Integrating smartphone and sensor data in a quantified self app

Kenneth Brüel Falmark

Kongens Lyngby 2013
M.Sc-2013-69

Technical University of Denmark
DTU Compute
Building 303B, DK-2800 Kongens Lyngby, Denmark
Phone +45 45253031 , Fax +45 45881399
compute@compute.dtu.dk
http://www.compute.dtu.dk/ M.Sc-2013-69

Abstract

The purpose of this thesis is to combine the fields of data visualization, quantified self and app development to create an approach which can be used to develop quantified self apps which avoid the problems of today's data representations in quantified self apps.

The problems are determined by conducting an early validation test where a number of data representations in current quantified self apps are tested. The results of the test show that there are 6 categories of problems: icons, colours, information overload, size, comparisons and explanations.

The literature and studies related to the fields of quantified self and data visualization are examined in order to understand the cause of the problems.

An approach to solve the problems is derived based on the examination of the literature and studies and the results of the early validation test. This approach is used to develop a prototype of a quantified self app. This app is then tested to see if the approach can solve the problems.

The results of the prototype test indicate that the approach could be a solution to the problems in quantified self app's data representations.

Resume

Formålet med denne afhandling er at kombinere områderne datavisualisering, quantified self og app-udvikling til at skabe en fremgangsmåde, som kan bruges til at udvikle quantified self apps, som undgår problemerne forbundet med nutidens datarepræsentationer i quantified self apps.

Problemerne bestemmes ved at gennemføre en test, hvor et antal af datarepræsentationer i nuværende quantified self apps testes. Resultatet af testen viser at der eksisterer 6 forskellige kategorier af problemer: ikoner, farver, information overload, størrelse, sammenligninger og forklaringer.

Litteratur og undersøgelser relateret til områderne quantified self og datavisualisering undersøges for at forstå årsagen til problemerne.

På baggrund af undersøgelserne og resultaterne af testen udledes en fremgangsmåde til at løse problemerne. Denne fremgangsmåde anvendes til at udvikle en prototype af en quantified self app. Derefter testes denne app for at se om fremgangsmåden kan løse problemerne.

Resultaterne af prototypetesten indikerer at den udledte fremgangsmåde kan være en løsning på problemerne i datarepræsentationerne i nutidens quantified self apps.

Preface

This thesis was prepared at the department of DTU Compute at the Technical University of Denmark in fulfilment of the requirements for acquiring an M.Sc. in Digital Media Engineering.

Lyngby, 2-August-2013

Kenneth Falmark

Kenneth Brüel Falmark

Acknowledgements

I would like to thank my advisors, Michael Kai Petersen and Jakob Eg Larsen, and the members of my study group, Mustafa Olgun, Roshanth Ranganathan and Sujeevan Kuganesan, for their assistance and advices through the work on this thesis.
I would also like to thank the participants of the early validation test and prototype test for their time and cooperation.

Contents

Chapter 1

Introduction

"Don't measure yourself by what you have accomplished, but by what you should have accomplished with your ability."
John Wooden

In today's world everything is quantifiable. A person can be quantified by his number of friends on Facebook, followers on Twitter and his activity throughout a day. Furthermore people are interested in the possibility of being able to quantify or track their behaviour. This has created **the quantified self movement**. The quantified self movement's members' interest in self tracking has led to a market which consists of over 200 different devices which track peoples' activities. These devices track everything from your sleeping patterns to your step count. The most popular device for self tracking is the smartphone due to its built-in accelerometer, GPS, etcetera. This makes it possible for your smartphone to track your behaviour every day. It tracks your communications, location and usage. A smartphone also gives its owner access to either App Store, Google Play or Windows Phone Store depending on the operating system of the smartphone. This makes it possible to download apps designed to track peoples' behaviour: quantified self apps. Some of the most well-known quantified self apps are Endomondo Sports Tracker, Mobilbank and Nike+ Running. These can help the user to track his behaviour and display the results in order for the user to understand his behaviour.

The majority of the world is familiar with both smartphones and apps since over 1 billion people own a smartphone. However apps have only been around since the launch of the App Store in 2008. Therefore it has only been around for 5 years and new technologies are often accompanied with new problems and difficulties as history has proven many times. One example is the scientific community's visualizations after the invention of the computer. The new possibilities created a new type of visualizations where the focus was on the design rather then the data. The result was that the viewers ended up saying *You can do that?* instead of *What interesting data.*

It is unknown if this has also happened with the data representations in quantified self apps. They are supposed to help the user understand his data. But it is possible that the developers are focusing more on the design than the data. If this is true you cannot help to wonder: How many problems has this created for the users of quantified self apps?

1.1 Objectives

The first objective of this thesis is to determine the problems with data representations in quantified self apps. Therefore it is necessary to research the current quantified self apps available on the various app stores in 2 stages.
The first stage will be a study of the quantified self apps and their data representations. However this study will only focus on apps which focus on sleep and exercise data. The result of this study will be the most common types of data representations and their problems.
The second stage will be an early validation test of the most common types of data representations. This test can confirm that there are problems with today's data representations in quantified self apps.

The second objective is to examine the literature and studies related to data visualization and quantified self in order to understand the problems determined in the test.

The third objective is to derive an approach based on the early validation test and the examination of the the literature and studies. The goal of this approach is to help developers to develop quantified self apps which avoid the problems of today's data representations in quantified self apps.

The fourth objective is to develop a prototype of a quantified self app with the approach. This prototype will be developed for the Android OS.

The final objective is to test the prototype in order to determine whether or not the approach has helped to develop an quantified self app which has no problems with its data representations.

1.2 Chapter overview

Chapter 2 describes the most popular quantified self apps related to sleep and exercise available on the various app stores and the field of quantified self.

Chapter 3 examines the scientific literature about visualizing data.

Chapter 4 describes the results of the early validation test of today's quantified self apps and the derived approach to eliminate the problems. Furthermore the results of developing an app with the approach is described. Finally the results of the test of the developed app is described.

