially dietary change) then critically went over technology features that are employed in current system designs. Our critical review of previous systems pointed out that the persuasion-based approach might not work in long run and we argued that we need to leverage people's thoughts and beliefs about food to help people achieve sustained behavior change.

The very first step of this approach is to understand what kind of gaps exist between supports offered by existing interventions and supports people need so we can further explore how people's thought and belief about food might be leveraged to fill the gap. Our approach to investigate the question is through studying how and what people talk about food on two online forums. The forums we selected are both communities for food logging application users. We deliberately selected this type of forum because we mentioned in last chapter that interventions for dietary change often involves food logging. The food logging application offers a way to monitor behavior rather than thoughts so we expect to find out what kinds of supports for dietary change are missing in current interventions through the lens of people's conversation on the forums. Moreover, since forum is an open space for people to ask and answer questions or share their experience, it could be viewed as an open "diary" recording people's thoughts and belief about food. We, therefore, can summarize topics on forums to help us narrow our focus on what kinds of belief about food we want to further explore and leverage. Studying the forums also allow us to look at how people support each other. Since forum is a natural social space therefore we have the chance to understand how people talk to each other, what kinds of supports people expect from peers, and what kinds of topics

easily attract people to participate in the conversation. We then can understand how people's thought about food might be leveraged in future interventions through social approach. We might need social mechanism to collect people's thoughts or process people's thoughts to further assist dietary change.

3.1 The Study Design

In this exploratory study, we chose to study online discussion forums to understand what thoughts people have about diet and how they interact with peers. Our criteria for selection of forums were: First, the forum should offer a food logging product so we could investigate how those users of food logging technology talk about diet on the forum. second, the forum should be accessed by public so we were able to get the data. Two forums were finally chosen:TWYE's forum and precision nutrition's forum.

TWYE[1] is a free of cost service based on twitter, a popular social networking platform with 175 million users. TWYE allows users to log what they eat through twitter via web or SMS, or through a web form directly on the TWYE website. The log is formed of "food entries". Each food entry comprises an optional food picture, a textual food description, the number of calories associated with the food, and the time and date when the food was eaten. Precision Nutrition [2], is a commercial product which provides diet coach service; all users on this forum are paying clients. PN provides users with web tools to track own progress via answering heuristic questions, logging body measurement and taking picture of body and PN also provides users with a forum board to record detail diet journal. On PN, only clients are permitted to write on the forum, but the forum is public for reading.

3.1.1 Data Collection Method and Dataset Overview

We downloaded messages from the TWYE forum in the period from 1st June 2010 to 26th September 2010. In this dataset, there are in total 553 posts including 180 original posts(thread starters), and 373 replies. Six of the posts are not written in English, and were therefore excluded from this analysis. The messages were authored by 163 users, 114 of them initiated discussion threads. The average messages per user is 8.22, the average threads initiated by each user is 2.13, and the average reply per user is 6.1.

A number of different forum boards are associated with PN, we choose to analyse the "Nutrition Question" board because it's content is the most similar to the TWYE forum's. We analysed messages from the same period as the TWYE selection. The data

[1] http://www.tweetwhatyoueat.com
[2] http://www.precisionnutrition.com

our findings	**example post**
Describe contextual information and personal condition to point out barriers to diet	"My schedule at work is unpredictable"
People like to connect with other people who share similar conditions	"I am 45 and have about 40-50 lbs to lose, as well. Let's help each other along!"
Experiences in diet are shared to either solving similar problems or showing similarity between peers.	"I have a friend who uses liquid orange-flavored fish oil in her supershakes and she says even her kids like to drink the shakes."
Express feeling to food	"...I hate veggies..."
Celebrate success or complain frustration on diet	"I feel fantastic, and alot of body fat is gone!"
Ask other people to review public information or confirm own understandings and practice	"Just wondering if people agreed that a small handful of mixed nuts (almonds, brazil etc.) was about 50 cal?"

Table 3.1: Finding Summary Table

contains in total 248 threads, including 1620 posts and 236 posts have replies. The messages in dataset were authored by 291 users, 155 of them initiated discussion threads. The average posts per user is 5.57, the average threads initiated by each user is 0.85 and the average reply per user is 4.71.

3.1.2 Data Analysis Method

To analyze posts on two forums, we used content analysis method inspired by grounded theory (Strauss and Corbin, 1990). The analysis started by tagging posts at the sentence level through open coding. Two researchers coded the first 100 posts independently then cross-examined the coded posts and agreed on codes we would use to tag posts. We eventually had 41 codes and then grouped them into 13 categories: *how-to question*, *knowledge question*, *contextual information*, *personal condition*, *emotions toward diet*, *emotion to food*, *sharing opinion on information*, *similarity-buddy*, *similarity-solution*, *goal*, *knowledge*, *alternative* and *method*. Among those categories, we found *how-to question*, *knowledge question*, *goal*, *knowledge*, *alternative* and *method* were less surprising and uninteresting as they were already documented in previous literature as reported in Chapter 2. Therefore, we focused on only those 7 categories.

3.2 The Findings

In this section, we report our findings on what issues people discuss on forum and how they discuss those issues (see Table 3.1 for a quick summary of each category).

3.2.1 Contextual Information and Personal Conditions

On both forums, we observed that people ask for advice on how to change the way they eat. When users ask questions, they describe contextual information like location to frame their own eating patterns. For example, on PN, one person reports he spends *"A LOT of time in airports"* and asks whether it is possible to "eat better" at airports. Other contextual information frequently mentioned is food-related events or activities. On the TWYE forum one user asserted that she really wanted to start diet , but she *"got to go to dinner for daughter's birthday and she wants to do fondue. no lo-cal choices there at least I don't think there is anyway"*.

In addition to contextual information, people also mentioned their *lifestyle* as a barrier to a specific eating style, generally because of being short of time or on an unusual schedule. For instance, one person on PN wrote: *"...I work in the evenings (3-12) and am often stuck in a car during my shift. My schedule at work is unpredictable.....Anyone have any really portable food options for dinner?..."*. We found similar examples on TWYE: one man who thought that he was not overweight but want to change the way he eats said he follows a very *"busy lifestyle"*, which makes him feel that he has no time to eat better.

Food preference is another topic people discussed on the forums. In one thread on PN it is asked: *" ....Has anyone tried to take the Gourmet Nutrition recipes and make them soy, dairy, wheat and peanut free? I can eat walnuts and almonds but no dairy, soy, fish, other nuts, peanuts, and most wheat products at all... sometimes even eggs are out..."* .

3.2.2 You and Me: We Are Similar

When people consider who will be their buddies (i.e. friends on forums who could support each other), they intend to find other users who are in a similar condition. For example, a post on PN mentioned: *"I'm 44 yrs. old, and it's my goal to lose about 45-50 lbs."* and someone replied: *"I am 45 and have about 40-50 lbs to lose, as well. Let's help each other along! I am sure we can do this."*. This also happened on TWYE. A user called for buddies writing: *"Female, age 56, weight 174..goal weight 144. Need a support group"*; another user replied: *"Female, age 59,weight 185. I, too, am good about sticking to a commitment, however, I've tried to do this alone too many times before. I'm puttting myself out there for the first time"*.

3.2.3 Sharing Similar Experiences

On both forums, we noticed that similarity is the strategy people usually employ to answer questions. Users tend to share their experience if it is similar to a question posed

by someone else, or to use their own experience to analyze the problem posed by other users and in this way provide suggestions and solutions. For instance, one post on PN mentioned *"Foods sometimes taste metallic to us"*; someone replied: *"...I get that taste frequently when I lower my carb intake. I know it sucks. Not sure how to get rid of it though. Maybe send a PM to Noel Cruz."* On TWYE, there are also similar cases. One common problem people shared on the forums is emotional eating: when people feel stressed or bored they want to eat. One person on TWYE wrote: *"Hey, I did the same as you: I was always hungry when I was bored..."*.

On PN we also observed that users answering questions point out similarity between the behaviour of the person asking the question and someone else they know. For example this may be based experiences of friends or family members. In these cases users would share an anecdote, they would analyse the posed problem and offer suggestions for a solution. One example of such answer from PN: *"I have a friend who uses liquid orange-flavored fish oil in her supershakes and she says even her kids like to drink the shakes."* Moreover, since people log food information online and this information is accessible by others on the forum, sometimes a person answering would suggest people asking questions to directly check the food log of specific users know to be in similar situations. One person on PN answered question in this way : *"Look up what Ryan Andrews eats on this site. He gets enough protein without eating any meat, but from legumes, vegetables, nuts, seeds, etc."* The same pattern was not observed on TWYE. This might be because TWYE's users are less professional than PN's users and therefore usually don't refer to others' logs and opinions.

3.2.4 Emotion: Loving Food

The link between emotion and food is another thing we clearly found in the data. On both forums we observed people writing how they love food and reporting their pleasure in eating. This is the natural reaction of humans to food. Because food meets humans' needs of getting energy, the perception of being full and the good taste of food can cause brain to trigger pleasure signals, so people feel happy. In forums, when people talk about food, they use emotional words to express their relationship with specific items. For example, on PN we read about how a user loves the a specific product: *" the taste of the chocolate one [GNC brand Whey Protein Isolate] is pretty good."*. Beyond talking about feelings of food, people also express how emotions can make it difficult to change the way they eat. One person on TWYE wrote *"Want to get down to 9st 7lb and tweeting is helping although I love food especially chocolate"* . The passion for specific food makes it hard for people to get weight down and reach specific goals. This kind of conversations were also found on PN. For instance, one person on PN wrote: *"...I hate veggies..."* while they are required by the PN programme.

3.2.5 Emotion: Enjoying Diet

On both forums, we observed people either celebrating their success or reporting their frustration about the ability or inability to follow a diet and observe results. One user on PN wrote a post to announce he will try some new plan, after a period of practice, he replied to his own post to celebrate his own success: *"just wanted to post that this is going GREAT! I feel fantastic, and a lot of body fat is gone!"*. Another example from TWYE showed people expressing frustration about diet: *"My current weight is 197, I'm so mad because the amount food, and calories, I put in my body was cut in half. I even work out, but no results"*. Besides expressing emotion explicitly, people also used some implicit cue to celebrate own success , for example, one person on TWYE showed off that *"Today I'm wearing jeans which had not fit the last fix months..."* and the author wrote that such good feelings really *"...help me going on(diet)..."*

Another aspect related to emotion that we found on the PN forum is that some people consider diet an experiment which is playful and enjoyable. For example, one person wrote on PN: *"I'm just experimenting with some stuff to see what else is out there"*. For those people, diet is a process like research which means they enjoy the exploration of food-health relationship to see how food will affect their own health and still keep them happy. However, similar words were not found on TWYE, we think it is because PN users have a clear diet plan and a specific attitude towards it. In contrast, TWYE users have goals that are not clearly defined and they might not be following a diet at all, so they usually care more about whether they eat less or equal to a specific calorie value and do not focus on testing new ways.

3.2.6 Share Opinions on Products and Information

We found that people use forums as a place to share information about methods or products and ask others' opinions about them. For example, one person on PN asked: *"Hey everyone, I was reading an interesting article on 'the myth of meal frequency'; over at elite and will like to know your thoughts on it. Here's the link http://articles.elitefts.com/article...eal-frequency/Cheers,pshake"*.

Besides asking opinions about published articles, people also confront others about their own thoughts on diet methods (e.g. specific techniques, attitude towards products, recipes, etc.). One reason for this is that people have doubts. For instance, one person on TWYE asked: *"Just wondering if people agreed that a small handful of mixed nuts (almonds, brazil etc.) was about 50 cal? "*. Another reason might be that people find something is not covered by books, so they ask others to provide suggestions and confirmations, as one post on PN demonstrates: *"My concern is that i noticed that alcohol is not addressed in the books and when I did a search I didn't find much. I'm*

just curious what leave of alcohol (if any) is allowed on the PN plan for someone with my goals".

3.3 Discussion

In this chapter, we presented an exploratory study of online discussion forums related to two diet services: TWYE and PN. Through analysing the posts on two forums, we gained better insight into: (1) what diet related problems people discuss online (2) how they interact with other people to ask for supports. Those findings presented above potentially open the space for design for dietary change. We discuss potential design implications for our research idea in detail below:

3.3.1 Rethink about Food Logging Technology

The current design of food logging technology has three major parts: (1) Recording what people eat in multimedia (text and/or photo) ;(2) Presenting semi-auto/auto generated nutritional assessment (usually calorie counting) and (3) Presenting the progress towards the goal, which is set by the user or the application. Our findings in this study identified two types of information that we believe current food logging system and even most food-related systems do not well leveraged: emotion information and context information. We believe these two types of information demonstrates the gap between what people think about diet and food and what applications usually could capture. Thus, the findings encourage us to rethink how could we re-design the food logging technology and food-related application in general, especially those designed for dietary change.

First, we think about what role the emotion information plays in food logging technology. As our finding on emotion to food/diet suggest, emotion is one piece of missing information in current design.We believe that logging how you feel about the food could assist people in:

1. Reflecting on whether you are happy with current diet and want to continue

   As we discussed in related work, people's reflection on the process is one important aspects we want to support. How people feel about the process of behavior change reflects whether the behavior change intervention itself suits a person or not. If a person feel quite negative to what the intervention suggests to do then likely the person will not be able to succeed in behavior change. Therefore it is important to log how people feel about the intervention (i.e. the food he/she ate in the context of diet). This piece of information should be logged and presented along with what people eat to show people's progress so people not only understand how well i eat (i.e. whether what i eat is healthy or not) but also understand am i happy to

continue. This information also could help dietitian or other professionals to offer suggestion on how the individual might tweak the intervention plan to satisfy the individual: not only achieving the behavior change but also feeling enjoyed during the whole process. However, we acknowledge that logging emotion is not easy. The self-report method might be used first in research to understand how emotion information might effect people's performance on diet but more objective method is expected in future.

2. Spotting emotional eating and try to break the bad link.

   Emotional eating refers to "the tendency to overeat in response to negative emotions such as anxiety or irritability" (Van Strien and Ouwens, 2007). The ability to monitor such link between emotion and food intake, thus, is important for people to understand how they eat in respond to a particular emotion and how they might to break such link in future to avoid overeating. This requires the system has not only the feature to collect emotion data but also the feature to visualize and offer insight into link between emotion and food intake. Regarding the visualization, we suggest the future system should help people (1) understand how one particular emotion might affect food intake: this information assist people in understanding why they overeat or choose "bad" food in a particular emotional condition and (2) how food intake might affect emotion: this information help people identify those foods arousing positive/negative emotion. Then people might better choose food in response to a particular emotion.

Second, we think about the role of contexts in food logging technology. As we found in the study, contexts such as location (where the person eats, e.g. the airport) and time (when the person eats, e.g. during night work) do matter when people think about what to eat. Current food logging system designs start logging location and time, but the link between those contexts and food intake is not properly visualized and used. Current food logging systems more focus on nutritional assessment and do not yet value the social and environmental factors as pointed in Chapter 2. Therefore, even though some of food logging systems log the contexts information (usually,location and time), those information are not yet fully leveraged to offer users insights into how those specific time and location context affect their food choices and how they could make changes based upon the insights they gained from those context information. We here, thus, strongly encourage future food logging systems should try to think beyond "how good is my nutrition intake" to "how well do I eat". The latter question is much informative and could help people, novice or experts, to frame the definition of healthy eating properly. To achieve this goal, the value of contexts should be addressed in future design and the link between contexts and food intake should be visualized to help people know how certain contexts affect food choice and reflect on what they could do. A simple example could be a map of how healthy you eat at different place: if the system measures how

healthy you eat by calculating calorie, then a map could be generated to illustrate where you consume highest and lowest calories.

3.3.2 Think about the Social Supports

Through our analysis, we found there are two types of social interaction patterns in diet-related social networks: one is social review mechanism and the other is similarity-based interaction. The first one, social review mechanism, means people need to validate thoughts and actions by discussing with other people. Such discussion could help people in two ways:

1. It could help people validate their own thoughts and actions by letting them know how many people support the thoughts or actions.

   For instance, we observed a person listed own healthy condition, goals and a typical week's meals to ask whether it is a good plan. People who replied the post listed the cons and pros of the plan and offered suggestions on how to tweak the plan. The social network, in this case, serves as a review system to judge people's thoughts and actions in order to offer new perspectives on how to make improvement. Considering the idea we mentioned above about "how good I eat", we think introducing such social review mechanism could be a new way to measure how good a person's diet is. Unlike the nutritional assessment, people's reviews on a person's diet practice could offer more information and became more meaningful. The factors like environmental factor or emotional factor usually are not taken into account by machine in the case of nutritional assessment, but they could be addressed by humans. Therefore, we suggest future research could explore how to utilize social network to help people judge their diet practices in order to help people make real changes. How to coordinate such social review and How to manage and present those review data are two open challenging questions for future research.

2. It could help people to validate information from other sources

   Huge amount of information on how to eat healthy could be accessed by people via various media including newspapers, tv, internet and so on. However, without certain level of knowledge, people can not always be sure whether a piece of information is credible and reliable. As we reported above, people on forums posed questions regarding information they found and asked other people what they think about the information. The discussion around the information help people think about whether the information could be trusted and how to use it.

The second one is similarity-based interaction which means people want to connect and interact with people who are similar with them. Such similarity between people could

be similar life experience (e.g. both working at late night), similar personal condition (e.g. both having heart disease), or/and similar plan(e.g. both aiming for losing 10 kg). We observed that people deliberately mentioned such attributes in their conversation on forums to build friendship and offer suggestions. However, it is a challenge to think about how to coordinate such similarity-based interaction in other settings other than forums(e.g. web application or mobile applications). PatientLikeMe [3] might be a good example to demonstrate how certain basic similarity attributes could be identified: similarity between people are recorded by asking people detail questions on what is the patients condition, what medicine the patient takes, what is the patient's age and gender. Those basic data help system create a basic profile or "persona" to describe a group of people in order to link them with each other. Furthermore, the more detail data such as daily medicine intake, weekly symptom update help to further "quantifying" the patient and more detail similarity could be identified in a long term interaction. In the context of dietary change, such approach could work too. By logging what people eat every week and how they respond to it physically (e.g. nutrition needs) and emotionally, we have a great potential to better quantifying people in terms of how they eat and could better help people who are similar with each other build friendship and offer supports.

3.4 Conclusion

In this chapter, we explored our initial design ideas by looking at what people talk about on forums. We see forum is a space recording what people think and reflect, thus, the issues discussed on forums indicate what type of support people need. In addition, we also are interested in how people interact with each other since social support is crucial for behavior change. We identified two type of information: contextual information and emotional information as important content should be leveraged in future food logging systems in order to better define and assess the concept of healthy eating and help people better understand how they eat. We also identified two types of social patterns: similarity-based interaction and social review mechanism. We see these two social interaction patterns fit our initial design ideas well in terms of how we might assess people's diet and how we might coordinate the support between peers. In next chapter, we will present our first design prototype to demonstrate our design ideas which incorporate some findings in this work.

[3] http://www.patientlikeme.com

Chapter 4

Design Prototype

We reported our initial design idea: a social food logging system for people to reflect upon own dietary behavior in Chapter 2 and reported a preliminary study that explored what people think about diet and how they interact with each other to generate further design ideas in Chapter 3. In this chapter, we report the first prototype we developed on iPhone to demonstrate how we implement the design idea. In the first two part of the chapter, we first describe the rationale behind our design idea and then explain the details of the design prototype. Then, in third part, we report the deployment of the prototype and user studies we organized. Finally, we report the real commercial applications that employ similar idea to move on to our next study.

4.1 Rationale Behind The Design

To support people's reflection during the process of dietary change, we propose to use the traditional diary method to help people assess and reflect upon their experience. Specifically, here we are interested in how people reflect upon their previous eating experience: thinking critically about how healthy their food choice is. In this particular context, the diary means a set of logs that recording how people eat: what people eat and how healthy it is. To log what people eat, we consider the tradeoff between how easy people could log it and how complete the eating experience could be captured. As we discussed in Chapter 2, the traditional food logging method is to write down food items in text but this method is quite time-consuming and hard for people to review the food item later. Picture becomes a popular media now to capture a relatively rich information about what people eat but it also leaves certain uncertainty for people (e.g. some details might not be clear in the picture). In our design, we plan to use picture as the main media to record what people eat. This is because we observe mobile photo applications like Instagram [1] are now quite popular and people are enjoying taking and

[1] http://instagr.am

sharing pictures in such applications. We think the popularity of photo sharing could be leveraged to make positive impact on people's logging behavior. In addition, taking picture itself could be viewed as a small nudge for reflection. In the design, we also offer optional text description for people to write down some important details which might not be obvious in the picture. For example, a list of ingredients of a sauce on steak could be wrote down to help people know much better about the food.