Chapter 5 states the problems with data representations in quantified self apps, how the problems were solved, the success of the solution and how to improve the solution in the future.

CHAPTER 2

Related Work

"The goal isn't to figure out something about human beings generally but to discover something about yourself. Their validity may be narrow, but it is beautifully relevant."
[Wol10, p. 7]

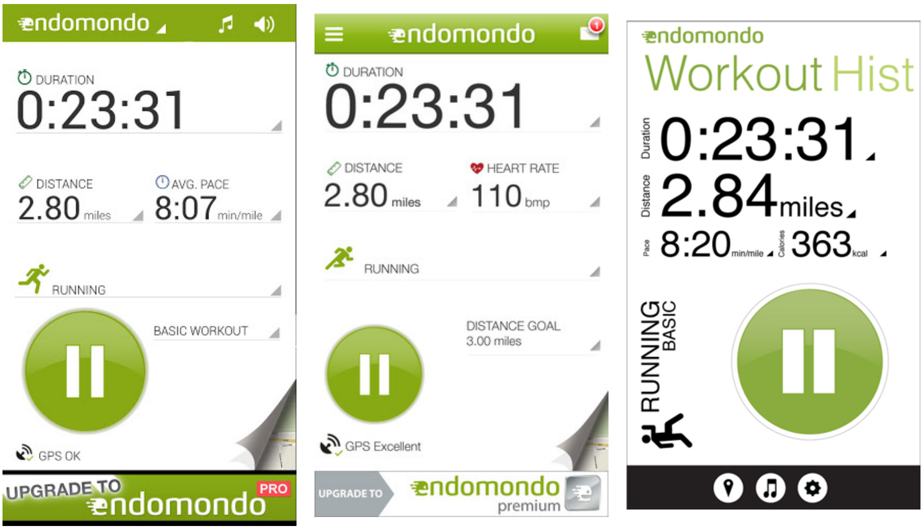
This chapter describes the most popular quantified-self apps related to sleep and exercise available on the various app stores and the field of quantified self.

2.1 Apps

In this section quantified self apps are described with a introduction to their popularity, functionalities and usage. Each app's data representations are then examined.

Endomondo

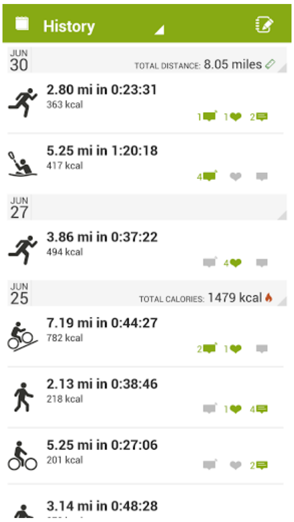
Endomondo is one of the biggest exercise quantified self apps with *"10 million registered users"*. The majority of the *"250,000 user workouts"* tracked everyday are runs [Vel12].
The design is similar on all 3 platforms in Figure 2.1. The data presented to the user is also similar. All 3 show the user the duration, distance and exercise type. Both the Android and Windows Phone app show the pace and both the Android and iPhone app show the average speed. The iPhone app further shows the goal and the heart rate whereas the Windows Phone app shows calories burned.



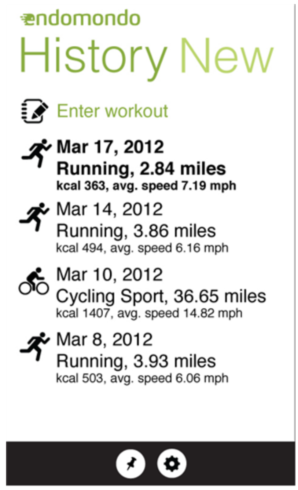
Android iPhone Windows Phone

Figure 2.1: Session view of *Endomondo* app

The Android app in Figure 2.2 shows the **social networking aspect of Endomondo** with **the history of comments, live comments and likes for each exercise** besides the data(exercise type, distance and calories) whereas the Windows Phone app only shows the data (exercise type, distance, calories and speed). The Another difference between the 2 is that **the Android app only shows the exercise type with an icon** whereas **the Windows Phone app also writes it in addition to the icon**.



Android

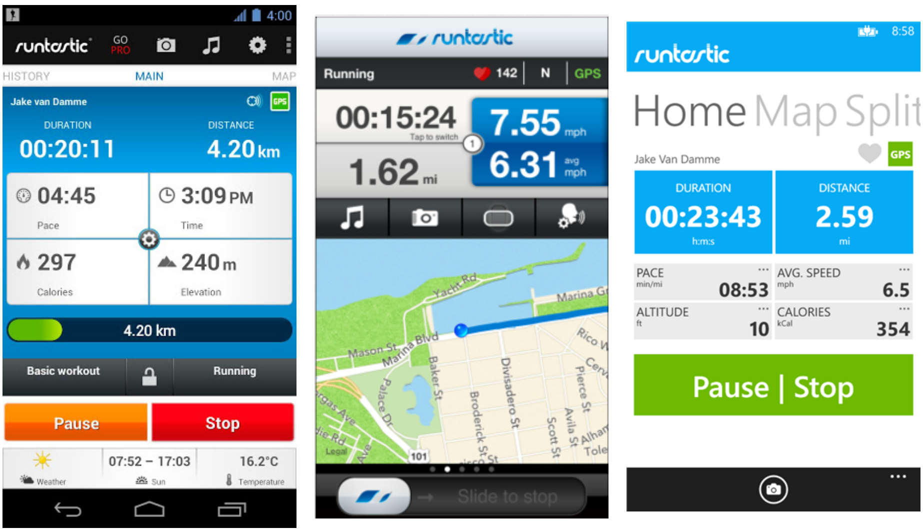


Windows Phone

Figure 2.2: History view of *Endomondo* app

Runtastic

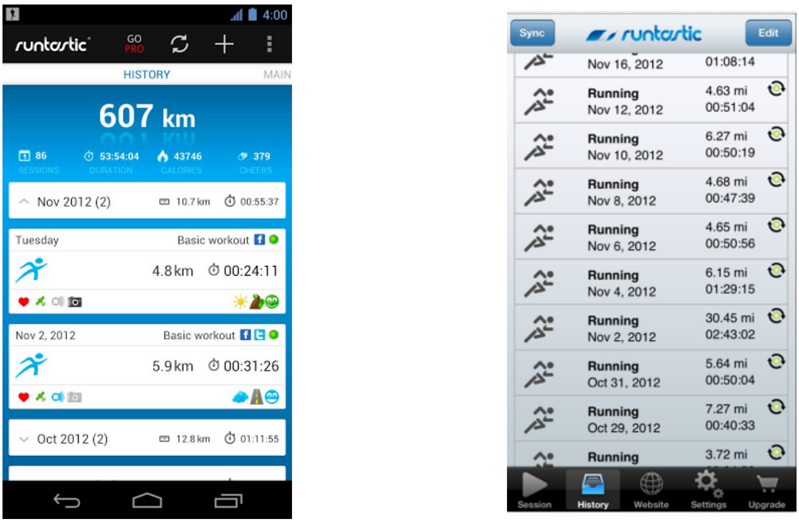
Runtastic is an international mobile fitness company which *"has 25 million mobile users across all of its apps"* [Web13]. Runtastic is responsible for a series of apps such as *Runtastic Pedometer*, *Runtastic SitUps*, *Runtastic PushUps*, etcetera. The one described in this section is the *Runtastic Running & Fitness* app.
Figure 2.3 shows the Session view of Runtastic on 3 different platforms. All show the duration, distance and speed. The biggest difference is the additional data being tracked by the Android app. The Android app **tracks the context of the run by tracking both the weather and the temperature**.



Android iPhone Windows Phone

Figure 2.3: Session view of *Runtastic* app

Figure 2.4 shows the influence of the additional data. The Android app **shows the weather, surface and temperature** in addition the exercise data whereas the iPhone app only shows the exercise data. **The Android app only uses an icon to show the type of the exercis**e compared to **the iPhone app which also writes the type of exercise**. This is similar to the approach of the Endomondo app.
One could **assume that the smileys in the Android apprepresent either temperature or mood**. But **without a legend it is impossible to be certain**.

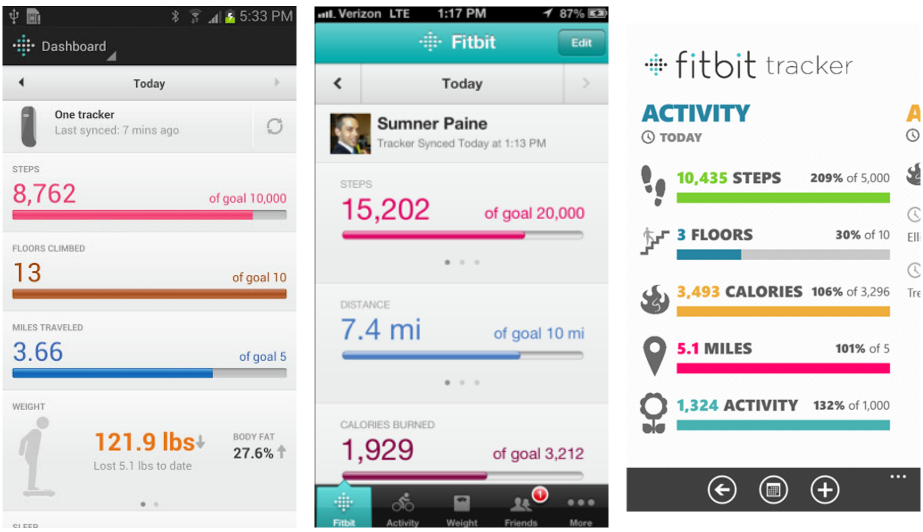


Android iPhone

Figure 2.4: History view of *Runtastic* app

Fitbit

Fitbit is a company which offers a series of quantified self devices. The data available to the users depend of the owned device. Most of them are able to track steps, distance, calories and length and quality of sleep[1]. Furthermore it is possible to view the tracked data real-time on select smartphones with both Fitbit Flex and Fitbit One on the Fitbit app. It is this app that can be seen in Figure 2.5.
The dashboards in the apps are customizable. Therefore the user can choose which data he wants to view daily. In the dashboard **the chosen data is compared to a set goal**. The comparison is done with the values of the current status and the goal, a progress bar and a percentage. The progress bar and percentage show the user how close he is to reaching his goal which can also be customized by the user.
The only difference between them is **the use of icons on the Windows Phone app**. These icons are accompanied **with the unit of measure of the activities** in order to eliminate confusion. Otherwise the meaning of the location marker without MILES could be difficult to determine by users.



Android iPhone Windows Phone

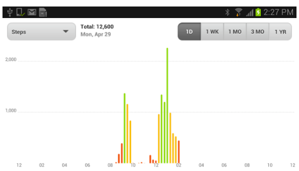
Figure 2.5: Session view of *Fitbit* app

[1] See http://www.fitbit.com/ for more information.

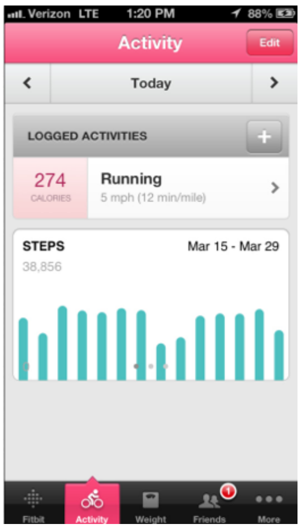
Figure 2.6 shows the various approaches used by Fitbit to display the step count history on 3 platforms. All approaches include a **bar chart without axis titles**. In the Android app **different coloured bars** have been chosen to show the user's daily steps along with the total number of steps. However it is **impossible to determine the meaning of the different coloured bars due to the missing legend**.

In the iPhone app the bars are the **same colour** but it also has difficulties. The total number of steps during the 14 day period is present but **the number of steps each day cannot be determined due to missing axes and values**.