Regarding how to log how healthy the food is, we propose that instead of using traditional nutritional assessment that is usually done by machine automatically, we want to use a crowdsourcing method to solicit a large group of people's thoughts to judge the food's healthiness. Such approach is not new in web 2.0 since it has been recognized as social rating system (Feng et al., 2012) and has been used in various settings such as movie review scenario. IMDB, as one example, allows people to rate how good a movie is by their own standards and the aggregated rating will be used to rank the movies. Similarly, we propose to use the similar mechanism in our design. As we reported in last chapter, the social review mechanism is one social interaction pattern we observed on forums. We think it will be good to let other people judge the food for each other and the aggregated judgements could be used to potentially indicate how healthy one meal/food could be. However, we here do not use the simple rating as the review system;instead, we use a set of heuristics to indicate the healthiness of food. This is because a set of heuristics might be used to inspire and guide people to think about how to assess the food in a pre-defined way and also heuristics narrow various standards of healthy eating down to a fixed set of criteria which allow us and users easily interpret the result. The set of heuristics we use in the design are borrowed from Precision Nutrition. And it includes five basic heuristics for healthy eating assessment:

1. Does the food have healthy fat in it?
2. Does the food have complete protein in it?
3. Does the food have carbs in it and you eat it after workout?
4. Does the food have any fruits or vegetables in it?
5. Does the food have whole food in it?

For each heuristics, we ask people to rate "no", "some" or"rich". The introduction of "some" and "rich" is used to indicate whether the portion size of specific item in heuristic is big enough to be healthy. If it is not, then people should select "some", otherwise "rich" is the option.

Since we here introduce the social review mechanism into our design, it implies that we want people share their eating experience (i.e. the pictures of food) with each other. We believe that people could benefit from the sharing in two ways:

1. it offers people an opportunity to look at what other people eat and potentially learn from other people's experience

   In Chapter 2, we reported that social and environmental constraints make people feel hard to eat healthier. We think it might be a good idea to inspire people what to eat under certain constraints by exposing them to other people's experience. In such way, people get the opportunity to look at what other people eat and learn new ways to eat: cooking the same ingredients in a different way to avoid getting bored, finding unnoticed healthy foods at local, trying same type of food (e.g. cake) made in a new type of ingredient.

2. It help people to identify similar peers to connect and interact with each other. By looking at what other people eat, the user has the opportunity to identify similar peers who share the same eating preference, diet goals or other attributes in the process. It help people to connect with each other and further interaction between them could be encouraged in the future.

4.2 The Detail Design and First Prototype

The first prototype of the design idea is implemented as iPhone application. We developed the web server on Django, a python web framework, and a web application having same features as iPhone application was also implemented. The iPhone client then communicates with the server via API designed on the web server.

According to the design idea presented above, the concrete features the application has are: food logging by taking picture, food assessment by checking heuristics, feedback, sharing the photo in a timeline view, building friendship with peers, and progress checking.

1. Food Logging

   As shown in Figure 4.1, users are instructed to first taking a picture of what they eat (or selecting an existing picture from albums) then they are directed to write down additional information about the food and check whether it is a post-workout meal. The information about whether it is a post-workout meal is useful for the heuristic "Whether the food has carbs in it and you eat it after workout? and will be shown along with the food's picture and text description.

2. Food Assessment

   The food assessment occurs at two different time: when people log the food or when people look at a particular food log and make the assessment.

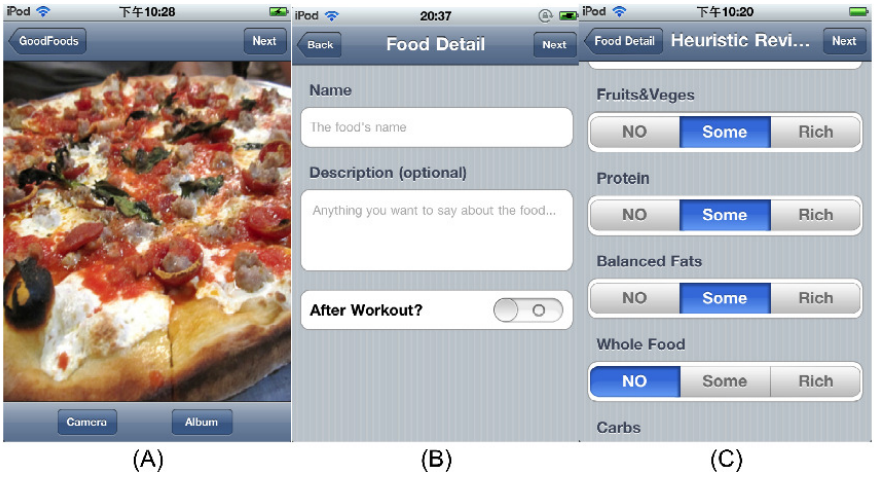

Figure 4.1: Food Logging Screenshot: (A) taking picture of food, (B) Log details (C) Heuristics Review

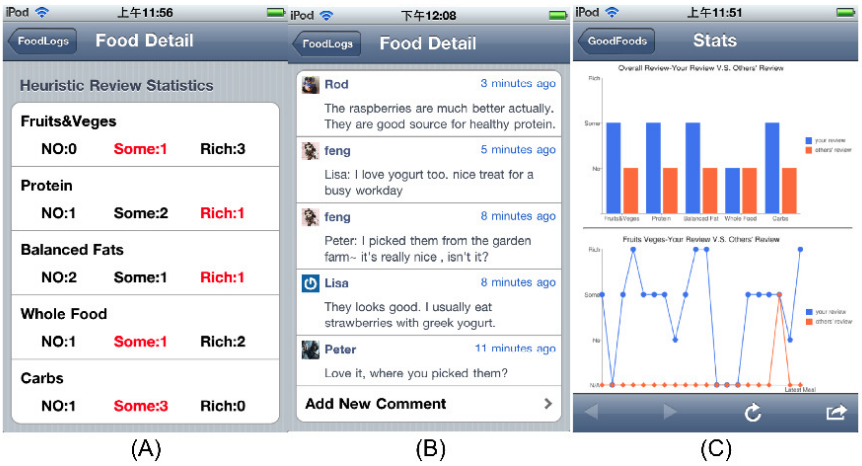

Figure 4.2: Feedback and Progress Screenshot: (A) Heuristics Review Result (B) Comments (C) Progress

For the first case, after people take the picture and write down additional information, they are instructed to check the five heuristics. The brief explanation of these five heuristics is available for people to read.

For the second case, when people look at a specific food log, they could choose to assess the food and then they will be directed to the assessment screen to complete the five heuristics checking.

3. Feedback

   The assessment contributed by people is one kind of feedback in the system. Here the feedback is not only useful for the person who actually logged the food but also for anyone who assessed the food log. For the person who logged the food, notifications of new evaluation for the person's food logs will be pushed to iPhone. The screenshot Figure 4.2 shows how the result of assessment is presented. As it shows, the distribution of rating of each heuristics (i.e. how many people vote for

one option) is presented and the option which the person, who views the result, chose is highlighted in red. This design is used to create the potential "conflicting" in order to let people notice how different other people think about the food and help people reflect upon their own understanding.

For other people who assessed the food log, after they assessed the food log, they automatically "follow" this food log. By "follow", it means any feedbacks to this food log will push a notification to people who assessed the food log. This "follow" mechanism is designed on purpose to help people, who assessed the food, keep monitoring the food log they contributed to and have the chance to validate their thoughts.

In addition to the assessment (i.e. heuristics checking), we also offer a free-text comment system to allow people freely discussing the food log. This feature is expected to stimuli discussion on how to eat healthy and allow people to build friendship through communication.

4. Friendship and Timeline

   In the system, we implemented the basic friendship feature that allows people to follow each other. By "follow", it means the user is interested in what food another person eats and want to keep monitoring it. All food logs created by people whom the user follows are presented in a reserve chronicle order; this is so called timeline. Each item in the timeline includes the food picture, food name, user's id and a timestamp. People could navigate from timeline to specific food log by tapping the log.

5. Progress

   To help people monitor their progress and identify long-term flaws in their eating behavior, we provide simple visualization for each heuristics. As shown in Figure 4.2, the visualization of each heuristic plots both the user's own heuristic judgement and other people's judgements together. By viewing those visualizations, people are expected to note when they eat bad, what problem is (i.e. which heuristic gets a bad assessment result) and what problems of their healthy eating understanding is (i.e. comparison between own judgement and others' judgements)

4.3 Deployment and User Study

After we developed the system (both iPhone app and web app), we attempted to deploy the application two times. At first we conducted a user study that did not deploy the system into open market. In the meanwhile of this study, we submitted our iPhone app to Apple Store for review and successfully deploy our iPhone app into the open market later.

4.3.1 First User Study

During the summer of 2011, we started recruiting participants for our first study. In the study, we mainly investigated two issues: first is the usability of this application and second is to look at how this new intervention help people make changes.

4.3.1.1 The Study Design

The study started by conducting a pre-study interview with participants to get basic understandings on their previous eating behavior, their knowledge on healthy eating and their experience in using technologies. Then we asked participants to use the application for 4 weeks (the detail deployment is discussed below). After it, we planned to have a post-study interview to see how people made change during the process and what they thought about the application.

4.3.1.2 Participants

In three weeks, we recruited 8 participants by sending emails to several university email lists including ECS msc students, PhD student and staffs and also other faculties' students and staffs. In invitation email we required that the participant should own an iPhone and was interested in healthy eating. The 8 participants we finally recruited included 1 staff member of General Hospital, 1 ECS undergraduate, 3 ECS msc students, 1 staff member of international office, and 2 PhD student in Institute of Sound and Vibration Research (ISVR). Two of them are female and others are male.

4.3.1.3 Pre-study Interview

The pre-study interview began by briefly overview the study plan and got participant's consent. Then the participant was asked to complete a nutrition knowledge assessment questionnaire. This questionnaire was used to get a base line of participant's nutrition knowledge level. 2 (P3, P6) of our participants have high level knowledge, 4 have intermediate level knowledge, and 2 (P1,P7) have novice level knowledge. Then a semi-structured interview was conducted, which covered four topics:

1. People's understanding on what is healthy eating

   We asked people how they understand healthy eating? to start the conversation. Participants' answers are not quite surprising since they mentioned those rules covered in mass media such as: eating vegetables and fruits, avoiding fats, eating natural foods, balanced eating, and avoid eating too much.

2. Where they get information on how to eat healthy?

   Participants said they get information on healthy eating by reading newspapers/magazines, talking with friends, watching tvs, and looking at information on internet. The rules they talked about above were mainly learned from mass media because those medias frequently advertise those rules.

3. What efforts they make to eat healthy and what problems they have?

   Two participants (P2,P6) made great efforts on healthy eating since they have own special diet requirements and restrictions. Unlike other participants, these two usually bring packed home-made lunch to university. Other participants usually buy fast food at university shops or eat at Chinese restaurants around the university. Those participants mentioned that they want to eat healthy when they actually go into the shop, they feel lost in those choices and just want to satisfy themselves. Other problems our participants faced are, for instance, losing interest in eating healthy foods like fruits. P5 said he tried to eat only healthy food for a week and give up because it is "too boring".

4. Their experience in using social networking service and photo-based application

   P3 and P4 were not experienced app users and did not join any social networking service. Other participants at least used one type of social networking service such as Twitter and Facebook. Only P5 and P7 used photo-based food logging application before, and P3,P4,P1,and P8 did not use any type of photo-based application.

4.3.1.4 The Deployment

In this study, we did not choose to deploy our app into Apple Store because the application was at alpha stage where we needed to fix bugs frequently but the review time of Apple Store is relatively long (averagely it takes about 3-4 days, sometimes even a week). Thus we chose to use a new iPhone app test platform, TestFlightApp, to manage our deployment.

The TestFlightApp platform provides developers with a web application to manage app testers and allows developers to assign different version of apps to different group of testers. It also provides a iPhone app to collect testers' UDID and allow testers to manage those test apps. In our invitation email, we asked all participants to install and register this app on their iPhone then we could use the collected UDIDs to register those participants as testers on Apple Developer Portal. After successfully registering those participants, we use the TestFlightApp to send notifications to participants to notify them that the app is ready for download. They then could download and install the app by the TestFlightApp on their device. The TestFlightApp platform monitors all

participants' activities including receiving our notification, reading our notification, and installing our app thus we could ensure all participants installed the app.

4.3.1.5 Lesson Learnt from this study

This user study unfortunately did not end with very informative results on how it changes people's behavior due to several reasons: (1) 3 participants ((P1,P4,P8) dropped off the study because losing own phone, upgrading the iOS to unsupported beta version (i.e. iOS 5.0 at the time of the study), having commitment to other important work. (2) people got bored after keeping logging food for around 2 weeks (3) people forgot to do the logging because of a busy schedule.

However, even though we did not collect enough data to allow us to look at whether it successfully change people's behavior, we still learnt some lessons from this study, especially on how people use the application. In our post-study interview, three of our participants (P3, P5, P7) were able to help us. The interview discussed following three issues:

1. Usability Issues

   Overall, they think the app is quite interesting and potentially could help them think more critically about what they eat. P3 complained about that".. the app is interesting but I hope it could allow me to duplicate previous logs if I buy the same food". P5 mentioned the feature to allow them to re-post the content to other service since P5 believed that it is always interesting to let others know what you are doing, but it is less fun to share with just a small group. All of them suggested it would better to allow them to log in plain text in case they forgot to take pictures.

2. Heuristic Assessment and Comments

   Regarding with the heuristic assessment, P5 thought it might be good to allow them define the heuristics since certain rule like "carbs only after workout" is not quite important to him. P7 suggested that the heuristics did help him notice certain aspects he did not know and it was interesting to view what other people thought about the food.

   P3 expressed strong interest in comments:"I enjoyed so much in discussion with a guy [P2] on the pasta I made at home.[I] was given great suggestions on how to tweak my recipe by reducing the amount of cream and replacing the corn oil with olive oil." Overall, three participants thought the discussion on a particular food did help them build friendship and know the food much better.

3. Sharing What I eat

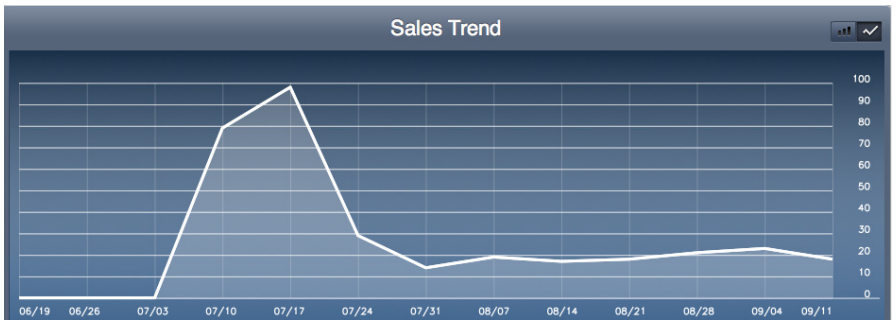

Figure 4.3: The Sales of the Application on App Store

Overall, all of them were happy to share what they ate with others. They thought they got inspiration by looking at what others ate and it was really a pleased experience. P5 commented:“ you know it was always hard to think about what to eat for lunch but those pictures from others inspire me to think new ideas and it is good”. However, a problem was raised by P7 “sometimes I ate bad and did not want to allow those people to see it. The app should let me hide them from other people.”

4.3.2 Open Deployment

We submitted the iPhone app to Apple Store while we were conducting the study. The app was approved in the early July and upon launch we got around 200 downloads in the first two weeks. After it, we got around 20 downloads per week in the following 2 months (see Figure 4.3). However, no users of our application were willing to participate in an interview to discuss how they used it.

The application was available in the Apple store for nine month due to the expiration of the Apple developer license, and during the nine month time, we had overall 637 users registered the application. Among them, 84% of them only used the application for one time, usually for a testing purpose. 12% of users used the application for more than once but stopped using it for a couple of tries. Only 24 users used the application for couple of meals in a row. Some of them used it to track the lunch everyday for several days and stopped using it while some of them used the application to track whole day meals including snacks and stopped using it after a few days. Those uses, however, did not use the social function but primarily used this application as a food logging system to track own eating behavior and probably reflect upon own behavior through answer heuristic review question. One reason behind it is we did not leverage any existing social networking in the process so people feel unconformable to comment on other people’s foods even though they can see their food pictures in the application. And also we in the process did try to operate the social network to encourage people to do more social interactions so the application in real world failed to encourage people to interact with each other and leverage the social feedbacks.

4.4 Similar Applications in the Real World

It is interesting to note there are two applications implementing the similar idea and being deployed into App Store after we developed the prototype: Pic Healthy [2] and the Eatery [3]. Both of them encourage people to take picture of food and evaluate food by rating them on a Likert scale for each other. The difference between our implementation and theirs is on how people evaluate food: these two applications use simple rating system to assess food while ours use a set of heuristics. This difference might be significant but we think it might not be worth to further test our own app at this time point. This is because that based upon feedbacks we received, a major concern is around the assessment UI and to understand this issue further, our focus will not be on studying the application as a whole but rather on a specific question that how people actually assess food and convert their thoughts into the assessment UI? In other words, we need to find out what attributes do they take into account to assess food?

By looking at the data generated on these two applications, we think it is also important to understand above questions in order to design and develop this intervention idea further. For instance, as shown in Figure 4.4, without comments regarding the rationale for the rating it is difficult to understand why one salad (Figure 4.4 A) is rated 4, but another one (Figure 4.4 B) is rated 3. Likewise it's unclear why toast and sausages (Figure 4.4 D) receive the same rating as the salad in Figure 4.4 B. Without capturing any information about the rationale behind rating, it is not clear what factors people take into account to assess a food's supposed healthiness. Motivated by out observation and the fact that there are two commercial application out there with more participants, we decided to focus on exploring our research question via taking advantage of materials from their applications. In next chapter, we will report our investigation into the particular question we mentioned above with real world food pictures from PicHealthy. In addition, we will also discuss how knowledge we gained from the study could help future designs to promote healthy eating.

[2] http://www.medhelp.org/land/photo-food-app
[3] http://eatery.massivehealthy.com

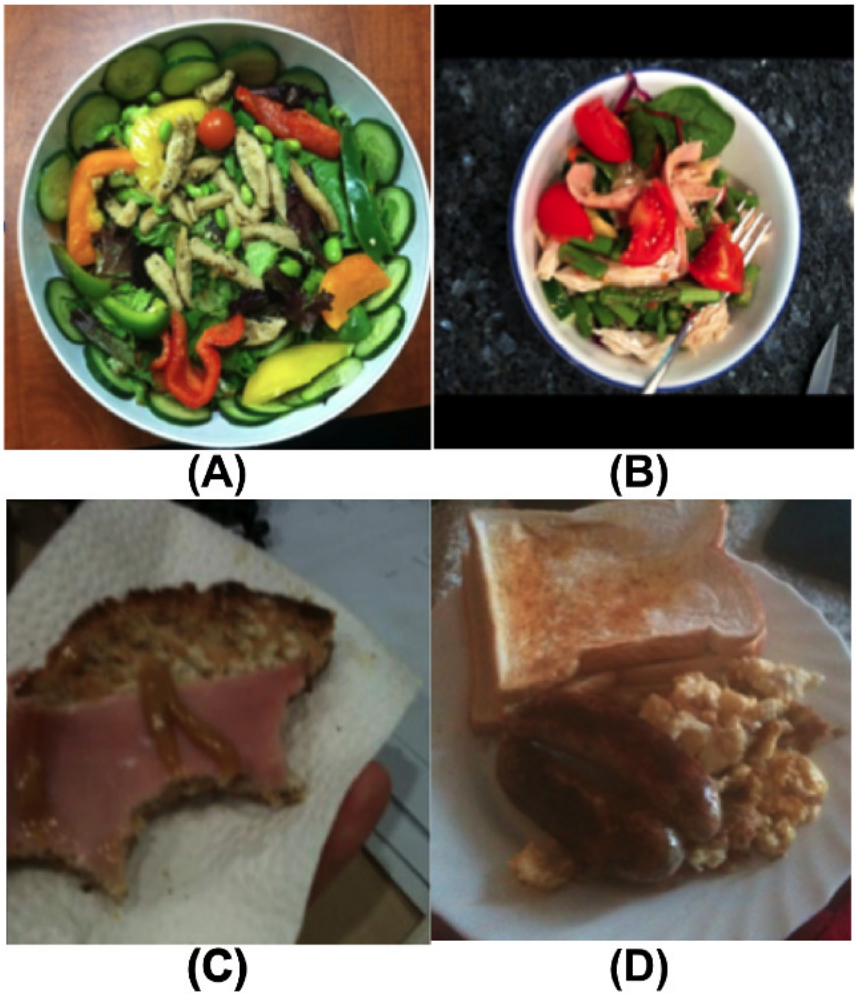

Figure 4.4: Sample food pictures from Pic Healthy and their ratings: (A) Chicken cucumber and pepper salad, 4 stars; (B) Chicken, tomato and spinach salad, 3 stars; (C) Beef sandwich, 2 stars; (D) Sausage, egg and toast, 3 stars

Chapter 5

Understand How People Evaluate Foods Healthiness and Its Design Implications

In the Chapter 2, we pointed out that most current dietary-related applications predefine what is healthy eating in their design and the dominant idea seems to be counting and controlling the amount of calorie you consume everyday means eating healthy. We argued that in contrast to those predefined healthy eating in interventions, people have their own understanding of what is healthy eating and in Chapter 3 we demonstrated that providing people with a way to express what they think about food could help them reflect upon own understanding. In order to facilitate the expression of people's understanding, our design featured the social sharing of food logging and crowdsourcing people's judgement of foods to help people practice their knowledge and express their thoughts for peer-review. However, as we shown in the end of last chapter, the reasons behind those evaluations are not clear and "invisible" to people in current design of the systems (e.g. Pic Healthy and the Eatery). Therefore, it makes the result hard for people to interpret and use for further decision-making. We think it is important to understand how people understand food's healthiness and apply their understanding in evaluation of foods in order to better design systems to facilitate the expression of people's definition of healthy eating (i.e. the understanding of food's healthiness). Furthermore, we believe the understanding on this issue also could inform systems to think about how to define healthy eating: pre-defining it based upon designers' view or adapting to people's views. To address the issue, we report a survey study to understand how users and potential users of food support technology take decisions about the relative healthiness of food options. Our aim for this work is to identify attributes that people take into account to evaluate healthiness of food and inform the design of better tools to support people in their choices. One major outcome of the study suggests that there is no one correct answer for food's healthiness. In other words, people pay attention to

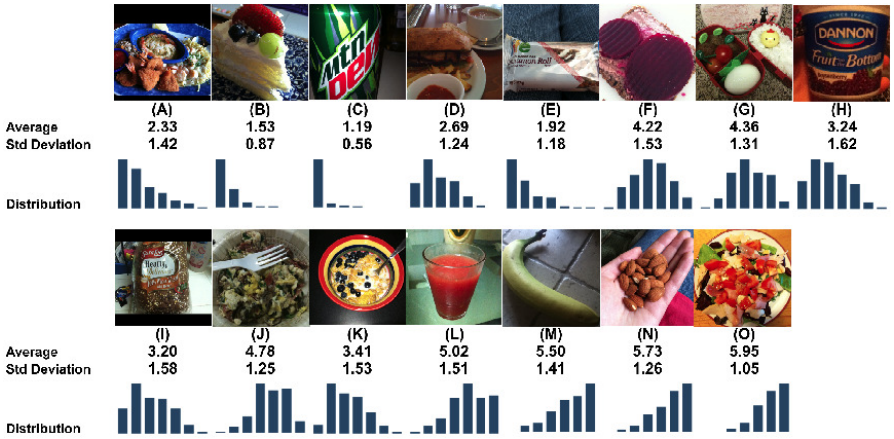

Figure 5.1: The 15 Food Pictures in the Survey with Their Average Rating, Std Deviation of Rating and Distribution of Rating

different attributes and weight them differently thus the correct answer does not exist based upon the crowd's opinion. To further verify our finding, we conducted a follow-up study to recruit a group of experts to take the same survey and then the findings of both study lead us to the discussion of "it depends" in the particular design space of dietary-related applications. In the following sections, we first present our study method and discuss how it differs from previous works. Then we report our results, and based on the analysis of the data we collected we discuss the "it depends" problem in designing for dietary-related applications and propose a suite of intervention opportunities for future system designs for healthy eating.

5.1 Method

In order to investigate what kinds of attributes people take into account when assessing food's healthiness, there are two major methods used in previous research. The first method is to use large-scale questionnaire or in-depth interview to elicit what people think. This method allows researcher to get large-scale data but fail to capture people's detail thoughts on concrete food items and ends up with a set of abstract and high level attributes. Alternatively, presenting people concrete food items and asking people to judge the healthiness could help researchers capture more details about what people think. For example, Carels et al. (Carels et al., 2006, 2007) tried to solicit food attributes by asking participants to rate 16 foods and explain their ratings in writing. In a first study (Carels et al., 2006), they asked 75 undergraduates to list 10 healthy and unhealthy foods they know. Then they selected the top 8 healthy and unhealthy entries to construct a "Food Healthiness Questionnaire", which asks to rate the healthiness of the 16 foods (described in plain text) and justify the rating with a written explanation. They then recruited 55 participants from an obesity care program to complete the questionnaire. In a second study (Carels et al., 2007), the same questionnaire was answered by 101

undergraduates. Both studies reported a set of nutritional attributes such as low/high fat and low/high protein as well as perceived quality attribute and portion size attribute. However, these two studies are limited in two ways: first, foods in the questionnaire are simple foods (e.g. an apple) and no multi- food items (e.g. chicken, spinach and tomato salad). Second, foods are described by plain text therefore people are required to imagine those foods while evaluating them. Social food photo applications like the Eatery and Pic Healthy use images and take images not only of single items but seemingly more often of meals, that is, of plates of multiple food, rather than single food items. It is not clear, therefore, either seeing an actual image rather than text, or dealing with multiple food items rather than single items inform health assessments/food choices.

To better understand how people evaluate those arbitrary food and meal, which are presented in images, we adopted a method similar to the one used by Carels et al. (Carels et al., 2006, 2007) to solicit attributes of food people take into account. In our study design, we first randomly selected 15 food pictures from Pic Healthy to construct the survey. The pictures we selected, illustrated in Figure 5.1, include not just simple and raw food (e.g. banana) but also complicated everyday meals (e.g. triple shrimp meal). For each food picture, we use one 7-star Likert scale for rating healthiness of the food and one open question asking people to list attributes of the food to explain the rating. To mitigate order effects, the presentation order of the pictures was randomized. After the rating and explanation task, the survey asks for demographic data: age, country of residence and job.

5.2 Study Results

After we designed the survey, it was deployed on a public free service, SurveyGizmo, and an invitation to complete it was sent out to our university mail lists (accessed by staff and students), health-oriented forums, Twitter, and Facebook. In order to identify where our participants were recruited, we set distinct URL variables to track responses. In two weeks, we received total 153 responses. Participants' age ranged from 18 to 65, with the majority (35.3%) between 26 and 35. Geographic provenance: 45.8% (n=70) of participants live in North America and 39.2% (n=60) in the UK, the rest in European countries. Jobs: 12.4% (n=19) of participants work in the health-related industry, and other participants work in various industries from IT (n=27,17.6%) to education (n=20,13.1%) to sports (n=5, 3.3%). To analyze the results of the one open question on food attributes, we applied an open coding method (Strauss and Corbin, 1990): answers were coded at the sentence level. A total of 26 codes were initially generated, then grouped into the following 8 more general categories: Equivalence Labelling, Brand Association, Nutrient, Portion Size, Quality, Health Effect, Comparison, and Uncertainty. Each category is described in the following subsections.

Equivalence Labelling

Equivalence labelling is a special attribute of food perceived by people: people simply attach a label (e.g. "healthy", "fatty", and "super") to food, nutrition or other attributes without detail explanation. In our analysis, we found several interesting labels such as "honey=sugar", "cheese=fat", and "fruit=super". As expected based on the literature (De Almeida et al., 2001; Eikenberry and Smith, 2004), the label healthy was regularly attached to vegetables and fruits. In addition to above labels attached to one type of food, participants also attached label to certain nutrient they identified within a complex food. For instance Figure 5.1 (A), the three shrimp dish, was rated 6 stars with the simple explanation: "Protein". So in this rating it seems that protein outweighs any other consideration. Therefore, protein is perceived as the equivalence to health for this particular participant.

Brand Association

Participants used the brand information to judge foods healthiness. The Figure 5.1 (C), (E), (H), (I) all contain brand information and participants did use this data as one criterion for assessment. For instance, one answer to the Figure 5.1 (I), a bagged loaf of whole wheat bread, said:

> ...Sara Lee brand = not identified with healthy; cellophane wrapper = not identified with healthy

As shown in this example, this particular answer did highlight the fact that the participant does not trust that Sara Lee could offer healthy food. The trust issue could be further explained by another answer to this bread:

> ...Most foods labelled whole wheat are in fact anything but processed carbs

This answer helps us further understand that people actually doubt whether commercial brands tell the truth on their package. How the future information systems could offer information related to brand and verify certain claims of those commercial food products would be an interesting open question. Interestingly, we also found that people also identify certain packaging material as unhealthy. For instance, the cellophane wrapper, the package of the bread, was identified as "unhealthy". It might suggest that packaging material information could also be leveraged in future designs to help people identify healthy food. Brand Association is an attribute that we did not find in the literature, probably because prior studies used textual descriptions, while in our study food was presented through photos.

Nutrient

As we expected, the most used attribute (category) is Nutrient Five major nutrients are mentioned in all answers: sugar (n=447, 32.67%), fat (n=371, 27.11%), protein (n=287,20.98%), carb (n=208, 15.20%), and calorie (n=55,4.02%). It is interesting that calorie is the least mentioned nutrient in participants' explanation and one possible implication based upon this finding is that calories are not the focus when people think about what is healthy eating. This finding particularly contrasts the current trend in system designs that favor calories counting (Chi et al., 2008; Noronha et al., 2011). Finally, in terms of the amount of nutrient, our findings accords with literatures (Croll et al., 2001; Oakes and Slotterback, 2001; Paquette, 2005) that people pursue low fat, low carb, low sugar but high protein food.

Portion Size

Portion Size is a common attribute identified in previous literature (Eikenberry and Smith, 2004; Falk et al., 2001). In our analysis, we also found that participants mentioned the portion size of certain ingredient/food as one factor influencing their ratings. Usually participants focused on the amount of fruit or vegetables in the food. If the perceived amount was low or there is no fruit or vegetable, then participants gave a negative rating.

Quality

The quality of food relates to two issues: (1) whether the food is natural (not processed) and (2) whether the food is fresh. Generally, our findings resonate with previous literature (Povey et al., 1998; Sage Research Corporation, 2003; Santich, 1994) that people believe natural and fresh food is better than processed food. For example, one participant suggested eating raw grapefruit instead of juice to get fiber in Figure 5.1 (L), a cup of fresh juice, because:

> Freshly squeezed juice is natural, and better than from a carton, although it lacks the fiber you would get from eating the whole fruit.

Our study, however, captured more details in terms of why people believe natural food is better. One major reason is people think that chemical additives and preservatives in processed food are bad:

> Too much confectioned sugar. Too much chemicals and other sweeteners. Chemical flavouring

In addition to the consideration about additives, some participants also raised the issue that how animals or plants grow before they are processed in the factory: "Also the rearing of these animals [shrimp] is mostly bad and they are fed chemicals." These two reasons potentially provide us with more insights into what type of quality information around food, especially processed food, should be given to people to judge food's healthiness.

Health Effect

Health effect is one interesting attribute identified by us. In our analysis, participants reported their concerns around specific health effects of certain foods on human body. For example, one answer to the Figure 5.1 (A), the three shrimp meal, mentioned: "...Shrimp are filter feeders and are thus not especially good for human consumption." Participants also explained the health effect from the perspectives of what nutrition the human body needs: "Fizzy drinks: lots of calories, carbonic acid and sugar to rot your teeth, various additives (like aspartame) of dubious effect on long-term health, and nothing your body actually needs (except water, but it's better to get that from the tap!)" Another effect participants mentioned is the relationships to disease, for example:

> This type of refined carbs leads to heart disease. I wouldn't eat this

This effect was also further explained as the interaction between medicine and certain food in the answer to picture (l):

> I gave it 5 stars, because grapefruit juice has interactions with many prescription meds that could adversely affect your liver. If a person is 100% healthy and med-free, then I would give it 6 stars. Eat the whole fruit for 7 stars.

Based upon this finding, we suggest future designs for health eating should offer personalized information that highlight potential health effect to assist food choices.

Comparison & Relation

Some participants rationalized their ratings by comparison to other foods. For instance, one participant compared the rye bread with wheat bread in the answer to Figure 5.1 (F):

> Rye bread is likely better than a 'wheat' bread as it's made with rye flour and definitely better than white bread

This type of comparison illustrates that people have certain knowledge that which type of food is better than another type of food and could apply this knowledge to their food evaluation. Participants also compared portion sizes of different ingredients in a food picture. One answer to Figure 5.1 (B), a slice of cake, mentioned the small portion size of fruit as a topping cannot overcome big portion size of those less healthy ingredients:

> The small amount of fruit on top isn't enough to overcome the white flour, white sugar and plenty of fat larding this cake.

This interesting point should be paid attention in future designs because relative proportion information will be very useful when people need to consider the basic concept of balanced food and help people to know which part of the food is not good enough and suggest possible solution. For example, if a person want to eat the cake shown in Figure 5.1 (B), the system might suggest that you should also get additional fruit to add vitamin and other good nutrients.

Uncertainty

In some cases, participants mentioned in answers that they wanted to get more details about the food or the person who ate the food in order to make reasonable and fair judgment. The most common uncertainty we found is around ingredients (64% out of all answers were coded as Uncertainty). Even though picture already provides more details but certain information like how one ingredient is processed (e.g. grilled or roasted) or how much oil was used or what type of meat was used are still not possible to be always clear to people. Participants wanted to know those details in order to identify nutrient or attach label (i.e. Equivalence Labelling). For example, one answer to Figure 5.1 (G), rice, one boiled egg, two small tomatoes and 2 meatballs, said:

> I'll give 4 stars... more detail on what's in the 'meatball' could change things...

Interestingly, sometimes participants offered their own assumptions on ingredients to rationalize their rating. For instance, in another answer to Figure 5.1 (G):

> Rice has too high of a GI. Egg is good. Assuming the meatball is just ground beef, then that is good. Tomatoes are great. I give it 3 stars due to all of that white rice.

As we mentioned above, participants also questioned how a food was processed and the focus was on whether any additional flavor or additives has been used. For example, one participant mentioned the processing method: (s)he thought the nuts would be healthy

in the answer to Figure 5.1 (N), a hand full of nuts only if "Unsalted and no oil" Portion size was also at times a point of uncertainty. For example, the answer to Figure 5.1 (J) mentioned:

> Yes, however healthiness depends on the fattiness of the cold cuts and the proportion of cold cuts in the dish.

Another type of uncertainty is about the person who ate the food. Such type of uncertainty raised questions like: what type of lifestyle the person has, what type of physical activity the person did, what type of diet the person is on, what else food the person ate throughout the day and etc. For instance, one answer to the Figure 5.1 (B), a cake with fruit on top, mentioned:

> It depends on what else you are eating, and your lifestyle. You don't need much of this type of food if you [are] sedentary. For your average person, it would be unhealthy to eat too much of this food. It is likely to be high in fat, and sugar; despite the fruit on top.

As this example shows, whether the cake is healthy or not is very context-dependent. It is different from other answers that did not consider the person who ate the food because it highlights the concept of healthy eating/diet/meal instead of the concept of healthy food, which is the mainstream concept encapsulated in most current systems. Therefore, we propose that future designs could think beyond healthy food to think about the big concept of healthy eating by capturing and modeling those rich contexts about the person who eats foods.

5.2.1 The Relation between Attributes Usages, Ratings and Demographics

For ratings of foods, we investigated three statistical variables: average, standard deviation as a measure of agreement and distribution of each food's ratings. The average rating indicates the aggregated numerical result of how healthy a food is. We define foods that are rated greater than 4 stars as healthy, and all the other as unhealthy. Then as shown in Figure 5.1, foods (A, B, C, D, E, H, I, K) are unhealthy foods and the rest are healthy foods. This result basically accords with common sense that packaged or fast food are bad and foods containing more fruits/vegetables are good. For complicated foods (A, D, G,J, O), we expected that people would find it overall difficult to distinguish them and rate all of them around 4 stars (i.e. neutral). Instead, the ratings for these 5 foods clearly indicate healthiness. Thus, we think this result demonstrate our participants are able to apply their knowledge to evaluate those complex foods.

The standard deviation and distribution indicate the variation of ratings of each food pictures. As shown in Figure 5.1, it is not surprising that the canned drink (Figure 5.1 C), the cake (Figure 5.1 B) and vegetable salad (Figure 5.1 O) have relative low standard deviation compared to other foods. This is because they are very common foods and most people know how healthy they are from mass media or public education. But it is very surprising that the banana, a fruit that is commonly viewed as healthy food, has a relatively high standard deviation. By looking at detail explanations contributed by participants, 82.4% people, who rated it lower than 4 stars, mentioned Uncertainty attribute and raised the issue that eating banana only will be bad for health and it should be served with other foods containing rich protein. From this we confirmed that context is important when people are evaluating food's healthiness.

Next, we investigated the relation between ratings and demographics and looked at how attributes we identified could explain those relations. Based upon where participants were recruited, we categorized participants into two groups: those coming from health-related forums and those coming from elsewhere (university mail lists, Twitter and Facebook). The reason to categorize people into these two groups is because people who are in health-related forums are relatively active in thinking, talking and practicing healthy eating. They might have deeper food knowledge than others, and their judgment may be informed by specific nutritional theories or diets. Our goal in making this distinction between groups was to see if and how far apart ratings and rationales of healthiness in these two groups might be. And the results might help us to understand how to better mediate possible communications between two groups to discuss healthy eating and what type of information should be delivered to which group. In total 93 participants were from health-related forums: we refer to these as health community group (H). The remaining 60 people from other places: we refer to these as lay population group (L).

5.2.1.1 Differences in Ratings between Two Groups

A chi-test revealed that the ratings of the foods in Figure 5.1oats with pecans and blueberries. As shown in Figure 5.2, the L group rated these four foods (E, H, I, K) on average higher than the H group did. The reason behind such difference, as indicated by attributes analysis, is the H group people are more sensitive to natural food. All these four foods are commercial food (three of them have the clear brand label in picture, however the Figure 5.1 (K)'s brand information is shown in the text description as honey bunches of oat) and the H group people expressed their strong concerns about added sugar and additives and doubted whether those foods contain real natural ingredients (e.g. fruit in fruit yoghurt) as claimed on package. In contrast, only 5.2% participants from L group had same consideration and only took into account macro nutrients to assess the healthiness.It is worthy to note that the chi-test on the ratings of all other foods between the two user groups did not reveal any other significant differences.

	(E)	(H)	(I)	(K)
Average(H)	1.4	2.63	2.54	2.7
Average(L)	2.72	4.18	4.22	4.5
Std Deviation(H)	0.57	1.42	1.36	1.23
Std Deviation(L)	1.4	1.46	1.33	1.3

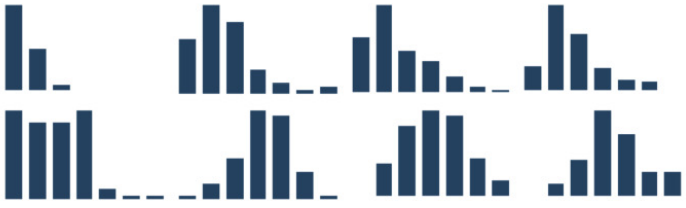

Figure 5.2: The Four Food Were Rated Significantly Different by Two Groups

5.2.1.2 Differences in Attributes Usages between Two Groups

In addition to investigating differences in rating, we also looked at how these two groups used those attributes differently. First, we investigated the two groups' different interests in nutrients. The H group showed interest in sugar (n=312, 34.32%) then protein (n=208, 22.88%) and then fat (n=179, 19.69%). In contrast, the L group is more concerned about fat (n=192, 33.10%) then sugar (n=135, 23.28%) and then protein (n=79,13.62%). By conducting a t-test, we found that the difference in mentioning fat is significant between these two groups (t(28)=2.40, p¡0.05, Cohen's d=0.88). The presence or absence of fat for the L group was a very important feature. For instance, when judgingFigure 5.1 (L), a cup of fresh juice, in one answer from L group, the low amount of fat was highlighted: "Contains lots of minerals and vitamins while low in fat so healthy." In fact, it is not surprising that the L group is more concerned with fat than the H group based on previous research (Carels et al., 2006, 2007). In addition, low fat followed by low salt are currently the most popular features advertised in stores in relation to healthy food (e.g. low or no fat versions of classic products). Likewise cutting fat is perhaps the most often-repeated strategy to lose weight, a popular reason for people to be interested in nutrition (Ricciardelli and McCabe, 2001).