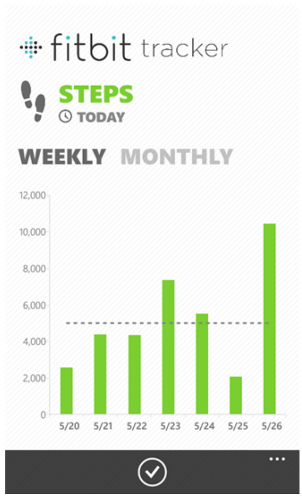
In the Windows Phone app the bar chart is accompanied with a **horizontal line representing the daily goal**. The user can then **compare his daily steps with his goal** due to both the axes' values and the horizontal line.



Android



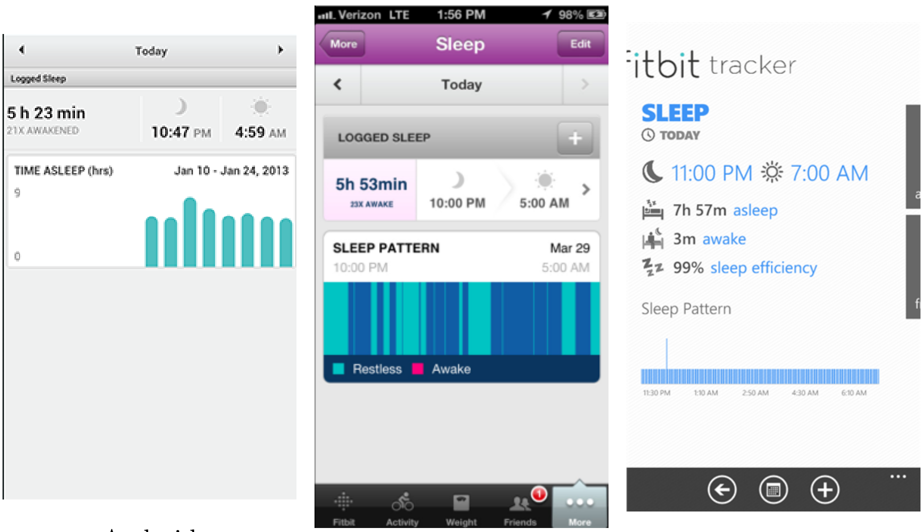
iPhone



Windows Phone

Figure 2.6: History view of *Fitbit* app

Figure 2.7 shows the various approaches used by Fitbit to display sleep data on 3 platforms. All 3 show when the user went to sleep, woke up and the duration of sleep. The differences are in the visualizations. The Android app shows the sleep history over a fortnight. It is **difficult to determine the time asleep due to both the distance between the data and the y-axis and the limited number of values on the y-axis**.
Both the iPhone and Windows Phone app have sleep pattern visualizations. The only difference between them is the legend on the iPhone. It is **possible to determine when the sleep was restless due to the legend** which is not possible on the Windows Phone app. However on the iPhone **the user can only assume the darker blue represents deep sleep due to this colour being absent from the legend**.



Android iPhone Windows Phone

Figure 2.7: Sleep view of *Fitbit* app

Just 6 Weeks

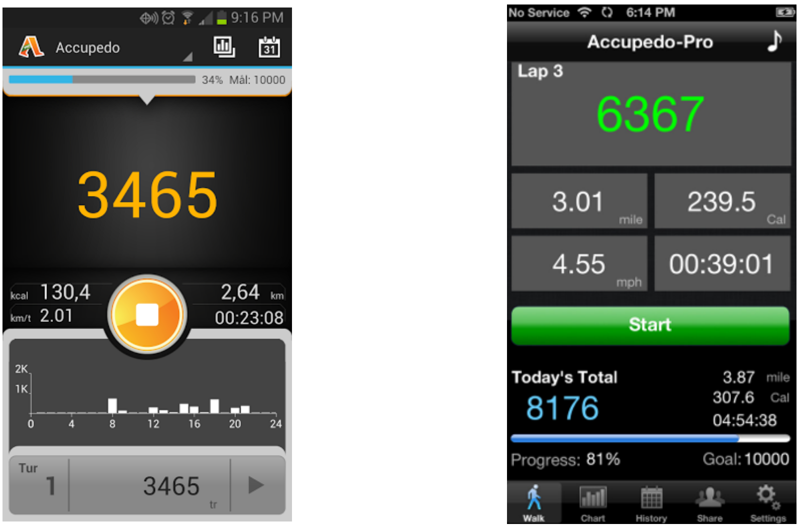
Just 6 Weeks is an app which provides it users with 5 different programs. The goal of each program is help the user to be able to do a certain number of repetitions of an exercise: 100 push-ups, 200 sit-ups, 200 squats, etcetera.
For each program it is possible for the user to view his progress. An example of this can be seen in Figure 2.8. It is possible to see one's development as **both a chart and a list**.

Android iPhone

Figure 2.8: History view of *Just 6 Weeks* app

Accupedo

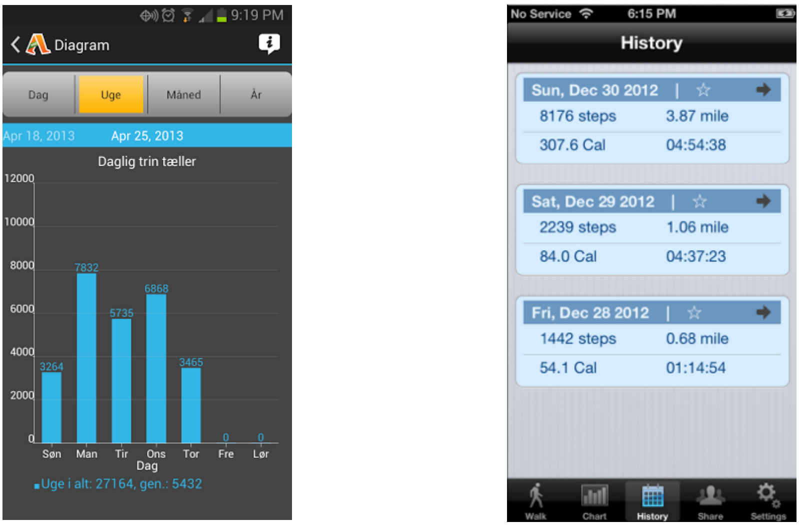
Accupedo is an app which uses the accelerometer in the smartphone compared to the Fitbit app which uses an external device. It uses algorithms to distinguish between steps and none-step movements. Accupedo uses the same approach as most exercise apps: The user chooses when the app tracks him in order to prevent excess battery consumption. The start/stop button used for this can be seen in Figure 2.9.
The only difference between the Android and iPhone app is the visualization on Android. A user could assume that it is showing one's steps during the day. However due to it **missing both axis titles and title it is not certain**.