In terms of fat, we also observed differing degrees of nutrient sophistication. The H group mentioned saturated or unsaturated fat 79 times (44.1% of all answers mentioning "fat") and mentioned complete protein 20 times (9.6% of all answers mentioning "protein"). In contrast, the L group only mentioned the type of fat 21 times (44.1% of all answers mentioning "fat") and the complete protein 4 times (5.1% of all answers mentioning protein). This result indicates that L groups knowledge on fat is limited: not a full understanding of the benefits of healthy fat (i.e. unsaturated fat). Finally, we also noticed that the H group mentioned carbs much more frequently (18.7% of this group

responses) than the L group (6.55% of this group responses). Therefore, we suggest for people in H group, the carb is the new evil nutrient and this result should be taken into account for future design to help this group of people consume healthy carbs or avoid carbs. Our next analysis was about the ways in which people from the two groups use the eight attributes: individually or in combination. It is perhaps not surprising that the H group has a more nuanced view of food choices than the L group, reflected in their greater use of multiple attributes to explain a rating, while the L group tends to use a single factor in assessing a food picture. One-attribute explanations featured in 51.2% of the H group responses, but in 68% of the L group.. Based upon our observation, we suggest that future design should offer more guidance in judging food's healthiness. Potential method could be sharing thoughts from more expert groups with the general population in order to help users understand what attributes they might be missing. This is indeed inline with the method proposed by Mamykina et al. (Mamykina et al., 2008).

5.3 Follow-up Study

In the main study, we found limited agreement in participants' views of a food's healthiness. It seems that it is hard to produce one correct answer from crowd's input but is it possible to produce such a correct answer based upon a group of experts? Our anticipation here is if we could produce such a correct answer for food's healthiness, then we could create a simple application that might be to show participants how close or far their assessments would be from knowledgeable experts. In other words we assumed that there would be a consensus around experts where there was not with non-experts. However, as we will report in this section, we were wrong. The following results shows the kinds of divergence we see in expert rankings, even among those with the same certifications or training.

5.3.1 Participants

We recruited expert volunteers by sending emails to 70 registered dietitians from freelance dietitian website [1] and 22 coaches who are listed on Precision Nutrition website [2] and have certified in at least one nutrition program. The reason to recruit two different types of experts is based upon our assumption that PN experts could more easily reach consensus because they are trained by Precision Nutrition so they should hold same or similar food beliefs. In the end, we recruited a total of 14 experts: 8 registered dietitians (RD) and 6 precision nutrition coaches (PN). We then re-ran the same protocol as our main study.

[1] http://www.freelancedietitians.org/. The freelance dietitian is in conjunction with British Dietetic Association.

[2] http://www.precisionnutrition.com/

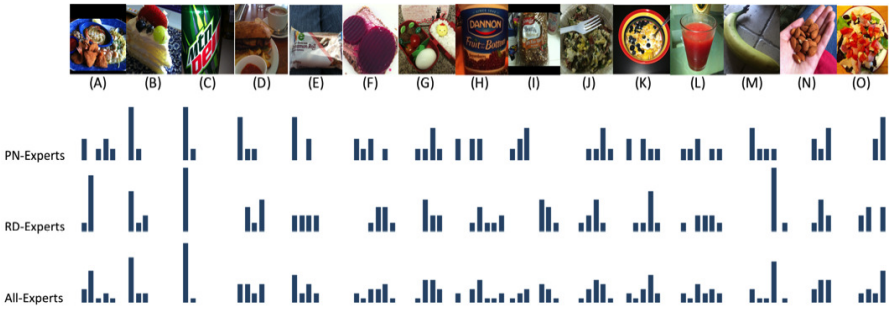

Figure 5.3: The Rating Distributions of 15 foods for follow-up study. The foods are: (A) Three shrimp meal; (B) Cake topped with fruit; (C) A can of carbonated drink; (D) Grilled veggies with pesto sandwich, fries, potato soup, fresh squeezed and strawberry lemonade; (E) Quaker Life soft baked cinnamon bar; (F) Rye bread with beetroot; (G) Meatball, egg, rice and tomato; (H) Fruit Yogurt; (I) Wholewheat with honey bread; (J) Feta cheese, cold cut meats, mushrooms, two eggs and one egg white; (K) Honey bunches of oats with pecans and blueberries; (L) Freshly squeezed grapefruit; (M) Banana; (N) Almonds; (O) Salad. See a larger version in Appendix A

5.3.2 Results

Figure 5.3 shows the rating distribution, and overall we could see experts did not reach consensus on food's healthiness. An ICC test we ran to measure agreement between all experts shows poor agreement (ICC(2,15)=0.0669) both within and between RD and PN expert groups. Looking at the rating distribution of each expert group, we noticed that for food (I), a pack of wholewheat bread, the two expert groups hold contradictory opinions. For example, the PN experts rated the food below 4 indicating they believed the food is unhealthy while RDs thought it is healthy. The open text data revealed that registered dietitians believed the food is healthy because they trust that the bread is wholewheat bread as said on the package and therefore contains high fibre. On the other hand, only one precision nutrition expert mentioned the wholewheat but at the same time acknowledged that the bread is processed. Most precision nutrition experts emphasized on that the bread is highly processed and it is a mass produced, packed food. Therefore the wholewheat bread was not as healthy as it sounds and it should be considered as unhealthy.

The experts also did not reach agreement within each of the two individual groups. The results of ICC for RD and PN coaches were 0.0351 and 0.0726, respectively. For example for food (L), labelled a cup of fresh juice, we noticed that the ratings from registered dietitians varied from 2 to 7. We also noticed that even when two registered dietitians both agreed on that the food is healthy but whole fruit could be better, these two still gave different ratings (4 vs. 6).

In summary, the disagreement on ratings by experts suggests that there is possibly no simple ground truth for food's healthiness rating. Different experts focus on different sets

of attributes of food and so weight these differently. Furthermore, people's perceptions of food such as healthiness is affected whether there is real fruit in packaged yogurt (see food (H) in Figure 5.3) or not, or healthiness may depend on goals and timing: eating a high sugar fruit like a banana may depend on whether someone trying to lose weight has just worked out or not.

5.4 Discussion

The aim of this study is to understand how people apply their nutrition knowledge when they assess food. It is important to understand this because unless we could know exactly what kinds of attributes of food people concern and use while assessing food's healthiness, we can tell whether the popular food's healthiness measurements such as calorie counting can truly reflects how healthy food is and help people make right decision. Moreover, those attributes will reflect how people think about food's healthiness and that could be the basis to guide design for dietary-related application to provide right people with right information.

With our survey study, we were able to solicit attributes of food that participants used. In the process of analysis, we avoided making judgments about the correctness of any assessment contributed by participants; instead, we focused on identifying attributes of food people wrote down to explain their judgments. Those attributes of food is thus the details thoughts on food's healthiness people hold and apply when they are assessing food.

Of the 8 attributes we reported, 3 (*Nutrient*, *Portion Size*, and *Quality*) are the same as those identified by previous literature (see details in Chapter 2) (Santich, 1994; Sage Research Corporation, 2003; Edwards and Hartwell, 2002). Most of our findings around these three attributes are in accordance with previous research. In relation to *Nutrient* in our analysis, we found that calories are not the focus in participant's answers. This also confirms previous findings in health field: calories are just one aspect of food's healthiness. In terms of *Quality*, we noticed two concerns people reported in relation to quality: how vegetables, fruits and animals are grown and how much additives and processed ingredients are added.

Since we used complicated food items and presented foods in image, we observed people talked about brand and compared the proportion of different ingredients in food items. The two attributes: *Brand Association* and *Comparison*, therefore, are new to HCI and Health filed. In addition, the usage of *Equivalence Labelling* poses the question that how future designs might leverage those "labels" (e.g. honey=sugar) to adjust or inform people's healthy eating understanding. Finally, we also reported the *Uncertainty* and *Health Effect* attributes, which to our knowledge were not reported in previous literature about interpretation of healthy eating. The findings raise the question that

how future systems could concern more about contexts related to the individual (e.g. health condition).

Considering how people use those attributes, we found that our participants used those eight attributes in combinations. This fact suggests that the concept of healthy food is not one-dimensional as defined by many current system designs (Chi et al., 2008; Mankoff et al., 2002; Noronha et al., 2011) but multi-dimensional, which means healthy food is not just about moderate calorie value but also about good amount of macro nutrients, quality of food, portion size, brand and health effect. Moreover, we found that participants referred also to issues related to the person who eats the food. This suggests that in system design we need to think beyond food and to consider the more general context around the user.

The nature of food's healthiness is multi-dimensional and context-based and it is reflected in the ratings given by people: people's ratings on food can not reach consensus. In the main study, we recruited two types of users: one is general population from our university, and the other is people who subscribe to a specific diet method by "Precision Nutrition". Our expectation was that there should be difference between two groups since the PN group are trained to hold specific set of food beliefs so they possibly judge food differently from general population. On the other hand, we did not expect there was disagreement within the PN group since they are same type of dieter with similar minds. The result surprised us that even people sharing similar minds still disagree each other and that led us to think that the simple strategy to group people based upon what type of dieter they are (i.e. Precision Nutrition dieter) or based upon other demographic profile does not work well to really cluster people into like-minded group that could reach consensus.

The findings from main study were later confirmed by the follow-up study with experts. In the follow-up study, we found that even experts who have advanced knowledge in nutrition and foods can not reach consensus on how healthy the food is. The disagreement was observed not only between two groups of experts (dietitians v.s. PN coaches) but also within each group of expert. Such a broad divergence among and within experts lead us to think there is possibly no simple correct answer for food's healthiness. Perhaps the best way we can inform people whether the food is healthy is to show them where their assessments of food sit in terms of beliefs around what makes a food healthy or not. In other words, we think the definition of food's healthiness varies for different people and different context, then the best way to tell the person whether the food is healthy or not possibly is trying to show the person how his/her assessment of this particular food's healthiness is compared with other people's assessments. Note here, other people should mean the group of people who share similar food beliefs with the person and also possibly also share the similar context in which they eat the food. The comparison of assessments will let people know how well they understand the food and what they

ignore in their own assessment so they could find out what they should do to improve own knowledge and improve own decision.

Reflecting upon current design of dietary-related applications, we think we need to acknowledge that food's healthiness is always "it depends" and we believe it opens up new design space for dietary-related applications. In fact, as we discuss above, the healthiness of food depends on two things: the persons' specific understanding of food's healthiness and the specific context in which the person eats the food. Specifically, the persons' specific understanding of food's healthiness includes the attributes people concern when assessing food and also the weight people assigns to the particular attributes. In following sections we explore how future design might leverage the attributes we identified to support "it depends" design for dietary-related applications.

5.4.1 Negotiable System for Dietary Change

Current interactive systems to support healthy eating define healthy food mainly based upon calories counting. Most systems for food logging (Noronha et al., 2011), shopping (Mankoff et al., 2002) and preparation (Chi et al., 2008) try to promote healthy eating by recommending people to reduce calories intake. However, healthy eating is more than just calories. Meeting the suggested daily calories amount does not really mean a person consumes enough fruit/vegetables or has balanced nutrition intake or gets enough energy to recover from heavy physical activity. Indeed, as our results show, people have their own understanding of healthy eating: it is a multi-dimensional and context-dependent concept. Thus the healthy eating is always "it depends". We suggest that future system designs should rethink the role technology could play in the process of dietary change. It is not anymore the authority that people should strictly follow and listen to; instead, it should be the coach who is willing to negotiate with people on what is healthy food/eating to reach a more rich and reasonable definition that is suitable for individual case.

To design a negotiable system is not easy, especially when people hold different opinions on what is healthy eating. Take what is the good quality of food as one example: some people think commercial foods are all evil and some people think some types of packaging materials just mean unhealthy. The negotiation will be on what is healthy eating and that means two things: first, what type of food should the person eat and second, what the detail dietary change actions the person should do to eat healthy.

Regarding what type of food people eat, the system could possibly begin negotiation with people by discussing whether the person accepts some basic general rules that are widely accepted as elements of healthy food. For instance, fruit and vegetables are usually considered as healthy food and recommended by USDA that people should usually include them in the meals. And the focus of negotiation could be on how much

portions of the fruit and vegetables should be eaten and how frequently they should be included in the meal.The system could try to offer people the recommended portion size and frequency and ask people whether they would like to adjust the settings to help people easily adopt new habits to eat more fruits and vegetables.

One challenging question should be asked in the design for negotiable system: should the system always respect people's thoughts and comprise the planned interventions to help people eat healthy? One possible issue behind the question is the progress of changing people could feel and see. We know the ultimate goal of dietary-related application is always to help people finally eat healthy and thus the progress does matter. People have their high expectation to achieve their own dietary goals when searching for a good interventions. And especially for technology-based interventions, people might have even higher expectations to achieve it fast. Though we know the fact that the healthy eating is not something could be achieved quickly, how to let people feel comfortable during the process and still feel happy about their progress is the problem here. One possible strategy to address the problem is that the system should let people know the possible consequences if the users choose to adjust the planned interventions. For instance, if the person thinks "5 portions of fruits and vegetables everyday" is not a easy goal to achieve then possibly the person could negotiate with the system to start eating fruits and vegetables for one or two portion everyday. The system should show the person that the adjustment will double the time he/she takes to achieve the final goal.

There is also another issue behind the above challenging question: how much should the system compromise in the negotiation? When a person want to negotiate with the system to adjust the plan, how the system decides whether to compromise or not? One possible way to solve this issue is to ask people justify their decision by selecting reasons from a prepared list. This list could be based upon a set of known barriers in dietary change such as lack of money and shortage of time (see Chapter 2 for details). If the system tries to encourage people eating more fruits and vegetables everyday but the person faces financial problem thus can not afford the fruits and vegetables everyday, then the system should accept this excuse and re-sets the goal for people as just eating fruits and vegetables every week. Note here, since the financial problem is unlikely solved by the person self so here the system re-set the goal completely and just asks the person aim for an easier goal. On the contrary, if the person does not face such an unsolvable barrier but instead having other excuses for an easier plan, then the system could consider compromising a bit first but later reminding people of aiming for higher standards to achieve the final goal (i.e. eating fruits and vegetables everyday).

The new negotiable system for dietary change will also pose new question on how can we evaluate such systems. We think there are at least two aspects of the negotiation should be evaluated:

First, we concern the user's experience of the negotiation. That is mainly related to the user's subjective thoughts and feeling about the whole negotiation process. The possible evaluation method is to interview users after they use the system for a while to collect their feedbacks on how well do they think the negotiation with the system works and what problems they encountered. The feedbacks from users could help researchers to get qualitative information on how well the negotiation works.

Second, we concern the outcome of the negotiation. In the above discussion, we had the example that the system might re-set its goal based upon the users excuse. The new goal then is one type of the outcome of the negotiation process. Then the challenging here is how can we evaluate whether the outcome is appropriate? Three different aspects might be taken into account: first, whether the user think the new goal is more appropriate than the original goal predefined by the system? second, whether experts such as dietitian think the new goal is appropriate?

5.4.2 Support Reflection on Healthy Eating Understanding and Decision-Making

As evident from our data, people sometimes hold wrong or bias beliefs about healthy eating. The use of equivalence labelling, as one example, shows that people sometimes simply attach labels to certain foods without considering the larger context, such as whether the food is balanced by other foods or by physical activity. Therefore, it is important that systems could help people to improve their understandings of healthy eating by creating space and time for their reflection. As argued by (Purpura et al., 2011), current systems offer too much automation supports in behavior change and it might be bad because people have no opportunity to actually think about whether their thoughts and actions are right. Not only we agree with (Purpura et al., 2011), but we further argue that systems for healthy eating should offer the opportunities for people to actively express their own thinking about what is healthy eating and create a space for people to discuss and improve their understandings.

Our results suggest that future systems should attempt not just to capture abstracted ratings, but also to capture the rich thoughts behind them, so those information could be used to create a space for people to reflect upon own understandings. Such space is important for people because In Chapter 3, we already reported that people do seek for a space to further verify what they understand in order to improve own knowledge and do better in the future. And we believe allowing people further exchanging what attributes of food make it healthy or unhealthy and how important each of them is could help people get better understandings in nutrition and improve their potential food decision making ability. For example, a person who believes all fats are bad might benefit from seeing the argument that unsaturated fat are good for health.

To create such a space, there are two design challenges should be explored and solved. First, designers should think about how to design lightweight user interfaces to capture those thoughts behind ratings? More specifically, we here concern two things. The first thing is about lightweight UI, which means the interfaces should not cost user's too much efforts (both time effort and cognitive effort) but still be capable of capturing enough information to represent people's thoughts behind ratings. Regarding the thoughts behind ratings, it specially refers to both the attributes of foods and the weight of each attribute. The 8 categories of food attributes we reported in this chapter thus could be helpful for future design to frame the user interface to capture those possible attributes of food. In chapter 6, well report our attempt in designing such UIs. But how to capture the weight of attributes still remains a question.

The second design challenge is related to how can those thoughts behind ratings could be visualized so end-users could learn new knowledge from them and reflect upon own understanding? Since people hold different views on foods healthiness then how the visualization could break down all data it collects from crowds to help people make sense of which part of understanding is right and which part they have problem.

Furthermore, we also think that such a space for reflection could also help people reconsider how well own choice of food is and help them refine the decision based upon the new inputs from crowds. It is quite different from the current practices where applications usually offer one correct answers for food's healthiness and let people make decision directly based upon it. We believe this is the right move since food's healthiness is always "it depends" and we should acknowledge that one correct answer can not work well to help people make right decision. This is because that the correct answers defined or generated by applications usually are based upon authority such as USDA but can not acknowledge individual uniqueness in dietary change. Although those so-called correct answers are considered more scientific, the crowdsourced assessments could convey much richer information and acknowledge possible context issues.

The support of decision making raises another challenge for visualization of those information. Since people need those information in real time to support their decision either buying food in supermarket or ordering food in restaurant, the visualization should be designed in the way where people could quickly get the most relevant and important piece of information to make the decision quickly.

5.4.3 In Situ Information Support: Shopping and Preparation

Finally, we discuss how the eight categories of attributes we found might be leveraged to offer in-situ information support. Two application areas are discussed here: food shopping and food preparation.

In the related work, we mentioned the Mankoff's system (Mankoff et al., 2002) could scan grocery receipt to offer suggestion on what foods should be bought to get a balanced meal. There are also applications like fooducate [3] could scan food product's barcode to offer alternatives information and details nutrition information. Based upon our analysis, systems for food shopping should also consider other information.

One type of information is whether a brand could be trusted (see Brand Association above). Systems should try to look at how people determine a brand is trustworthy and collect information about those brand or product characteristics. For example, as we mentioned in result, some people think the cellophane wrapper is one unhealthy sign. Thus, if the food product is also available in another type of package (e.g. paper wrapper) then it might let those people perceive the product much healthier. Moreover, the brand's product history might also matter. Sara Lee, for instance, is one company usually produces sweet cakes. Thus it leaves people who know the brand a bad impression that this company is not able to product healthy food. For people who do not know the brand, a list of Sara Lee products or a summary review of previous products will be helpful for people to determine whether to choose the food or not.Finally, it will be good to offer people information regarding context. If a person just comes back from gym to the market for snack, it will be good to offer suggestions on how to pick up food that could help the person recover from heavy workout.

The in-situ information support also could help people prepare food. For example, in (Chi et al., 2008)'s smart kitchen, the system offers calorie estimation for ingredients in order to help people be aware of calories intake and swap some ingredients. Based upon our results, calories are not the focus. But the portion size might be one important thing people care during cooking. It will be good to guide people how to cut meat in a proper size or how much oil should be used. In addition, such smart kitchen also could suggest possible alternatives for user during cooking. For example, olive oil is better than vegetable oil because it provides unsaturated fat. If the system could detect whether a person has some olive oil at home and offer suggestion on swapping vegetable oil for olive oil. Then the person could cook the food in a much healthier way.