Android iPhone

Figure 2.9: Session view of *Accupedo* app

Accupedo provides its users with a History view which allows the user to see his progress. The free Android version of the app offers its users visualizations and a list as data representations whereas the free iPhone version is limited to a list as seen in Figure 2.10.
The visualization in the Android app provides the user with **both a bar and a value representing each day**. This ensures the user can quickly determine his steps for each day and compare it to his goal. Furthermore the user is provided with the week total and average.
The list in the iPhone app also shows the step count for each day. In addition it also shows the distance, calories and duration. The **star's meaning is however questionable** since it could mean both rating and bookmark.

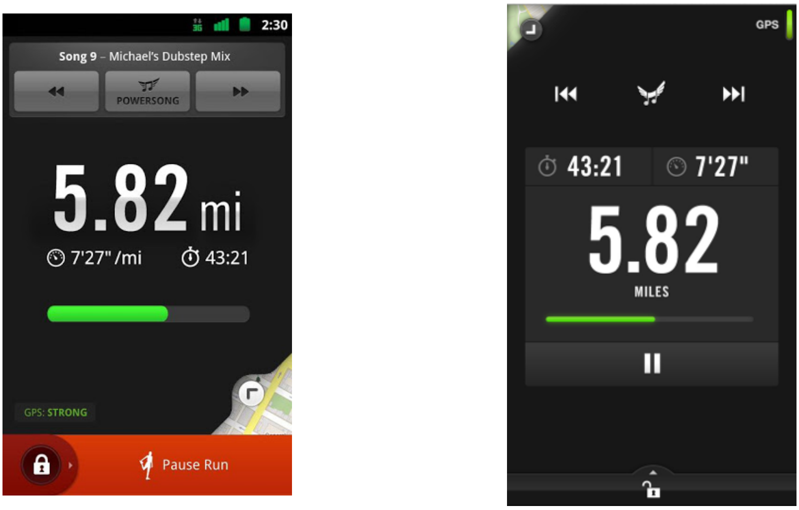


Android iPhone

Figure 2.10: History view of *Accupedo* app

Nike+

Nike+ uses the same approach as Accupedo: No additional devices besides your smartphone are needed. You just press a start/stop button and run. As it can be seen in Figure 2.11 there is almost no difference between the Android and iPhone app. Both show the distance, speed, duration and allow the user to view a map of his route.



Android iPhone

Figure 2.11: Session view of Nike+ app

The History view of Nike+ and Runtastic is similar in both design and data. However the Nike+ app in Figure 2.12 and the Android app in Figure 2.4 do differ. Only the distances in Nike+ are given a unit whereas the unit is displayed for the majority of the data the Runtastic app. It is **unclear what both *13,612* and *10'23"* represent** in the Nike+ app. Furthermore a user could assume the total distance run in both April and March is *25.16* miles. But this is **uncertain due to the use of an icon instead of *Total distance***.

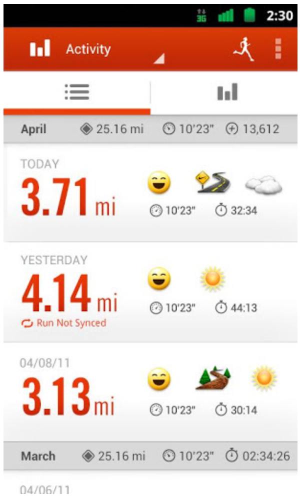


Figure 2.12: History view of the Android *Nike+* app

Sleep as Android

Sleep as Android is an app which tracks your sleep and wakes you up during light sleep in order to wake you up gently. After a night's sleep it is then possible to see one's sleep pattern during the night. An example of this is given in Figure 2.13. At first glance it is **difficult to distinguish between light and deep sleep due to the small size of the labels on the y-axis**. If this is not discovered by the user the meaning of the green and blue lines in the bottom visualization can be difficult to understand.