5.5 Summary

In this chapter we have explored what attributes people used to interpret whether a food presented in picture and text is healthy or unhealthy. The key outcomes of the this study are:

- Evidence shows participants do have and apply (varying degrees of) food knowledge to make judgments about food healthiness.

[3] http://www.fooducate.com

- We identified 8 attributes in assessing food healthiness; these attributes are used singly or in groups when making assessments.
- Evidence shows even experts do not reach consensus when judging food's healthiness
- Evidence shows calories are not the focus and we propose future design should shift focus from calorie counting to multi-dimensional healthy eating assessment.
- Evidence shows context is one key to healthy eating and we propose future design should try to leverage context information about a person in order to offer reasonable suggestions.

Based on these findings, we reflected on current system designs and explored three future design opportunities: negotiable definition of healthy eating, supporting reflection on healthy eating understanding and in situ information support. These three opportunities all leverage those eight attributes we identified and could help future systems to assist people in making real dietary change throughout all stages of eating, from shopping to preparation to eating food.

Chapter 6

Capturing and Visualizing What People think about Food's Healthiness

In the last chapter, we drilled into the detail about what kinds of food-related attributes people take into account when assessing food's healthiness. The 8 attributes we identified led us to reflect on whether the single aggregated rating, which is used in the crowdsourcing application such as the Pic Healthy and the Eatery, still fairly reflects the healthiness of the food. In most cases, we as designers tend to have the expectation that such crowdsourced, single answer can be successful because we can deliver an answer to people wanting to make a decision about something (Consolvo et al.; Noronha et al., 2011). if one wants to know how many calories are in the pizza, then the application tells the user whether the crowd thinks that piece of pizza is healthy or not so the user knows whether to eat it or not (e.g. The Eatery or Pic Healthy) - or to feel guilt if we already have.

However, as the evidence we collected through previous studies, we know that not all questions may have such reducibly factual answers. The Chapter 5 reported that people did not reach consensus on the healthiness of food and pay attention to different kinds of attributes and in particular, they showed their interest in the food context: the conditions under which one consumed a food also informed whether it could be construed as healthy or not. That led us to think that how to design applications to support more nuanced aspects of health, where the closest answer to a ground truth is always "it depends" - no matter how large the crowd that's sourced and whether they are experts or not.

Taking context into account presents a conundrum for proactive health interaction design: we are used to getting fast answers in our pockets to health queries; we are often lead to expect there is a *correct answer* to simple questions like "is this sandwich

healthy?". What we have found in our work is that how we can answer that "truthfully" depends on taking into account the questioner's context, where context includes not only physical or social context, but one's beliefs and biases, too. A need to take into account such varying contexts may suggest that food healthiness is relative. Could we really be saying there is no such thing as agreeing say, a vegetable is healthy? Apparently not: we have found well-followed scientists (not necessarily food scientists) who argue strongly that vegetables are problematic; legumes are toxic and that sugar and orange juice in abundance are an excellent way to diet, lose weight and be healthy.

A design challenge of interest is then what to reflect back to a user about their assessment/understanding of a food? For instance, one may be aligned to a particular nutritional belief system and have no idea that either that is the case, or that choices are being made from that frame or that other views are also possible. It may be that there is value in being able to show someone how their choices align to a particular system or systems, and also provide alternative views to help them assess/explore their own understanding of a food. Our goal in the work we report here is to explore the first stages of this process: to assess possible interface designs that let us capture, in the most efficient and effective way possible (irritating the user the least) a person's rationale for a particular food assessment. We assume that if we are capable of capture such rationale then we are able to possibly model people's food beliefs and biases and then we could leverage such model to offer better food related information service.

In the following sections, we present our evaluations of three interface types for capturing food choice rationale: open text box, predefind list and heuristic checks. We assessed these against three criteria (1) efficiency: time people spent on the UI; (2) data quality: the amount and quality of food attributes (i.e. the rationale) each user interfaces could capture; and (3) potential for presentation of the collected data: how the way the data is collected through each interface may influence people's judgement of food's healthiness. We found that each UI affords a surprisingly distinct set of non-overlapping qualities. We show how these attributes may be used deliberately in design as a foundation for the next step: helping people see their current food belief biases. Finally, we present implications for future healthy eating application designs.

6.1 Study: User Interface Design for Capturing Rationale behind Ratings

In this section, we first introduce our user interface design to capture rationale behind ratings. Then, we report four experiments we carried out to compare three user interfaces from three perspectives: efficiency, data quality and potential for presentation of the collected data.

Figure 6.1: The Screenshot of Predefined List UI

6.1.1 UI Design: Data Capture

The challenge of data capture is not new and the tradeoffs are familiar with usual types of data capture interfaces. One may be presented with an open text box or some kind of ordered list or rating scale. Each has costs in terms of sufficiency of data captured and efficacy for the user of the capture process. The structure of the UI may influence or restrict what people may say. Such "structure" is a concept borrowed from questionnaire design (Peterson, 1999), which makes clear, as is well understood in HCI, that the response type influences how freely people are able to express their thoughts as well as how easy it is for people to complete/use a system. Take the open text box as an example: it is unstructured. People can input any data in any order. People are familiar with the open text box so it is easy to use, but also has a cognitive cost in determining what and how much to say. The quality of the data generated by open text box, therefore, is dependent on users' knowledge and motivation to contribute the data. Especially, in a mobile context, people are reluctant to put too much effort on answering open text questions so the quality of data is in doubt. The predefined list on the other hand offers checkboxes that reduces effort to come up with terms (Sinclair, 1984), but may also nudge someone to prioritize something that they would not have considered important if it had not been explicitly in front of them.

Because our goal is to best understand ways to expose judgment attributes we want to compare the most structured - lists - with the least - open text. Our lists are informed by our previous work on attributes of importance (see Chapter 5 for details) including both

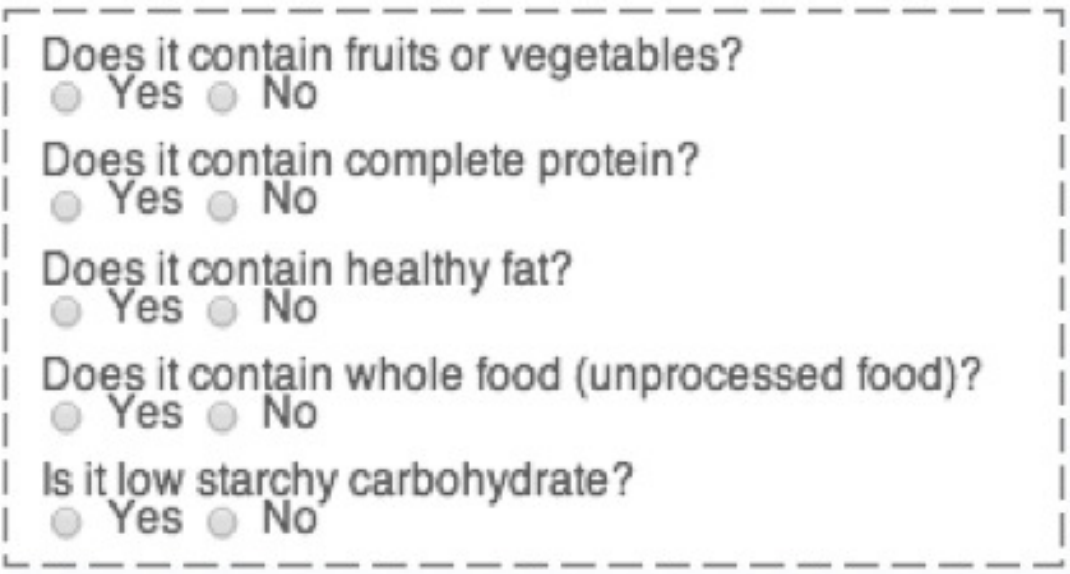

Figure 6.2: The Screenshot of Heuristic Check UI

positive and negative attributes like good brand, fresh, bad portion size, no fruit-&-veg and high protein (see Figure 6.1 for details).

Beyond attribute lists of nutrients and open text boxes, we assessed a third type of structured query: what we call heuristic checks. It is designed as five heuristic questions that ask people to check whether each heuristic applies to the food and answer yes or no. The questions are based on the nutrition program Precision Nutrition[1] developed by Berardi and colleagues specifically (1) to simplify food assessments/decision making and (2) to take into account context in which food is eaten. For example the heuristics include: each "feeding opportunity" (no distinguishing between a meal like lunch and a snack) should contain a complete protein. Each time one eats, also include a fruit or vegetable. Each day have sufficient healthy fats. Each time one eats, have a whole food (eliminate processed foods).

Ideally, the heuristic UI (see Figure 6.2 for details) should be the easiest and fastest to use interface because there are only five questions; on the other hand, it presumes a set of values about food that may not be shared by all, for instance starchy carbohydrates like potatoes, grains or sugars only after workouts if one is trying to lose weight. Our goal, however, in exploring this kind of compact interaction was less to postulate specific hypotheses about "best" UI, than to explore how people with a range of different perspectives about food might respond to such a specific interaction, and to learn whether its constraints/perspectives still delivered something useful in its efficiency or was over constrained.

[1] http://www.precisionnutrition.com/

6.1.2 Methods

We conducted four experiments to evaluate the interfaces against three criteria:

1. Efficiency: how much time people spend on each UI inputting the data
2. Data quality: the amount and quality of food attributes (i.e. the rationale) each user interfaces could capture; and
3. Data presentation: the influence of crowd wisdom, which is presented as visualization of data captured, on peoples judgement of food healthiness.

The goals of each experiment are as follows:

1. Experiment 1: How Does UI Type Impact Time on Task and Data Quality?
2. Experiment 2: Does Priming With Collective Crowd Wisdom Influence Ratings?
3. Experiment 3: Does Reflecting On Collective Crowd Wisdom After Rating Change Ones Decisions?
4. Experiment 4: Do 7-Point Ratings Differ If Provided Nutrition Label Instead of Visualization?

6.1.3 Experiment 1: How Does UI Type Impact Time on Task and Data Quality?

Our first experiment is to investigate how people use the UI to input data and examine the data quality. To do so, we revisited the same images from our previous study on food attributes: we used the same 15 images that included a mix of single food items and combinations of foods. Each image is also accompanied by a text description of the food. Participants were instructed both to rate the food's healthiness on a 7-star scale and justify the rating by using each of the three user interfaces, in turn. To test the different effect of each user interface, the whole quiz was divided into three sections, each of which contains five foods and one of three user interfaces. To mitigate the order effect, the assignment of foods and user interface into each section was fully randomized. For each food we recorded separately the time participants took to enter the quantitative rating and the rating explanation. In the end of the quiz, we also collected basic demographic data including: gender and age group. The whole experiment was deployed as a website.

In this experiment, we evaluated the interfaces according to the following three ways: time on task, qualitative assessments and relationship between qualitative and numerical assessment of food's healthiness.

1. Time on Task

   We are interested in two types of time costs: (1) time cost of assessment completion and (2) time cost of learning the UI. For the first cost, we set a timer on each food page to record the duration from entering the food page to submitting the assessment. Although this duration is not exactly the same as the time duration people really spent on the user interface to input their qualitative assessments, we here use it as a rough estimator. For the second one, we looked at how significantly the time cost of assessment completion could be reduced between the first task and the last task.

2. Qualitative assessment attributes

   The assessment data is composed of two parts: ratings and attributes generated by each UI. For ratings, we looked at whether there is any significant difference in rating distribution between three user interfaces to understand the effect of user interface on rating. Since we have postulated that there is no single ground truth of rating, we do not attempt to measure the accuracy of ratings. Rather, we tallied how many attributes per food each user interface could generate as one metric to assess the quality of the qualitative assessment. We also looked at how much agreement we could find among participants regarding the individual attributes and what are the most frequent mentioned attributed in different user interfaces to understand how different user interfaces affect users' input of qualitative assessment.

3. Relation between qualitative assessment and numerical rating

   Our last interest in the data is to unpack the connection between attributes (i.e. qualitative assessment) and ratings to inform the design of data visualization and future food-related system design. We looked the correlation between ratings and attributes to understand how people weight each attribute differently in the different user interfaces. We also further looked at how attributes are distributed on ratings of each food to find out whether there is any patterns in distribution that might indicate detail connection between attributes and ratings and help us tease out the "it depends".

6.1.3.1 Participants

We recruited two groups after the school's ethic committee approved the study to compare the results between paid turkers and a group of volunteers similar to the population using Pic healhty and the Eatery. Paid turkers (n=92) were recruited through Amazon Mechanical Turk, and we paid $0.3 per turker for the study. We added one additional question asking the colour of background (that is dynamically set and recorded) of the page to check that respondents were paying attention (answers for which this check

failed were excluded from further analysis). The volunteer group (n=82) was recruited from university by sending emails to mailing lists and volunteers were self-selected to join the study.

6.1.3.2 Results

In this section, we report the result of the experiment. Since most of the two groups responses have no significant difference from each other, we here mainly report the result from volunteer group but mention the difference between groups if needed.

Time on Task Analysis
A one-way ANOVA revealed that total time taken to complete one food assessment was significantly affected by the type of user interface ($F(2, 1242) = 38.7, p < 0.001$). To further unpack the difference in time completion and find out whether such significant difference exists in any pair of interfaces, we ran a follow-up Post-hoc Tukey HSD test. The results showed that differences in task completion time between predefined list (average time = 51.49s) and heuristic checks(28.4s), and between open text box (55.46s) and heuristic checks were significant. No difference was found between open text box and predefined list.

A t-test revealed a significant difference in task completion time between the first and the last time the predefined list was used by each participant ($t(136.5) = 6.3159, p < 0.001$). On average when using the list UI, participants spent 76.81s (SD:38.75) on the first food and 45.24s (SD: 23.90) on the last one. Similarly, for the heuristic checks UI participant spent longer when rating the first food (average: 40.28s, SD: 23.20), than when rating the last food (average: 24.95s, SD: 13.36), with statistical significance (t-test: $t(131) = 5.2152, p < 0.001$). No significant difference was found by the t-test between first and last times the open text box was used.

Analysis of Ratings
Even though an one-way ANOVA test (p=0.73) did not reveal any overall effect of the UI types on the numerical ratings, we observed a specific case that we consider worth reporting as anecdotal: the ratings of food (C), a can of carbonated drink. As visible in Figure 5.3, for this particular food, almost all the participants who rated it through the open text box rated it '1'; in contrast half of the participants who used the predefined list and one third of participants who used the heuristics UI rated it '2'. We think this is an interesting example in that it shows that people when using the predefined list and heuristic UIs may be distracted by features they would not otherwise consider. Perhaps when people are given a list of features to evaluate, they tend to assume that the food should be rated as '1' only when all negative features are selected, so some of them rated the carbonated drink as '2' because it does not, for example, contain fat. In contrast, when using the open text box people perhaps focus on just those attributes that they

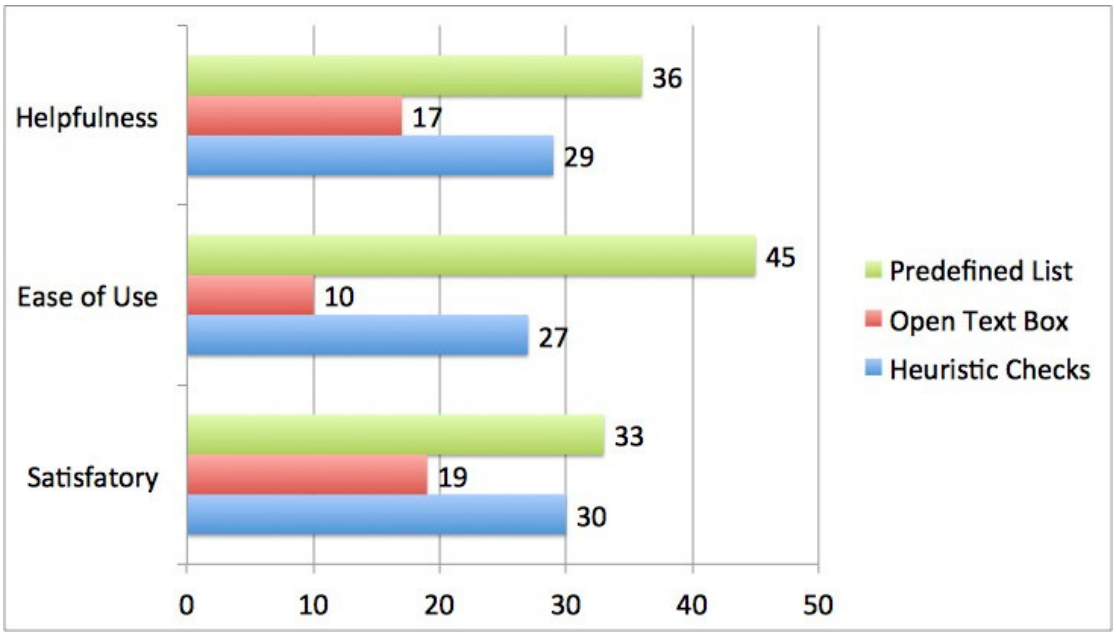

Figure 6.3: Helpfulness, Ease of use and Overall Satisfactory of Three User Interfaces

consider to be most important for the situation and based their judgement only on those. We suggest future research could try and shine more light on this phenomenon.

Subjective Ratings of the Three User Interfaces

The quantitative responses to the post-study questionnaire are reported in Figure 6.3. Specific quotes from the open answers are reported in the discussion section.

Assessment Attributes

We started by analysing how many attributes participants selected on average with each of the three UIs. For the heuristics-based UI the number of attributes is fixed to five by design, as it enforces users to answer all five questions it presents. The average number of attributes generated through the predefined list is 8, SD: 3.27, ranging from min to max and with interquartile range from 5 to 10. The average number of attributes generated through the open text box is 2, SD: 1.04, ranging from min to max and with interquartile range from 1 to 2. The difference between predefined list and open text box UIs was found to be significant through a t-test ($t(500) = 34, p < 0.01$).

Next, we examined which attributes where most frequently mentioned or selected through each UI. The top five attributes for the open text box are: food label (n=123), high sugar (n=92), high fat (n=77), fresh (n=42), and processed (n=35). In contrast, for the predefined list they are: good portion size (n=257), good for health (n=222), fresh (n=179), high sugar (n=178) and low fat (n=163). A Chi-squared test revealed that the differences in the frequency of the attributes 'portion size' and 'health effect' were both significant ($p < 0.01$) between the two UIs. Furthermore through the same test, we also

found significant differences in how frequently 'brand' and 'healthy fat' were mentioned between two interfaces (no other differences were found significant).

Lastly, we examined whether the different interfaces how much *agreement* we could find among participants regarding the individual attributes. For each individual food we calculated the percentage of responses that mentioned each attribute. We performed this analysis without taking into account whether the attributes were mentioned in a positive or negative way. For example, for food (a), three shrimp meal, 12 responses mentioned fruit&vegetable out of the 24 (50%) collected through predefined list; only 2 responses out of the 33 (6.06%) collected through the open text box mentioned the same attribute; all 25 out of the 25 (100%) collected through the heuristic questions mentioned fruit&vegetable. Once again, the heuristic questions constrain users to consider all five features, so the effect of that is constant. For the other two UIs, an one-way ANOVA and a Post hoc Tukey HSD test revealed that the average percentage of responses mentioning the same feature is higher for the predefined list (56.53%, SD: 16%, min: 21%, max: 82.65%) than for the open text box (14.54%, SD: 7.86%, min: 3.13%, max: 31.91%).

Attribute-Rating Correlation

We calculated the correlation between the each of the attributes mentioned or selected in the UIs and the numerical rating of the food healthiness using Kendall's correlation method. For both the predefined list and heuristic UIs all attributes are significantly correlated with ratings ($p < 0.01$), for the open text only 9 out of the 16 attributes were significantly correlated with ratings. For the predefined list the top 5 attributes correlated with the ratings are "health effect" with a correlation coefficient of 0.619, followed by calorie (0.476), fruit&vegetable (0.440), fresh (0.431) and sugar(0.389). In contrast, for the open text box the top 5 strongest correlated attributes are sugar (0.333), fat(0.294), whole food(0.282), calorie (0.204) and fresh(0.174).

For the heuristics the order of the attributes with respect to correlation is: whole food (heuristic 4, with coefficient 0.366), fruits&vegetable (heuristic 1, with coefficient 0.283), healthy fat(heuristic 3, with coefficient 0.282), complete protein(heuristic 2, with coefficient 0.18) and carb after workout (heuristic 5, with coefficient -0.122).

Attribute-Rating Distributions

To further investigate how attributes are linked to ratings, for each individual food and for each rating value (from 1 to 7) we counted how frequently participants who gave a specific rating selected each attribute. Figure 6.6 illustrates this analysis: each row R corresponds to a rating level, each column C to an attribute, the size of each bar indicates how many participants gave the rating R and at selected the attribute C for this food. For example 34 participants rated the food (H) and 28 selected good portion size as an attribute.

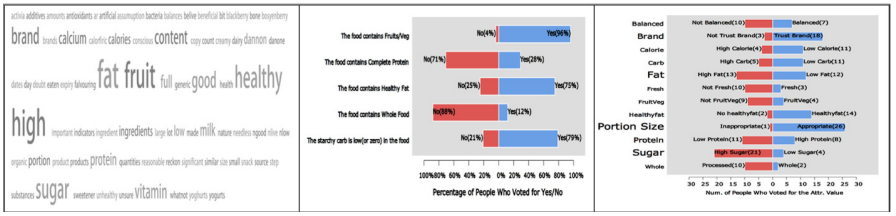

Figure 6.4: From left to right: wordcloud for open text, barchart for heuristic checks, and cloud-barchart for predefined list. See a larger version in the Appendix.

6.1.4 Investigating the Potential for Presentation of the Collected Data

After experiment 1, we decided to further compare our three interfaces by presenting potential end-users the data we collected through three interfaces. This is important because the data we collected need to be useful for end-users and we are interested in how different data collected by different user interfaces have different effect on people's judgment on food's healthiness. We conducted total three experiments: Experiment 2 targets receivers, who wants to know more about the food and did not assess the food before. Experiment 3 targets contributors, who assessed the food before and now return to see the result. And in the Experiment 4, we investigated how the knowledge we captured is different from nutrition labels for branded food.

6.1.5 Visualization Design: Present Food Knowledge

As shown in Figure 6.4, we designed three simple visualizations corresponding to data capture UIs to present those data we collected: wordcloud for open text UI's data, barchart for heuristic checks UI's data, and cloud-barchart for predefined list UI's data. We created the wordcloud for open text data by removing stop-words and punctuation from original text data we collected. Words then are presented in alphabetical order and the size of text is determined by word's frequency. In addition, we display all text in black to prevent people from paying attention to one attribute particularly because of its colour. For heuristic checks UI, it is simple to present the binary result of each heuristic in a barchart. Based upon this, for the predefined list data, we present each predefined attribute in a similar barchart but we also configure the size of the text of each attribute based upon its frequency so people could know which attribute is the most selected one and what's the crowd's judgment on this particular attribute via the barchart.The data we used to create those visualizations were from study described in Chapter 5.

6.1.6 Experiment 2: Does Priming With Collective Crowd Wisdom Influence Ratings?

The food knowledge we captured is in related to one particular food and thus the visualization is expected to help the end-user to tell whether the food is healthy or not. Receivers is a group of people who did not assess the food by themselves but will use the available information in the visualization to judge the healthiness of the food. Our expectation here is by showing the visualization, those receivers will gain better understanding of the food and produce better food's healthiness assessment. To test our hypothesis, we used the same set of study materials: 15 food pictures and text descriptions we used before, and three different type of visualizations we created for each food. A three-section quiz was then created. In each section, we have 5 foods with one type of data capture UI randomly selected and one corresponding type of visualization to present knowledge we capture. Regarding the task participants need to do, we show them food's picture, text description and visualization all together, then ask them to rate the food and give qualitative assessment. We used the same cheating check question as experiment 1 used for turkers we recruited.

6.1.6.1 Participants

Similar to experiment 1, we here also recruited participants from two groups: paid-turkers and volunteer group. We here used paid-turkers to quickly get responses but also used volunteer group to see possible difference between them. The way to recruit participant is the same for experiment 3 and 4. We paid each turker $0.3 for the study and recruited 93 turkers from MTurk and 58 volunteers from university. Note that for this particular experiment, we tried to exclude people who possibly joined in Experiment 1 because they already done the assessment so they are not receivers but contributors. Our method here was to check whether the IP address and email address were already in experiment 1's database, if so we kindly rejected the participants and asked the person to join experiment 3. We here acknowledge the limitation of this method but it seemed it was the best way we could do to exclude previous participants.

6.1.6.2 Results

We observed that among 1395 ratings we collected from turkers, only 454 (32.5%) ratings agree with the ratings in the visualization. A t-test showed no significance is found between ratings from this quiz and ratings from experiment 1. Then we looked at qualitative assessment participant gave and compare the data with data we collected from experiment 1. The Chi test result shows that compared with just showing food picture, showing people the predefined list visualization could help people mention significantly

more "brand" and "healthy fat" attributes ($p < 0.05$). Comparing the qualitative assessment from turkers with assessment from volunteers, we observed that volunteers mentioned "brand" and "healthy fat" significantly more than turkers ($p < 0.05$) in the predefined list condition. We did not observe that visualizations of heuristic checks's data and open text's data have such significant impact on people's judgement.

6.1.7 Experiment 3: Does Reflecting On Collective Crowd Wisdom After Rating Change Ones Decisions?

For contributor (who assessed the food before and now return to see the result), we want to check whether presenting them the information we capture changes their own understandings. The quiz used the same set of material as experiment 2 used and was created in the similar structure. The difference is the task people need to do: for each food, the participant first sees only the food's picture and text description to give rating and explanation. After it, we show the participant the visualization and ask whether the participant "will change" own assessment (the rating and qualitative assessment). If the participant chooses "yes", then he/she could modify it. Our survey recorded the "will change" choice, modified assessment, and original assessment.

6.1.7.1 Participants

We paid $0.3 to complete the quiz and got 92 turkers in the end. We also recruited 52 volunteers with 39 from different health-related forums and 13 from university.

6.1.7.2 Results

We finally collected 1455 ratings from turkers in total and among them, 28% of them originally agree with rating in visualization and 43% of them are close to (i.e. 1 point higher or lower) the rating in visualization. We found that 58 ratings (4% of total) were changed by participants after seeing visualization, with 17 are close to the ratings in visualizations and 30 are equal to the ratings in visualizations. Regarding food attributes, participants made no changes via open text UI and only 4 changes via predefined list UI. Among total 485 heuristic checks UI data, participants made 18 changes with only 4 finally agreeing with the result shown in the visualization. A Chi test showed that heuristic checks does significant better in terms of changing people's mind than other two UIs ($p < 0.01$). We further investigated how many heuristic checks are completely different from result in visualization. We found that only 10 heuristic checks are completely different from results in visualization while 72 originally agree with the result in visualization. 262 (54%) heuristic checks get 1 or 2 heuristics checked different from result in visualization. Regarding the data from volunteers, we see people in this

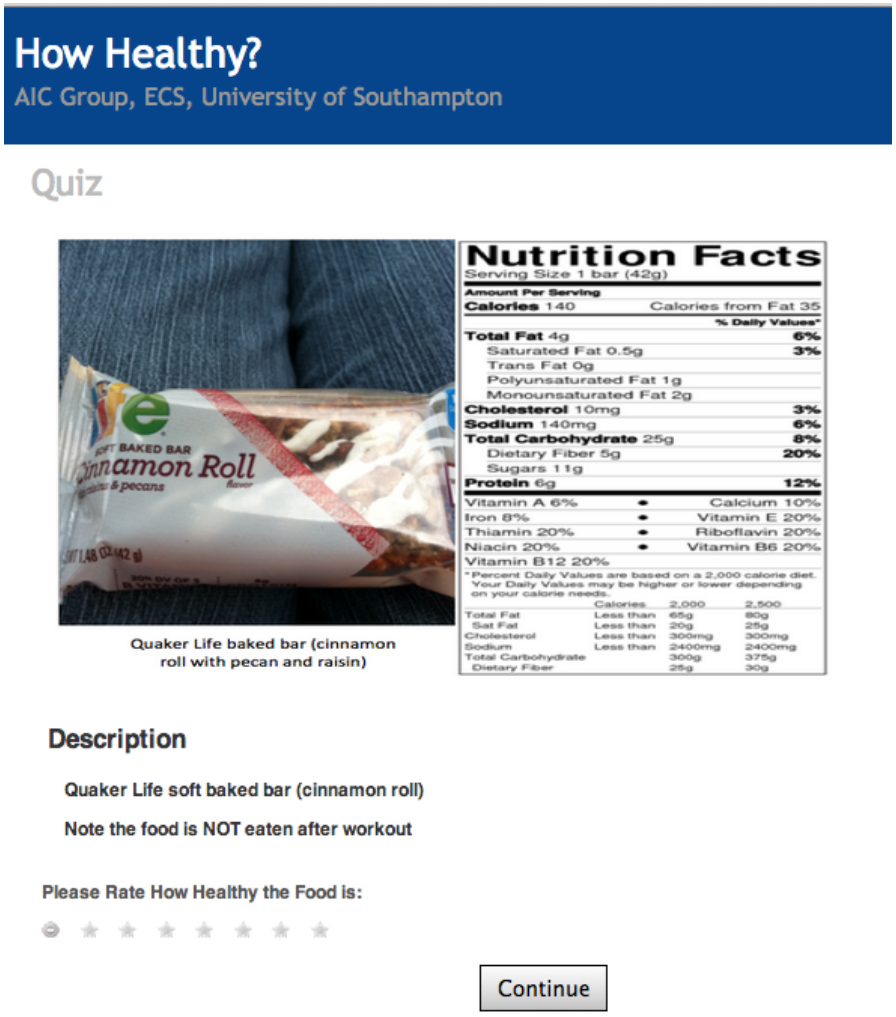

Figure 6.5: The screenshot of the experiment 4- nutrition label condition. The food picture and the nutrition label are shown on the screen for people to rate healthiness.

group rarely made changes and a Chi-squared test showed that it is significance between volunteers and turkers in number of assessment changed.

6.1.8 Experiment 4: Do 7-Point Ratings Differ If Provided Nutrition Label Instead of Visualization?

The last experiment we conducted was to look at difference between showing a nutrition label, and showing the visualization of food knowledge we captured. We used only four branded food from 15 foods (i.e. food (C) , (E), (H) and (I)) here to construct a survey since only those foods have accurate nutrition label available on manufactories' websites. On each foods page (see Figure 6.5 , the participant sees either nutrition label with food information or one random food knowledge visualization from three visualizations we designed with food information. The task is to rate the food on 7-star scale.

Figure 6.6: The Attribute Distribution of Food (H), Fruit Yogurt. The Y-axis is the 7-star scale value from 1 (top) to 7 (bottom), and the X-axis is attribute values of this particular food. For each two columns, they represent a pair of the attribute (e.g. fat) with different two values (lowfat .v.s highfat). For Mark 1, it is the Factual pattern on the portion size attribute with value: badportion and good portion . Mark 2 is the Disagreement pattern on the nofruitveg and fruitveg.See a larger version in the Appendix.

6.1.8.1 Participants

For this experiment, we recruited 47 turkers with $0.1 per person and 38 volunteers from university.

6.1.8.2 Results

A t-test revealed that there is no significant difference in ratings between these two conditions. A further one-way ANOVA test showed that there is also no significant difference between Nutrition label condition, food knowledge visualization condition and control condition (just food picture and description,we used the rating data from experiment 1). The results were confirmed in both participant groups.

6.2 Discussion

In this work, we reported our efforts on designing user interfaces to capture the rationale behind ratings to provide end-users feedback on food's healthiness but also to be a mean to collect data on how people judge food's healthiness so we could extract people's beliefs and biases towards food from it. In following subsections, we first discuss the trade-offs of three user interfaces we proposed. Then we discuss two patterns we found in the data collected through predefined list to guide future research. In the end, we discuss how our research offers the opportunity towards building a data platform for future food-related system.

6.2.1 User Interface Design for Capturing Rationales behind Ratings

From our analysis, participants were able to complete tasks in a shorter time when using the heuristics UI, than when using either of the other two interfaces. The differences in task completion time between the beginning and the end of the study indicate a learning effect for both the list UI and the heuristics UI. Participants probably became familiar with the available attributes and their location on each interface. No learning effect was observed on the open text UI, and while at the beginning of the study the task completion time for the list UI and the text UI are comparable, after the effect of learning the list UI is significantly faster to use than the text one. Another key advantage of the heuristics UI is that it supports easy comparison: since the heuristic questions are fixed for all foods, users can immediately see how the heuristics apply to different foods.

The number of attributes generated by each interface confirms that participants generally took advantage of the greater expressive power of the list UI, using this UI they generated on average 8 attributes per food, in comparison to the fixed 5 attributes generated by the heuristics UI. It might not be surprising that that the text UI led participants to produce significantly fewer attributes than the list UI given that people usually are reluctant to put much efforts on such heavy tasks. However, one thing we should note is that the open text UI demonstrates its power in capturing more diverse opinions and very details about food or particular attribute from participants. For example, for food (I), the packed bread, some participants included information about why they consider the brand as bad, such details obviously cannot be expressed through the list UI and will be pretty useful for other users to know much more about the brand and avoid other products produced by the company in the future.

The analysis of the frequencies of attributes generated by each UI indicates that some attributes ('portion size', 'health effect,' 'brand' and 'healthy fat') were mentioned significantly more frequently through the list UI than through the text UI. The list UI then elicits attributes that people would otherwise mention a lot less frequently using the text UI. Therefore the predefined list has potential to increase how informative the ratings are, to facilitate the end user in thinking out of their usual thinking patterns and to increase people's understanding of food. The larger number of selected attributes probably also explains the differences we found in terms of how widely participants agree on specific attributes expressed through each UI.

At the same time, when comparing the UIs it is important to look beyond quantity of attributes: even though no significant differences were found on the numerical ratings between the different UIs, we observed a specific case that we consider worth reporting as anecdotal: the ratings of food (C), a can of carbonated drink. As visible in Figure 2, for this particular food, almost all the participants who rated it through the open text box rated it '1'; in contrast half of the participants who used the predefined list

and one third of participants who used the heuristics UI rated it '2'. We think this is an interesting example in that it shows that people when using the predefined list and heuristic UIs may be distracted by features they would not otherwise consider. Perhaps when people are given a list of features to evaluate, they tend to assume that the food should be rated as '1' only when all negative features are selected, so some of them rated the carbonated drink as '2' because it does not, for example, contain fat. In contrast, when using the open text box people perhaps focus on just those attributes that they consider to be most important for the situation and based their judgment only on those. We suggest future research could try and shine more light on this phenomenon.

Regarding the effect of knowledge we captured, we found that there is no difference in providing nutrition label and providing food knowledge visualization (Note that only foods having brand were included in this particular experiment). Both of them failed to have significant impact on people's ratings of foods. It is interesting to see this result since we think 1) providing people any extra information (in the experiment, it means the nutrition label or food knowledge), people could make much better judgement and possibly reach consensus and 2) the rich information provided in the visualization could be more helpful than the nutrition label which only included nutrition information. The possible reason that both information did not make effect is that for branded food, people are relatively more familiar with them and already hold strong beliefs in whether they are good and will not change the opinion easily. The result thus further suggests the only Correct Answer for food's healthiness does not exist. Moreover, it also indicates that people are reluctant to change their impression of the food even given additional information to judge the food. This point was also confirmed in experiment 2 and 3.

However, it is interesting that participants were affected by the knowledge we captured when they did the qualitative assessment. More participants in experiment 2 selected particular attributes than in experiment 1 because those attributes were already shown in visualization (e.g. brand in food (I)). So what we observed is people were willing to learn new attributes and used them in their assessment but they were reluctant to re-form their total impression of the food. When considering the "contributor" group, we observed that heuristic checks visualization did significant better in terms of helping people change their mind in assessment. This effect was not shown by the predefined list visualization although it did show significant effect on helping "receiver" notice unusual attributes they might ignore. The result is interesting because we did expect predefined list could also help "contributor" reflect upon their own judgment because of richer information it conveys. The unexpected result implies that we need to find a better strategy to highlight certain information in predefined list visualization to attract a person's attention. We believe one way to explore it is to mine the person's food beliefs and highlight information in the visualization that might be related to those problematic beliefs the person has to foster the reflection. We also suggest that future research could investigate the strategy on how to combine those user interfaces into together to (1)

capture diverse and detail food's assessment data that could be useful for knowledge extraction and (2) be useful for both type of end-users to improve their knowledge on foods.

6.2.2 New Knowledge Captured by Predefined list: Patterns of Attribute Distribution

In our experiment 1 data analysis, we also observed that the predefined list potentially could reveal new knowledge by linking ratings with attributes. The feature of predefined list is not confirmed since we only observed it in the data from volunteer group but we think it is interesting to discuss it here.

For each individual food and for each rating value (from 1 to 7), we counted how frequently participants who gave a specific rating selected each attribute. Figure 6.6 illustrates this: each row R corresponds to a rating level, each column C to an attribute, the size of each bar indicates how many participants gave the rating R and at selected the attribute C for this food. For example 34 participants rated the food (H) and 28 selected good portion size as an attribute.

From our analysis, we found that there are two possible patterns could be noticed in the visual inspection: the **factual pattern** and the **disagreement pattern**. An example of the factual pattern is highlighted in Figure 6.6 with the number 1. We noticed that some features are mentioned by a large number of participants who rated the food: almost everyone agrees that this feature is important for the food in question. In some of these cases there is also a strong agreement about the value of the feature. For example in Figure 6.6 28 out of the 34 participants who rated this food mentioned "portion" (either bad portion or good portion), moreover, 27 of the 28 participants agree that this is a good portion. Even though these participants do not agree on how healthy the food should be rated (the rating values range from 1 to 7), we believe that it is important to highlight the agreement about the specific feature, which suggests a fact about this food. We refer to this, then, as the "factual" pattern. We found that these patterns are relatively common in the data we collected: if we define, for example, 67 as the threshold for the proportion of raters mentioning an attribute, and 90% as the threshold for the agreement, we can count 54 patterns out of 197 attribute distributions. These threshold values were selected just to provide an example, further research is needed to establish good values for those or some more clever machine learning methods to cluster such patterns. We also noticed a number of instances where a large proportion of the raters selects an attribute, but there is no strong agreement about its value. See for example the first two columns of Figure 6.6: in this case 13 raters selected low fat and 13 high fat, yet there seems to be no relation between these attributes and the ratings.

The second pattern we noticed is the "disagreement" pattern: two groups of raters select opposite values of the same attribute and rate the food at opposite ends of the scale. For example, in Figure 6.6 one group of 9 participants selected the attribute no fruit-&-veg and rated the food as unhealthy, while another group selected the attribute fruit-&-veg and rated the food as healthy. Our interpretation is that the two different groups assessed the food differently because of different understandings or beliefs about it. Some people think the yogurt is healthy because they believe it contains fruit (like the label says) while other people believes there is no "real" fruit in it. More formally, this pattern can be identified by (1) running a t-test on the ratings corresponding to two opposite values of an attribute, if the result is that the averages are significantly different and (2) checking that the distributions of two opposite attribute values are skewed and tend towards opposite direction. Using this definition we counted a total of 8 "disagreement" patterns out of 197 attribute distributions.

These two patterns are interesting and potentially useful when we consider extracting knowledge on food beliefs from the data we could collect. They are also be useful when we consider how to design visualizations to guide people pay attention to certain attributes of food. For instance, we might highlight in particular factual patterns on uncommon attributes, such as brand, whole food, and healthy fat, to try and help people gain the new knowledge on such attributes. Further more, if we could show people where the disagreement occurs and help people further explore the details behind such disagreement it will be helpful for end-users to look at different perspectives on the same food and decide which side they will believe, For instance, in the food (H), the fruit yogurt, we could show end-users that a group of raters believe that the food contains fruit and this group rates the food as healthy, while another group believes there is not fruit here and the food is unhealthy. In this way the visualization should almost ask end-users the question: "who do you believe?"

6.2.3 Towards a data infrastructure for future food-related systems

In this work, we envision that data about people's food beliefs could be useful to help people make sense of food's healthiness thus improve their healthy eating habits. The data patterns we discovered further demonstrate such potential and we believe that the future application around food and diet should build upon a data infrastructure that is able to capture people's food beliefs.