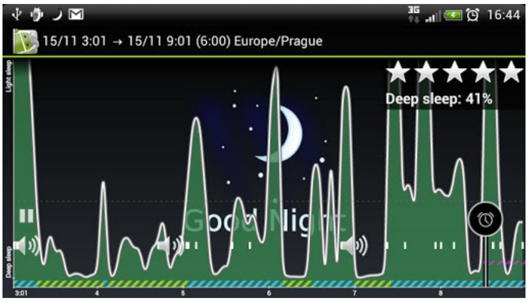


Figure 2.13: Session view of the Android *Sleep as Android* app

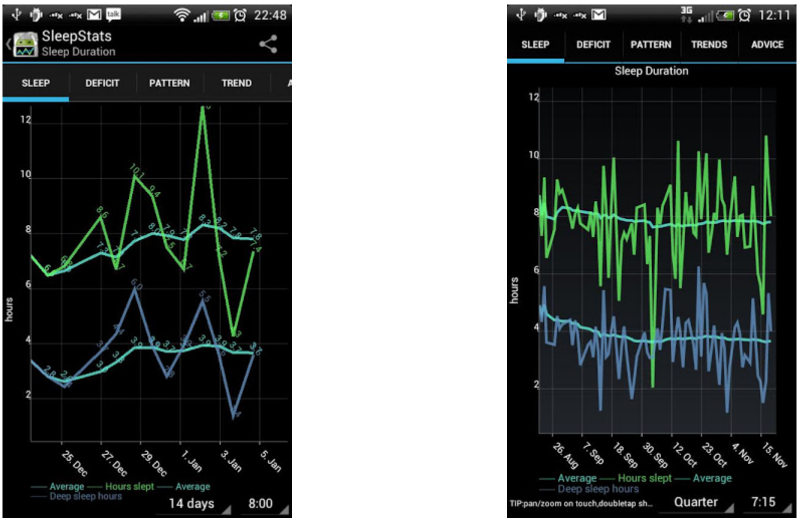
The user can also see his sleep pattern over a longer period. 2 examples of this are given in Figure 2.14. The visualizations of each night's sleep in View 1 are smaller but identical to the visualization in Figure 2.13. View 2 shows the user's sleep pattern for the last 14 days. The user can quickly determine the average for both the sleep duration and the deep sleep percentage. However **the meaning behind the sleep deficit is not clear**. The average deficit is +28 minutes the last 14 days. This means the benchmark for the sleep duration is 7.5 hours. But it is **unknown whether 7.5 hours of sleep is the recommended amount of sleep or the user's previous average**.
Traditionally the colours red and green mean bad and good respectively. This is also the case in Sleep As Android. It lets its users **distinguish between a good and bad night's sleep by highlighting them with a colour**. However the difference between a sleep deficit of -0:10 and -3:13 hours is unknown since the same colour is used for both. A user could assume both are equally bad.
The meaning of the blue colour used to highlight the deep sleep percentage is also unknown since no explanation is given. This also **prevents the user from knowing whether or not his percentage of deep sleep is too high, too low or sufficient**.

View 1 View 2

Figure 2.14: History view of the Android *Sleep As Android* app

SleepStats

SleepStats is an addition to Sleep As Android. It provides the user with more detailed statistics and visualizations. Examples of the visualizations can be seen in Figure 2.15. Both show the correlation between the sleep and deep sleep duration. In View 2 the value of each entry is absent which it is not in View 1. However **a number of chart values cannot be read because their colours is too similar to the lines' colours**.

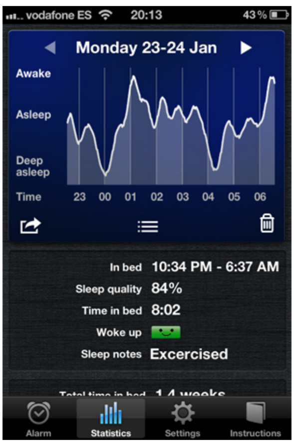


View 1 View 2

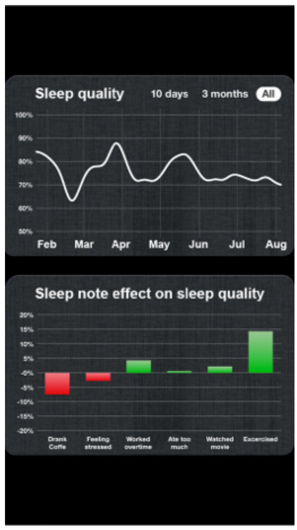
Figure 2.15: History view of the Android *SleepStats* app

Sleep Cycle alarm clock

Sleep Cycle alarm clock is similar to Sleep As Android. It also monitors your sleep and it wakes you up while you are in light sleep. It uses the built-in accelerometer to achieve this. One of the special features of this app is that it is possible for the user to **add a note to each night's sleep** in order to determine what effects his sleep quality.
Sleep Cycle alarm clock also provides its users with visualizations of the sleep data. As it can be seen in Figure 2.16 the user can see each night's sleep in detail, sleep quality over a period and what effects his sleep quality. The user can determine at which time he is in either deep or light sleep in the Session view due to the labels on the axes. The case is the same for the visualization of the effects on sleep quality in History view. The **colours also make it easy to determine what has a good effect on sleep quality if green represents a good effect on sleep quality**. This uncertainty is due to the missing legend explaining the meaning of the colours.



Session view



History view

Figure 2.16: iPhone *Sleep Cycle alarm clock* app

Comparison of apps

Table 2.1 shows the data representations the apps in this section use to display quantified self data. The majority uses a list to show the data in either the history or session view. 4 of the apps uses a list for the session view and a bar chart to show the users their history.
The apps using either colours or icons to highlight or distinguish data did so without explaining the meaning of them to the users.

App	List	Bar chart	Line charts	Scatter plots	Colours	Icons
Endomondo	✓					✓
Runtastic	✓					✓
Fitbit	✓	✓			✓	✓
Just 6 Weeks	✓		✓			
Accupedo	✓	✓				✓
Nike+	✓	✓				✓
Sleep as Android	✓		✓		✓	✓
SleepStats			✓	✓	✓	
Sleep Cycle alarm clock	✓	✓	✓		✓	

Table 2.1: The data representations used by quantified self apps

The recurring problems in these data representations are:

- Unexplained icons
- Missing axis titles and values
- Unexplained colour usage
- White space
- Small text

2.2 Quantified self

In this section the field of quantified self is described by listing the type of questions which people ask about their data, the life within the quantified self movement and a model which can be used to develop quantified self systems. The life within the quantified self movement is described by using examples from [Wol10] which lists the experiences people have with self tracking.

Questions

[LDF11] divides the questions people ask about their quantified self data into 6 categories. These categories are described in Table 2.2.

Status	Questions related to current status e.g. *What is my current step count?*
History	Questions related to long term data e.g. *How did I do this month compared to last month?*
Goal	Questions related to goals to pursue e.g. *How many hours of sleep do I need to be focused the following day?*
Discrepancies	Questions related to the comparison between one's current status and goals e.g. *How many steps am I missing to reach my daily goal?*
Context	Questions related to what was influencing one's current status *Is running with this person helping my run?*
Factors	Questions related to what influences one's status over a long period of time *Are my new shoes affecting my daily step count?*

Table 2.2: Categories of quantified self data questions

Goal questions is focused on either *"principle"* to *"program"* goals [LDF11, p. 5]. Principle goals are abstracts goals such as *I want to be healthy* whereas program goals are concrete such as *I must walk 15,000 steps a day to become healthier*. Typically one would start with a principle goal and transform this into a series of program goals. According to [LDF11] people use quantified self systems *"to help them set and complete program-level goals."*

Answers to **Discrepancies** questions can help people to *"make immediate decisions to address any differences"* between the current status and the goal [LDF11, p.

5]. However these can only be answered if the program goals have been set. If one is still at the principle level these questions are impossible to answer.

The life within the quantified self movement

Robin has constructed a quantified self which contained the data about his work, sleep and diet. One day he decided to stop drinking coffee. After 4 months of removing 20 millilitres of coffee from his cup on a weekly schedule he was successful. However one day he had a **Factor question**: Did coffee drinking improve his concentration? He could quickly determine if this was the case since he tracked his work efforts both before and after he stopped drinking coffee. **The numbers were clear**. It would not help. The data showed him that his work efforts were better without coffee.
This example proves **the weakness of the human mind**: *"People have limited memory, cannot directly observe some behaviors, and may not have the time to constantly and consistently observe some behaviors"* and *"people have such very poor sence of time"* [LDF10, p. 558], [Wol10, p. 1].
An example of the **poor sense of time** is Ben who thought he was spending a lot of time cleaning up after his roommate (Status question). He started to record his past activities to determine if this was the case. The numbers were clear. He only spent 20 minutes a day instead of the hour he assumed it was.

Technology has made it **easier to track** one's behaviour. Robin was able to track his work efforts on a computer and Ben his activities on a electronic notebook whereas trackers of the past were confined to either laboratories or notepads.
One of the other benefits of the new technologies is that the trackers **record their data honestly** because as Shaun, who recorded his drinking on `www.drinkingdiary.com`, says: *"It is silly to posture in front of a machine"* [Wol10, p. 8]. **A machine does not judge** you whereas family, doctors and dietitians could judge you.
The same is true the other way. A machine is honest. Sometimes brutally honest. *"Machines don't understand the value of forgiving a lapse, or of treating an unpleasant detail with tactful silence"* [Wol10, p. 9]. A victim of this was Alexandra who tracked 40 things about herself and stopped due to the **brutal honesty of the data**. It *"didn't respect her wishes or her self-esteem"* [Wol10, p. 9].

Another effect of the new technologies is that people have begun to share everything from pictures on Instagram to updates on Facebook. This trend has also effected the quantified self movement: *"Self-tracking culture is not particularly individualistic. In fact, there is a strong tendency among selftrackers*

to share data and collaborate on new ways of using it" [Wol09, p. 2].
An example of the **benefits of sharing** is the bipolar Jon. He built Moodscope and recorded his mood with it in order to answer his Discrepancy question: *When do I need encouragements to brighten my mood?*. Whenever his mood was down Moodscope automatically sent a few friends an email to notify them of his low mood. They would then send him an email wanting to know why he is down. This would brighten his mood.

Stage-based model

The stage-based model was *"derived from analysis of the survey of interview data"* [LDF10, p. 560]. The 5 stages of the model can be seen in Figure 2.17.

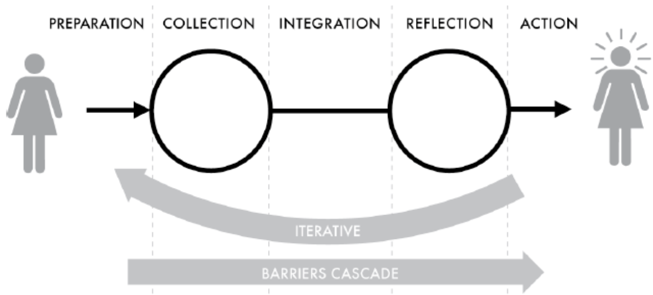


Figure 2.17: The stage based model

Source: [LDF10, p. 561]

The **Preparation** stage is when people get **motivated to collect data**, **which data to collect** and **which tools to use**. An example of this could be a person who wants to become healthier. He therefore chooses to collect data about his daily steps. For this he could use a device like Fitbit but he chooses to use an app similar to Accupedo.
After a while he realises his data is insufficient. He cannot determine whether or not he is becoming healthier. This is due to the barriers of this stage. His data is insufficient due to his choice of both data and tool. Step count is insufficient as data. Both exercise and food consumption data are needed which are not possible to track with the chosen app.

The **Collection** stage is when the quantified self data is collected. This stage has several barriers depending on the properties of the system. If it is **user-driven** more barriers exist since the user has to write the data down which is both time consuming and requires to user to remember to do it. If the system is **system-driven** the number of barriers is reduced since the data is collected automatically by the system. However the quality of the data may be reduced if the system is **uni-faceted** instead of **multi-faceted**. An app like Fitbit is multi-faceted since it collects information about both one's sleep and exercise where Accupedo is uni-faceted since it only collects information about one's exercise. A user-driven system can be both multi- and uni-faceted based on the number of different types of data people collect whereas system-driven systems tend to be uni-faceted.

The **Integration** stage is when the data is *"prepared, combined and transformed for the user to reflect on"* [LDF10, p. 561]. It is the time between the Collection stage and the Reflection stage. If the system is user-driven this stage can be long depending on the used tools. This shows how the barriers of each stage can effect other stages. This is illustrated in Figure 2.17 as **BARRIERS CASCADE**. If people choose to collect data with pen and paper (user-driven) this will lead to a long Integration stage since the data is difficult to transform in order for people to reflect on the collected data.

The **Reflection** stage is when people reflect on the collected data. Depending on the properties of the system the collected data can be presented as tables, lists and/or visualizations. This allows people to reflect on their current status and long-term data and answer their Status and History questions. But this is only if the data is presented correctly. One of the barriers of this stage is data representations preventing people from reflecting on the data. *"Great information visualization never starts from the standpoint of the data set; it starts with questions"* [Fry08, p. 4]. If the visualization is based on the data compared to one's questions it will not be able to provide the user with an answer. As one of the participants in [LDF10] stated: *"It's hard to get a holistic view of the data since the time filters are at most one month and I'd like to look at several months at once."* This participant probably has a History question he wanted to answer but due to the visualization's focus on the data rather than the question his History question remained unanswered.