Given the fact that most health applications nowadays focus exclusively on prescribed interventions (Baumer et al.), it is not surprising that healthy eating applications usually rely on official healthy eating guidelines or nutrition counting. The approach is problematic since guidelines and nutrition counting could be wrong (Novotny et al., 2012; United States Government Accountability Office, 2008) and often ignore the complexity of eating (Comber et al.; Maitland et al., 2009). The data infrastructure we

propose here has the potential to help applications develop personalized interventions based upon people's food beliefs, so the end user who concern own diet will not be forced to commit to any kind of healthy eating defined by third- parties, no matter whether it is government department or diet community. For instance, the traditional dietary coach system usually offers recommendation based upon a so- called user profile, which includes 1) the eating history of the person and 2) the general food type preference. But such recommendation systems do not dig into people's beliefs about food. We here believe if such applications could leverage the food beliefs data , it can take one person's food beliefs into account when making suggestions.

We also believe collecting data about people's food belief can help create better personalized educational intervention to identify specific knowledge gap and thus correct it or improve the understanding. For instance, there is one generally accepted belief that all types of fat is bad but actually it is not true. If a user holding this belief rates and assesses food's healthiness for a while in the system that we described in Chapter 4 then the system can detect there is pattern that those low score foods rated by this user are always associated with "contains fat" or "high fat". And this pattern can trigger the system to start implementing a set of rules we as designers can predefine: 1) show a small tip on the assessment interface to highlight that not all fats are bad 2) highlight opposite opinions whenever the user looking at the result of food's assessment 3) push educational message/tip to the user to motivate him/her read about more information on why not all fat are bad.

The belief data also can enable the system to be able to acknowledge and respect the user's proper beliefs that might be shared within a particular group of dieters but is different from the system's pre- defined guidelines on healthy eating. With this piece of information about the person, the system could better know the stakes and objectives of the user and thus it can be smart to negotiate with the user on how plans should be tailored.

Moreover, we think the food belief data can help future systems to better design the inquiry UI in food-related system to capture people's "it depends" questions. Current food systems assume people ask the general question "is it health" and try to offer an correct answer but actually people always are looking for information tailored to their specific needs and beliefs. With the food belief data, designers could possibly come up with new ideas on how people's inquiry such as "is this healthy after a workout if I'm on a low-fat diet with X, Y, Z food beliefs?" could be captured by the system.

The power of the food belief data is not limited in the interaction between human and technology. We think such food belief data will be as important as health electronic record to help dietitians to better deliver their services and to help policy makers better design the healthy eating guidelines to inform different group of population. It will also help the de- sign of nutrition labels on food package. As we discussed above, the nutrition

label might not affect peoples mind on branded food, if we could find out what kind of biases people form around such food, the key information to inform healthy eating decision making could be highlighted on the package and truly help peoples decision making.

6.3 Summary

A goal of mobile health applications is to provide decision support for people around making better informed, healthy choices. A single correct answer is always assumed by food-related application but our work with users and domain experts demonstrates that there are too many variables in the consideration of "food healthiness" and indeed about perspectives of what constitutes food healthiness to arrive at a consensus view. Consequently, we want to explore how we might collect people's judgement on food's healthiness to extract knowledge on people's food beliefs and biases and use such knowledge as context to better personalize food-related systems. We, therefore, developed three UI designs to explore best approaches to capture sufficient data about rationale for a given food judgement but light enough to enable that capture to be done without placing excessive burden on the assessor.

In this work, we believe we make two major contributions towards understanding the new design space of "it depends":

First, we propose three simple user interfaces to capture people's rational of food's healthiness judgment and present a detail comparison between three UIs to show the tradeoffs and values of each. Our analysis provides detail comparison between three UIs from three perspective: efficiency, data quality and potential for presentation of the collected data. The predefined list shows its strength on collecting relative large amount and diverse food attributes with reasonable efforts of crowds and its effect on assisting "receiver" to gain new knowledge. It seems that it also could produce two possible patterns: Factual and Disagreement patterns, that may be used to detect and show people's food beliefs and biases. The heuristic checks, on the other hand, shows people could take advantage of its simplicity to quickly input data and easily get the message conveyed by the data it captured. However, compared with predefined list, the food beliefs and biases it could detect are limited. Finally, the open text UI shows its ability to capture detail thoughts on particular attributes such as brand, freshness, food's effect on health, and etc. But the data it captured is hard to produce and useful for people to reflect upon own thoughts.

Second and finally, based upon our analysis, we suggest that towards solving the complexity of "it depends", a data infrastructure should be built to mine and identify possible food beliefs patterns in a group of people or individual. The two data pattern we identified in the data we collected: factual pattern and disagreement pattern are just the

initial result of our exploration. We strongly encourage future research tries to capture more data to find out whether there is more patterns could possibly be identified and we think such data will help not just technology design but also policy making to tackle the healthy eating problem.

Chapter 7

Conclusion

When considering the domain of behavior change technology, much attention has been paid to how can we integrate behavior change guidelines and methods into technology-based intervention to make people do the change much easier. In contrast, there has been relatively little investigation into how people actually understand what means healthy and how those understandings make each person as a unique user for such technology-based intervention. The work in this thesis focuses on the specific context: dietary change, and provides insight into what people's understanding of food's healthiness. The investigation opens up the new "it depends" design space and our attempt to capture those food knowledge help such food-related application advance in leveraging people's food beliefs to offer personalized interventions and services.

7.1 Thesis Summary

In order to suitably locate this work within its broader behavior change context and provide a point of reference to the theoretical grounding of many of the systems to be discussed in subsequent chapters, Chapter 2 reviewed four parts of related works. First, we discussed the difference between health application and wellness application to state it clearly that this thesis is focused on wellness applications, particularly on healthy eating application. Then we lay out the specific challenges of wellness applications in HCI, which will be covered and discussed throughout the thesis. Second, it overviewed the healthy eating problems and healthy eating guidelines to discuss why dietary change is so hard to be achieved. Then we reviewed three levels of behavior change theories:individual-based, interpersonal-based, and community-based to set the theoretical basis for following discussion on example systems. Third, we went over example systems developed to support behavior change. Since there were not too many examples specifically about dietary change category, we talked about a lot physical activity related interventions in that section. The deep analysis of those systems demonstrated

why persuasion-based technology is not enough. Finally, we reviewed the different design notions for technology-based interventions and reviewed related theories of reflection to start explore how we could leverage what people's thinking about foods to offer better interventions.

The related work we presented in this thesis also demonstrated our efforts on reflecting upon current wellness intervention practice (persuasion-based intervention) and identifying that current interventions, which are most persuasion-based, have not yet taken into account all complex of healthy eating into account and this gap makes those intervention can not work for life-long and often users drop out the intervention because they either do not think it will work for them or they think it is too easy to achieve the goal. This gap we identified during the literature review actually answers our first research question stated in introduction: what is the gap between nutrition science and current design practices of healthy eating applications? and that led us to consider alternative notion that how can we take into account what people think about food when designing the intervention.

In Chapter 3, we reported the very first study that tried to understand what people think about healthy eating and food when using current interventions. This qualitative study analyzed two online dietary change forums to explore what people think about food and how they support each other in the particular dietary change context. The result from the study provided insight into a set of information people like to share and discuss, and especially we highlighted the physical context information and emotion information in our result. We also identified two social support patterns: similarity-based interaction and social review mechanism.

Our effort in this very first study was trying to understand what kinds of thoughts about food we can possibly leverage in the application design and how can we possibly collect it. The result of study thus gave us the hint that we can set up a similar social space like forum for people to ask for help on answering whether the food is healthy for me. As we presented in the result, people expect far more beyond calorie counting and they expect the answer can address particular individual context. Thus we developed the idea that uses a popular photo-based community to attract people take a picture of own food and shared it in the community so others can looked at it and gave feedbacks on its healthiness. To make it easy for users to contribute and compare the result between foods, we employed a list of heuristics to help people assess foods.

The application idea and its implementations was documented in Chapter 4 and the pilot study evaluating the application was also reported. The details design of the application and rationale behind certain design decision were well discussed in that chapter. The pilot study we ran only involved a small group of people, however, we still got interesting feedback from our participants and have the preliminary evidence that people did find

expressing and exchanging their thoughts on food's healthiness could be helpful for them to reflect upon and improve own food knowledge.

It is interesting to see this simple idea was also implemented in the other two commercial applications. Having a close look at how people use those two commercial applications, we found that people generally appreciate such opportunity to exchange their thoughts about food's healthiness with peers and have time to learn through this interaction. However, we also noticed that people have different views on similar food and rate them differently. That motivated us to investigate a vital question behind the application: how actually people convey their thoughts on food's healthiness into the application through the user interface? To answer this question, we first need to understand exactly what kinds of food thoughts might be mentioned and can we actually group them into themes/topics. That exactly is the second research question we ask in the introduction.

We investigated this question via a survey study which included 15 food photos from real users and asked volunteers (both professionals and non-professionals) to assess food's healthiness and stated their reason in plain text. Chapter 5 presented the result of this study and went through all 8 categories of food beliefs we identified. Particularly, among those beliefs, we found the uncertainty part of food was interesting and we found that people did value context information about the person who eats the food, which included not just where and when the person eats the food but much more about the person's own health status and beliefs such as whether the person believes all fat is evil.

The result of the study did not just provide evidence that people do have knowledge about food and are able to use them when assessing food but also answer the other research question we asked in the introduction: how are people's thoughts about food different from the message conveyed by current healthy eating applications? Our research found that even nutrition experts can not agree on how healthy a food is and why it is healthy or not. That challenges the most widely held assumption behind healthy eating applications: there is one single answer for healthy eating and one guidelines or standards of what is healthy food can work for all. This assumption, in fact, is not true since in nutrition literature and behavior theories it is clearly stated that there are many complicated factors determining how healthy a food is and in Chapter 2, we already discussed those issues. But it is important to address this problem in our research because our findings actually point out the gap between what we know in nutrition science and how actually healthy eating applications are really designed today. We thus argued that healthy eating application design should acknowledge the uniqueness of each user and try to help people make possible changes based upon their thoughts. Furthermore, in Chapter 5, we also argued that calorie should not be the focus of food-related application design rather we need to find out what people most concern when they assessing food and use that as the main metric to motivate people change and help them make changes.

The main study we reported in Chapter 5 further led us to conclude that the design of healthy eating application should address "it depends" which means the healthiness of food is dependent on physical context and what people believe about food. In other words, It means healthy eating application should acknowledge that food's healthiness is a context-dependent and multi-dimensional concept. Therefore, the application design should find a way to negotiate what does healthy food means with the specific end-user to deliver personalized services. To achieve so, we believe we need to first be able to capture what people think about food's healthiness and find out the possible patterns in people's food belief. That led us to think about how to design such a user interface to capture what people think about food's healthiness. In Chapter 6, we presented our three simple designs: open text box, predefined list and heuristic checks. Those three interfaces were evaluated from three perspectives: efficiency, data quality, and effect of visualization. Our result offered a deep comparison between three interfaces and found that each interface has its own strength and weakness. For instance, heuristic checks takes less time to complete but its data are limited in the five choices we define. In contrast, open text relatively takes more time to complete but can possibly provide detailed information on certain aspect of the food. Based upon our analysis, we proposed strategies on how to combine those interfaces in suitable crowdsourcing setting and offer suggestions on how to possibly combine those interfaces into together to collect more detailed information. Moreover, we also provided suggestions on how to visualize those knowledge we capture and tested their effects. The result suggested that showing people the such knowledge could be useful but the effect size could be improved if we could better personalized the visualization based upon the person's specific understanding of the food.

7.2 Contribution

This thesis work reflects upon current designs for dietary behavior change (i.e healthy eating) and focus on dietary change to explore alternative design notions.The research question we asked in the beginning of the thesis is: how can we capture and then leverage people's thought about food to potentially better personalize the intervention and engage people in adopting to the new behavior? and that is broken down into four small questions:

- How are people's thoughts about food different from the message conveyed by current healthy eating applications? and how does it create a gap between nutrition theory and real wellness application design?
- What are people's thoughts and beliefs about food and what kinds of categories/-topics emerge from those thoughts about food?

- What kinds of user interface can we develop to capture those thoughts about food?
- How thought about food might be visualized and what is it's effect on influencing people's understanding of food?

Correspondingly, the thesis has answered above questions through our reflection upon literature review, the qualitative research of online healthy eating community, the survey study to understand how people assess food's healthiness, and UI development and evaluation. We, thus, made following contributions.

First, the thesis work provides insight into what people think about food's healthiness and shows the difference between people's thoughts on what is healthy food and what is defined in current designs. This type of contribution includes:

- Identified 8 categories food-related attributes people take into account when assessing food's healthiness. That provides the basis for designing new technology-based dietary intervention by addressing the possible gap between what is healthy food by people's own definition and by the designer's definition.
- Evidence that not only lay people but also nutrition experts cannot reach consensus on what is healthy when assessing individual food. This challenges the assumption behind existing healthy eating applications that there is one single answer for healthy eating and indicates that there is a gap between what is known in nutrition literature and how currently healthy eating application are actually designed.

Second, the thesis work opens up the design space of "it depends" and explored how to design user interfaces to start capturing what people think and how possibly we can use those kind of new knowledge to support behavior change. Those contributions in this category include:

- Designed and evaluated three user interfaces: open text box, predefined list and heuristic checks to show their strength and weakness in capturing what people think about food and presented two initial patterns from data: disagreement pattern and factual pattern.
- Based upon the patterns we identified, we designed, implemented and evaluated the visualizations for those data captured by three user interfaces. But we did not observe significant impact of visualizations on changing people's belief.

7.3 Future Work

There are several ways to continue the work presented in this thesis.

First, future work can further investigate how technology can be further developed to better capture people's food thoughts. And beyond than, what kind of new technology we need to transform those data into useful knowledge to assist end-users' decision making.

In this thesis, we used three simple UIs to capture food beliefs and from our analysis, we know it is possible to consider how to combine these three UIs together to capture more detailed information. For instance, we can combine predefined list and text-box so people can still use predefined list to quickly assess the food but they are also given the freedom to further explain certain points they most concern. Moreover, it is also important to further understand how we can collect rich information in real-time. Using a combination of UIs might be good for collecting rich information but it might take much longer time since every worker needs to provide very detail assessment. Therefore, one challenge here is how can we better design the crowdsourcing workflow to reduce each worker's work load but still get rich information from the crowds. One possible approach is to take the "Divide and Conquer" method: each worker in the crowdsourcing process does not need to assess every aspects of the food but instead, the worker will only be assigned a small task. For instance, we can design one micro-task to ask people to use predefined list to assess food's healthiness, and based upon the aggregated result of predefine list, we may automatically identify the disagreement pattern in it and create another micro-task to ask another group of workers to explain why that possibly happened and what is your explanation. Thus we can collect more information about the food but from a large group of people in a distributed manner. The related problem then is how can we better visualize the information and present them to the end-users in a relative real-time manner. To do so, we need to further understand what aspects of the food the end-user most concerns then we possibly arrange the order of the crowdsourcing workflow based upon it to show the end-user the most concerned information first to help the end-user make decision quickly. In addition, we did not observe significant impact of food belief information in our study, so better strategy of visualization should be explored to help people notice certain information they might neglect.

Also, more work should be done to further understand the food belief data we collect. In the thesis, we found two possible patterns we can further use to help design the visualization and identify specific group of people. The way to identify the pattern has not been confirmed by mathematical methods and it should be done with bigger size of data. The demographic of participant did not show effect on the food beliefs, but we think with more diversity data and large sample of participants, this kind of relationship might also be well studied and used.

Second, we also think more work can be done to understand how the food belief can be further leveraged and integrated into the application idea we described in Chapter 4. The application could be deployed into real market to reach large population and the data it potentially collects could help future researcher better investigate what kinds of

patterns of food beliefs might be mined from it. Also, a large scale study could help identify the difference in food beliefs between different demographic groups and that will be valuable to group people together to support each other and deliver information based upon it.

Third, we suggest future research could further investigate how people's concern about food's healthiness changes from one context to another and how these affect the weight of each attribute of foods we identified in Chapter 5. Such theoretical work maybe seems not to be directly relevant to the human-computer interaction but it is important to understand it since wellness application for healthy eating needs to use those information to create personalized interventions that match people's values and beliefs.

Finally, we suggest researchers can explore how the "it depends" question might be applied in other scenarios and how the crowdsoucing interfaces we developed in the Chapter 6 might be useful in other scenarios. in our research, we focused on the food healthiness. The question on whether a food is healthy or not is always "it depends" and so there is no single correct answer for it. The similar question might be asked in other contexts especially customer reviews such as whether a movie is fantastic or terrible. In such context, a single quantitative indicator of whether the movie is fantastic or terrible is often employed to help customers make quick decision. But such rating is usually also accompanied with text-based reviews for the customer to make informed decisions. Thus, the key here is not actually to produce a single answer but more importantly is to break down the reasons behind the rating. The interfaces we proposed and tested in Chapter 6 might be helpful in other scenarios such as rating movies to quickly break down the ratings into details that matter to individuals or groups. Future researches could investigate how the interfaces might be applied in such different context but also find out what should be changed as rating movies is not exactly the same as rating foods. Applying the interfaces we proposed into another scenario might also further inform how detail assessments could be presented/visualized to assist decision making especially real-time decision making.

7.4 Closing Remarks

This dissertation is motivated by growing trend in human-computer interaction whereby more applications are designed to help people change behavior to live in a healthier way. The over-arching goal of this dissertation is to explore how we can capture more data about what people think about what healthy behavior means to themselves and leverage such thoughts to deliver better personalized interventions and services to support behavior change.

In particular, this dissertation explored this design notion in the context of dietary change and investigated how people's thoughts on what is healthy food could be captured

and used. Food's healthiness is naturally multi-dimensional and context-dependent. That means there is no single correct answer to whether chocolate, for example, is healthy or not but it should always depend on who will eat it and what he/she wants to get from the food. The gap between what is known in nutrition literature and how exactly current healthy eating applications are designed, thus, is addressed and validated in this thesis. We argue that the interventions created to support dietary change need to acknowledge the "it depends" problem and offer right suggestions to right person in a right context, which is not only about physical context such as location and weather but also people's various food beliefs. We further suggest that there is value in thinking beyond calorie and leverage what people think about food's healthiness to design better personalized interventions.

However, to leverage such thoughts, we need to first understand what kinds of data we could collect from people and how could we capture them in the first place. In this thesis, the insights into people's knowledge about food are explored and discussed. And the way to capture that knowledge is presented and evaluated and suggestions on how to move on to design better tools to capture those knowledge are also discussed. Finally, the way to possibly visualize those information to help people reflect upon own thoughts to improve knowledge and motivate changes are also demonstrated in the thesis. We hope that the thesis could help researchers and designers in the domain of behavior change technology reflect upon how to design applications based upon what people think and feel and leverage what we have found and tried to create a data infrastructure to leverage those thoughts to better deliver personalized interventions.

Appendix A

Larger Figures

Figure A.1: The Rating Distributions of 15 foods for follow-up study. The foods are: (A) Three shrimp meal; (B) Cake topped with fruit; (C) A can of carbonated drink; (D) Grilled veggies with pesto sandwich, fries, potato soup, fresh squeezed and strawberry lemonade; (E) Quaker Life soft baked cinnamon bar; (F) Rye bread with beetroot; (G) Meatball, egg, rice and tomato; (H) Fruit Yogurt; (I) Wholewheat with honey bread; (J) Feta cheese, cold cut meats, mushrooms, two eggs and one egg white; (K) Honey bunches of oats with pecans and blueberries; (L) Freshly squeezed grapefruit; (M) Banana; (N) Almonds; (O) Salad.

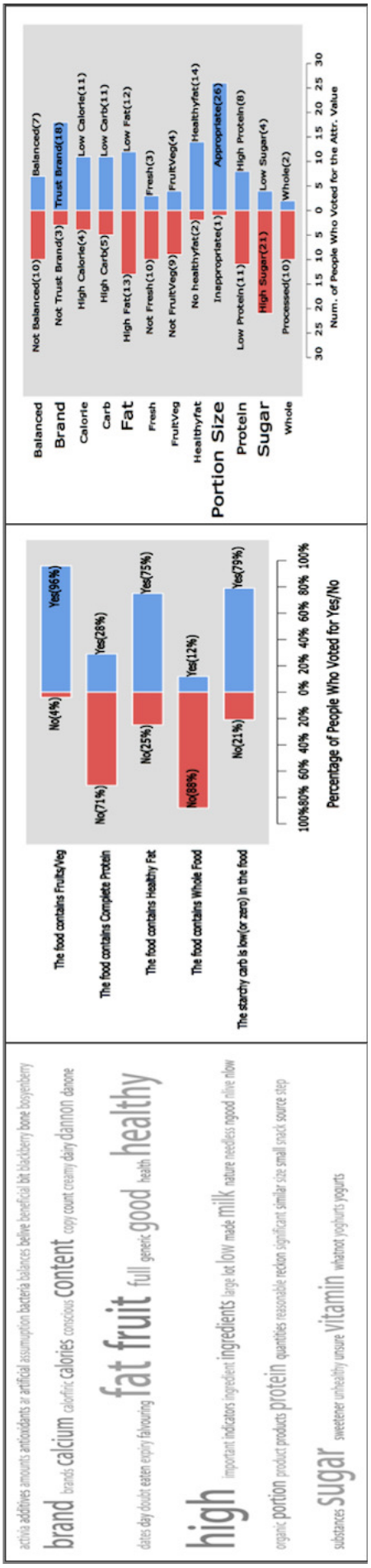

Figure A.2: From left to right: wordcloud for open text, barchart for heuristic checks, and cloud-barchart for predefined list.

Figure A.3: The Attribute Distribution of Food (H), Fruit Yogurt. The Y-axis is the 7-star scale value from 1 (top) to 7 (bottom), and the X-axis is attribute values of this particular food. For each two columns, they represent a pair of the attribute (e.g. fat) with different two values (lowfat .v.s highfat). For Mark 1, it is the Factual pattern on the portion size attribute with value: badportion and good portion . Mark 2 is the Disagreement pattern on the nofruitveg and fruitveg.

References

J Adams and M White. Why don't stage based activity promotion interventions work? *Health Education Research*, 20(2):237–243, 2005.

Icek Ajzen. Perceived Behavioral Control, Self-Efficacy, Locus of Control, and the Theory of Planned Behavior1. *Journal of Applied Social Psychology*, 32(4):665–683, 2002. ISSN 00219029.

Albert Bandura. *Social Foundations of Thought and Action: A Social Cognitive Theory [Paperback]*. Prentice Hall; 1 edition, 1985. ISBN 013815614X.

Albert Bandura. *Self-Efficacy: The Exercise of Control*. Worth Publishers; 1st edition, 1997. ISBN 0716728508.

Eric P.S. Baumer, Sherri Jean Katz, Jill E. Freeman, Phil Adams, Amy L. Gonzales, John Pollak, Daniela Retelny, Jeff Niederdeppe, Christine M. Olson, and Geri K. Gay. Prescriptive persuasion and open-ended social awareness: expanding the design space of mobile health. In *CSCW(2012)*. ISBN 978-1-4503-1086-4.

Timothy W Bickmore, L Caruso, K Cloughgorr, and T Heeren. It's just like you talk to a friend relational agents for older adults. *Interacting with Computers*, 17(6):711–735, 2005. ISSN 09535438.

David Boud, Rosemary Keogh, and David Walker. *Reflection: turning experience into learning*. Routledge, 1985. ISBN 0850388643.

Emily Brindal, Carlene Wilson, Philip Mohr, and Gary Wittert. Perceptions of portion size and energy content: implications for strategies to affect behaviour change. *Public Health Nutrition*, 15(02):246–253, February 2012. ISSN 1475-2727.

Colette Browning and Shane A. Thomas. Models of behaviour change and health promotion. In *Behavioral Change: An Evidence-based Handbook for Social and Public Health*, pages 7–16. Churchill Livingstone; 1 edition, 2005.

Peter Byass. Youth: choices and change. *Bulletin of the World Health Organization*, 84 (6):503, 2006.

Robert A Carels, Jessica Harper, and Krista Konrad. Qualitative perceptions and caloric estimations of healthy and unhealthy foods by behavioral weight loss participants. *Appetite*, 46(2):199–206, 2006.

Robert A Carels, Krista Konrad, and Jessica Harper. Individual differences in food perceptions and calorie estimation: an examination of dieting status, weight, and gender. *Appetite*, 49(2):450–8, September 2007. ISSN 0195-6663.

Frank Chen. What is a Mindful Technology? Design Mindfulness, 2011.

Pei-Yu (Peggy) Chi, Jen-Hao Chen, Hao-Hua Chu, and Jin-Ling Lo. Enabling Calorie-Aware Cooking in a Smart Kitchen. In *Persuasive(2008)*, volume 5033 of *Lecture Notes in Computer Science*, pages 116–127, 2008. ISBN 978-3-540-68500-5.

Gregory N. Choi, Jaz Hee-jeong, Foth, Marcus, & Hearn. *Eat Cook Grow : Mixing Human-Computer Interactions with Human-Food Interactions*. MIT Press, Cambridge, MA, 2013.

ChooseMyPlate.gov. USDA's MyPlate, 2011.

Julie A Chronister, Erica K Johnson, and Norman L Berven. Measuring social support in rehabilitation. *Disability and rehabilitation*, 28(2):75–84, 2006.

Rob Comber, Jettie Hoonhout, Aart van Halteren, Paula Moynihan, and Patrick Olivier. Food practices as situated action. In *CHI(2013)*, page 2457. ISBN 9781450318990.

S Consolvo, K Everitt, I Smith, and J.A. Landay. Design requirements for technologies that encourage physical activity. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, page 466. ACM, 2006.

Sunny Consolvo, Predrag Klasnja, David W. McDonald, Daniel Avrahami, Jon Froehlich, Louis LeGrand, Ryan Libby, Keith Mosher, and James A. Landay. Flowers or a robot army? Flowers or a robot army?: encouraging awareness & activity with personal, mobile displays. In *UbiComp (2008)*, pages 54–63. ISBN 9781605581361.

J K Croll, D Neumark-Sztainer, and M Story. Healthy eating: what does it mean to adolescents? *JNutritionEducation*, 33(4):193–198, 2001.

Rachel Davis, Danice Cook, and Larry Cohen. A community resilience approach to reducing ethnic and racial disparities in health. *American Journal of Public Health*, 95(12):2168–2173, 2005.

M D De Almeida, P Graça, C Afonso, J M Kearney, and M J Gibney. Healthy eating in European elderly: concepts, barriers and benefits. *The journal of nutrition health aging*, 5(4):217–219, 2001.

Helen Dixon and Susan Paxton. Eating and nutrition. In *Behavioural Change: An Evidence-Based Handbook for Social and Public Health*, pages 45–68. Churchill Livingstone; 1 edition, 2005. ISBN 0443073570.

J S A Edwards and H H Hartwell. Fruit and vegetables–attitudes and knowledge of primary school children. *JHumanNutrition&Dietetics*, 15(5):365–374, 2002.

PJ Edwards, KP Moloney, and JA Jacko. HCI Challenges Case Study: implementing an electronic medical record. In *CHI (2006)*, pages 1–4.

Nicole Eikenberry and Chery Smith. Healthful eating: perceptions, motivations, barriers, and promoters in low-income Minnesota communities. *JAmericanDieteticAssociation*, 104(7):1158–1161, 2004.

L W Falk, J Sobal, C A Bisogni, M Connors, and C M Devine. Managing healthy eating: definitions, classifications, and strategies. *Health Education & Behavior*, 28 (4):425–439, 2001.

Qinyuan Feng, Ling Liu, and Yafei Dai. Vulnerabilities and countermeasures in context-aware social rating services. *ACM Transactions on Internet Technology*, 11(3):1–27, January 2012. ISSN 15335399.

L. Festinger. A Theory of Social Comparison Processes. *Human Relations*, 7(2):117–140, May 1954. ISSN 0018-7267.

B. J. Fogg. *Persuasive technology: using computers to change what we think and do.* Morgan Kaufmann Publishers Inc., 1st edition, 2002.

Food Standards Agency. The Eatwell Plate, 2010.

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

Roland Gasser, Dominique Brodbeck, Markus Degen, and Jürg Luthiger. Persuasiveness of a Mobile Lifestyle Coaching Application Using Social Facilitation. *Persuasive Technology*, 3962(27 - 38):27–38, 2006. ISSN 03029743.

Karen Glanz, Barbara K Rimer, and K Vish Viswanath. *Health behavior and health education: theory, research, and practice*, volume 4th. Jossey-Bass, 2008. ISBN 9780787996147.

J S Greenberg. Health and wellness: a conceptual differentiation. *The Journal of school health*, 55(10):403–406, 1985.

Andrea Grimes, Martin Bednar, J.D. Bolter, and R.E. Grinter. EatWell: sharing nutrition-related memories in a low-income community. In *CSCW (2008)*, pages 87–96, 2008.

Andrea Grimes, Vasudhara Kantroo, and Rebecca E. Grinter. Let's play!: mobile health games for adults. In *Proceedings of the 12th ACM international conference on Ubiquitous computing - Ubicomp '10*, page 241, New York, New York, USA, September 2010. ACM Press. ISBN 9781605588438.

Sajanee Halko and Julie A. Kientz. Personality and persuasive technology: An exploratory study on health-promoting mobile applications. In *Proceedings of the 5th International Conference on Persuasive Technology*, PERSUASIVE'10, pages 150–161, Berlin, Heidelberg, 2010. Springer-Verlag. ISBN 3-642-13225-1, 978-3-642-13225-4.

Lars Hallnäs and Johan Redström. Slow Technology Designing for Reflection. *Personal and Ubiquitous Computing*, 5(3), 2001. ISSN 1617-4909.

W Hardeman, S Griffin, M Johnston, A L Kinmonth, and N J Wareham. Interventions to prevent weight gain: a systematic review of psychological models and behaviour change methods. *International journal of obesity and related metabolic disorders journal of the International Association for the Study of Obesity*, 24(2):131–143, 2000.

Harvard School of Public Health. Harvard Healthy Eating Plate, 2011.

T A Herzog, D B Abrams, K M Emmons, L A Linnan, and W G Shadel. Do processes of change predict smoking stage movements? A prospective analysis of the transtheoretical model. *Health psychology official journal of the Division of Health Psychology American Psychological Association*, 18(4):369–375, 1999.

Mary Margaret Huizinga, Adam J Carlisle, Kerri L Cavanaugh, Dianne L Davis, Rebecca P Gregory, David G Schlundt, and Russell L Rothman. Literacy, numeracy, and portion-size estimation skills. *American journal of preventive medicine*, 36(4):324–8, April 2009. ISSN 1873-2607.

A J Hutchison, J D Breckon, and L H Johnston. Physical activity behavior change interventions based on the transtheoretical model: a systematic review. *Health education behavior the official publication of the Society for Public Health Education*, 36 (5):829–845, 2009.

W P T James. WHO recognition of the global obesity epidemic. *INT. JObesity*, 32 Suppl 7(S7):S120–6, December 2008. ISSN 1476-5497.

Anne Keane and Anna Willetts. *Concepts of Healthy Eating: An Anthropological Investigation in South East London*. Goldsmiths' College, 1995. ISBN 0901542822.

Predrag Klasnja, Sunny Consolvo, and Wanda Pratt. How to evaluate technologies for health behavior change in HCI research. In *Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11*, page 3063, New York, New York, USA, May 2011. ACM Press. ISBN 9781450302289.

Wolfgang Kopp. High-insulinogenic nutritionan etiologic factor for obesity and the metabolic syndrome? *Metabolism*, 52(7):840–844, July 2003. ISSN 00260495.

Young S. Lee, Katie Siek, Carsten Röcker, Jochen Meyer, Susanne Boll, and Oscar Mayora. Wellness interventions and HCI. *ACM SIGHIT Record*, 2(2):51–53, 2012. ISSN 21588813.

J.J. Lin, Lena Mamykina, Silvia Lindtner, Gregory Delajoux, and H.B. Strub. Fish'n'Steps: encouraging physical activity with an interactive computer game. In *Lecture Notes in Computer Science*, volume 4206, pages 261–278. Springer, 2006.

Edwin A Locke and Gary P Latham. Building a practically useful theory of goal setting and task motivation. A 35-year odyssey. *American Psychologist*, 57(9):705–717, 2002.

S Mackey. Towards a definition of wellness. *The Australian journal of holistic nursing*, 7(2):34–38, 2000.

Julie Maitland. *From persuasion to negotiation in health promoting technology*. PhD thesis, University of Glasgow, 2009.

Julie Maitland, Matthew Chalmers, and Katie A. Siek. Persuasion not required Improving our understanding of the sociotechnical context of dietary behavioural change. In *Pervasive Computing Technologies for Healthcare, 2009. PervasiveHealth 2009. 3rd International Conference on*, pages 1–8, 2009.

Lena Mamykina, Elizabeth Mynatt, Patricia Davidson, and Daniel Greenblatt. MAHI: investigation of social scaffolding for reflective thinking in diabetes management. In *CHI (2008)*, 2008.

Jennifer Mankoff, Gary Hsieh, H C Hung, Sharon Lee, and Elizabeth Nitao. Using Low-Cost Sensing to Support Nutritional Awareness. In *UbiComp (2002)*, Lecture Notes in Computer Science, pages 371–378, 2002. ISBN 3540442677.

L. McAlpine, C. Weston, C. Beauchamp, C. Wiseman, and J. Beauchamp. Building a metacognitive model of reflection. *Higher Education*, 37(2):105–131, March 1999.

E. L. Moe. Promoting Healthy Lifestyles: Alternative Models' Effects (PHLAME). *Health Education Research*, 17(5):586–596, October 2002. ISSN 14653648.

Sean Munson. Mindful Technology vs. Persuasive Technology, 2011.

NHS Choices. 5 A DAY - Live Well - NHS Choices, 2011.

NHS Health Scotland. Profiling Health Behaviours in Scotland using the Health Education Population Survey (HEPS) 1996-99 dataset. Technical report, 2000.

C Ni Mhurchu, B M Margetts, and V M Speller. Applying the stages-of-change model to dietary change. *Nutrition Reviews*, 55(1 Pt 1):10–16, 1997.

Donald A Norman. *The Psychology of Everyday Things*. Basic Books, 1988. ISBN 0465067093.

Donald A Norman and D G Bobrow. On data-limited and resource-limited processes. *Cognitive Psychology*, 7(1):44–64, 1975. ISSN 00100285.

Jon Noronha, Eric Hysen, Haoqi Zhang, and Krzysztof Z Gajos. PlateMate : Crowdsourcing Nutrition Analysis from Food Photographs. In *UIST (2011)*, 2011. ISBN 9781450307161.

Janet A Novotny, Sarah K Gebauer, and David J Baer. Discrepancy between the Atwater factor predicted and empirically measured energy values of almonds in human diets. *American Society for Nutrition*, 96(2):296–301, 2012.

M E Oakes and C S Slotterback. Judgements of food healthfulness: food name stereotypes in adults over age 25. *Appetite*, 37(1):1–8, 2001.

Harri Oinas-Kukkonen. A foundation for the study of behavior change support systems. *Personal Ubiquitous Comput.*, 17(6):1223–1235, August 2013. ISSN 1617-4909.

Marie-Claude Paquette. Perceptions of healthy eating: state of knowledge and research gaps. *Canadian JPublicHealth*, 96 Suppl 3(Suppl 3:S15-9):S15–S19, S16–S21, 2005.

I M Parraga. Determinants of food consumption. *Journal of the American Dietetic Association*, 90(5):661–3, May 1990. ISSN 0002-8223.

Robert A. Peterson. *Constructing Effective Questionnaires*. Sage Publications, Inc, 1999. ISBN 0761916407.

D Porritt. Social support in crisis: quantity or quality? Technical Report 6, 1979.

R Povey, M Conner, P Sparks, R James, and R Shepherd. Interpretations of healthy and unhealthy eating, and implications for dietary change. *Health Education Research*, 13 (2):171–183, 1998.

James O. Prochaska, John Norcross, and Carlo DiClemente. *Changing for Good: A Revolutionary Six-Stage Program for Overcoming Bad Habits and Moving Your Life Positively Forward*. Harper Paperbacks, 1995.

Stephen Purpura, Victoria Schwanda, Kaiton Williams, William Stubler, and Phoebe Sengers. Fit4Life : The Design of a Persuasive Technology Promoting Healthy Behavior and Ideal Weight. In *CHI (2011)*, pages 423–432, 2011. ISBN 9781450302678.

K. Resnicow. Are precontemplators less likely to change their dietary behavior? A prospective analysis. *Health Education Research*, 18(6):693–705, December 2003. ISSN 1465-3648.

Lina A Ricciardelli and Marita P McCabe. Dietary restraint and negative affect as mediators of body dissatisfaction and bulimic behavior in adolescent girls and boys. *Behaviour Research and Therapy*, 39(11):1317–1328, November 2001. ISSN 00057967.

W T Riley, D E Rivera, A A Atienza, Wendy Nilsen, S M Allison, and Robin Mermelstein. Health behavior models in the age of mobile interventions: are our theories up to the task? *Translational Behavioral Medicine*, 1(1):53–71, 2011.

I M Rosenstock, V J Strecher, and M H Becker. Social Learning Theory and the Health Belief Model. *Health Education Behavior*, 15(2):175–183, 1988. ISSN 10901981.

Sage Research Corporation. *Qualitative investigation of Canadians' understanding of and attitudes towards nutrition and healthy eating: final report*. Sage Research Corporation, 2003.

B. Santich. Good for you: Beliefs about food and their relation to eating habits. *Australian JNutrition&Dietetics*, 51:68–73, 1994.

Phoebe Sengers, Kirsten Boehner, Shay David, and Joseph 'Jofish' Kaye. Reflective design. In *Proceedings of the 4th decennial conference on Critical computing: between sense and sensibility*, pages 49–58, 2005.

Shauna L Shapiro, Linda E Carlson, John A Astin, and Benedict Freedman. Mechanisms of Mindfulness. *Journal of Clinical Psychology*, 62(3):373–386, 2006. ISSN 00219762.

Mical Kay Shilts, Marcel Horowitz, and Marilyn S Townsend. Goal setting as a strategy for dietary and physical activity behavior change: a review of the literature. *American journal of health promotion AJHP*, 19(2):81–93, 2004.

By Robert S Sinclair. Thinking and Writing Cognitive Science and Intelligence Analysis. *Intelligence*, (January):41, 1984.

Leslie Spencer, Christopher Wharton, Sheila Moyle, and Troy Adams. The transtheoretical model as applied to dietary behaviour and outcomes. *Nutrition Research Reviews*, 20(1):46–73, 2007.

A Strauss and J Corbin. *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*, volume 270. Sage, 1990. ISBN 0803932502.

The Independent. Confusion over official advice on healthy diet, 2008.

T Toscos, A Faber, K Connelly, and A M Upoma. Encouraging physical activity in teens Can technology help reduce barriers to physical activity in adolescent girls? In *Pervasive Computing Technologies for Healthcare 2008 PervasiveHealth 2008 Second International Conference on*, volume 3, pages 218–221. Ieee, 2008. ISBN 9789639799158.

John C. Turner. *Social Influence*. Wadsworth Publishing; 1 edition, 1991. ISBN 0534169503.

United States Government Accountability Office. Food Labeling: FDA Needs to Better Leverage Resources, Improve Oversight, and Effectively Use Available Data to Help Consumers Select Healthy Foods. Technical report, 2008.

Tatjana Van Strien and Machteld A Ouwens. Effects of distress, alexithymia and impulsivity on eating. *Eating Behaviors*, 8(2):251–257, 2007.

W Kip Viscusi. Do Smokers Underestimate Risks? *Journal of Political Economy*, 98 (6):1253, 1990. ISSN 00223808.

Mark Weiser. The computing for the 21st century. *Scientific American*, 265(3):94–104, 1991.

Barry Wellman. Social Network Research - Substantive Issues and Methodological Questions - Knipscheer,Kcpm, Antonucci,Tc. *Contemporary Psychology*, 37(12):1315–1316, 1992.

Lauren Wilcox, Rupa Patel, Anthony Back, Mary Czerwinski, Paul Gorman, Eric Horvitz, and Wanda Pratt. Patient-clinician Communication: The Roadmap for HCI. In *CHI EA (2013)*, CHI EA '13, pages 3291–3294. ISBN 978-1-4503-1952-2.

W. C. Willett. Dietary fat plays a major role in obesity: no. *Obesity Reviews*, 3(2): 59–68, May 2002. ISSN 1467-7881.

World Health Organization. Health Promotion Glossary. Technical report, 1998.

World Health Organization. Preventing chronic diseases: a vital investment. Technical report, 2005.

Robert Yinger. The conversation of practice. In Renee Tipton Clift, Robert W. Houston, and Marleen C. Pugach, editors, *Encouraging Reflective Practice in Education: An Analysis of Issues and Programs*, pages 72–96. Teachers College Pr, 1990. ISBN 0807729906.

Robert B Zajonc. Social Facilitization: A solution is suggested for an old unresolved social psychological problem. *Science*, 149:269–274, 1965.