The Reflection stage has been divided into 2 separate phases in [LDF11] which differ in the type of questions people ask about their data. In **the Maintenance phase** people ask Status and Discrepancy questions whereas people in **the Discovery phase** are *"mostly asking History, Goals, Context, and Factors questions"* [LDF11, p. 7]. The difference in questions is due to the goals people have set. In the Discovery phase people's goals are at the principle-level and they are attempting to set their program-level goals whereas people Maintenance phase have set their program-level goals.

The **Action** stage is *"when people choose what they are going to do with their newfound understanding of themselves"* [LDF10, p. 562]. If the barriers from the Reflection stage has prevented people from reflecting on the collected this will prevent people from acting on their data.
Instead of depending on the users' own reflections *"some systems alert the user to take actions"* [LDF10, p. 562]. This could for instance be an step count app which encourages the user to go for a walk or to do better the following day. Despite one's scepticism of this *"a long line of research in human-computer interaction demonstrates that when machines are given humanlike characteristics and offer emotional reassurance, we actually do feel reassured"* [Wol10, p. 9].

If the stage-based model is used to develop a quantified self system [LDF10, LDF11] states that one should consider the following guidelines:

- Consider the system as a whole since barriers cascade. This requires that developers applies *"lessons from different areas of research, such as lifelogging, ubiquitous computing, information visualization, and persuasive technologies"* [LDF10, p. 565].
- Systems should be flexible and allow users to change what kind of data the collect
- Use a combination of both user- and system-driven approaches
- Allow users to track multiple facets of their lives
- Ease the transition from the Discovery phase to the Maintenance phase

Chapter 3

Theory

"A proper visualization is a kind of narrative, providing a clear answer to a question without extraneous details."
[Fry08, p. 4]

This chapter examines the scientific literature about data visualization.

3.1 Visualizing data

According to [Tuf83] the most important things to do when creating visualizations are to:

- show the data
- avoid distorting the data
- make large data sets coherent
- encourage the viewer to reflect on the substance rather than anything else
- use labels and legends to explain the visualization

In the following sections the principles behind **graphical excellence**, **integrity**, **chartjunk** and **the smallest effective difference** are examined in order to demonstrate the importance of the guidelines above. The principles are examined by using visualizations which are following these principles and visualizations which are not.
Finally a process for creating the best possible visualizations is examined.

Graphical excellence

The practice of distorting data to suit one's expectations or hypotheses instead of showing the data does exist. An approach which can be used to achieve this is to use aggregations which can both mask details and mislead the viewer.
An example of this is visualizations about the deaths during the cholera outbreak in London in 1854. In order to prevent more deaths John Snow suggested removing a handle from a water pump. According to the visualizations in Figure 3.1 it solved the problem.

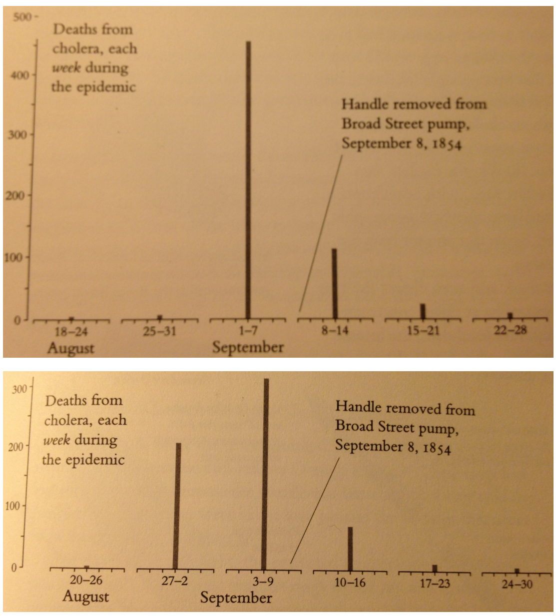


Figure 3.1: Visualisations of the outbreak of cholera in London in 1854

Source: [Tuf97, p. 36-37]

However if these visualisations are compared to the visualization in Figure 3.2 the effects of removing the handle is reduced. Furthermore it can be seen that the number of daily deaths had already begun to decrease.

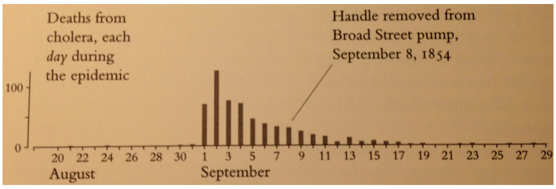


Figure 3.2: Visualisations of the outbreak of cholera in London in 1854

Source: [Tuf97, p. 37]

John Snow used a visualization similar to the one in Figure 3.2 to argue that cholera was transmitted through water and not air despite it showing that the removal of the water did not help. He **showed the data as it was instead of distorting it** to strengthen his arguments. He selected graphical excellence instead of the easy way.

Integrity

In order to keep the integrity of data intact one could avoid visualizations and use tables instead. *"Tables usually outperform graphics in reporting on small data sets of 20 numbers or less. The special power of graphics comes in the display of large data sets"* [Tuf83, p. 56]. However many visualizations contains less than 20 numbers such as the visualizations in the exercise apps described in *Apps* and visualizations can be *"preferable to tables because graphics showed the shape of the data in a comparative perspective"* [Tuf83, p. 32].

Chartjunk

E.B. White once said: *No one can write decently who is distrustful of the reader's intelligence or whose attitude toward those losers is patronizing.* This is also true for visualizations. If one thinks the **viewers cannot stay awake or are bored by visualizations which shows the data** one might resort to decorating one's visualization with chartjunk: Elements that do not add value to the visualization or distract the viewer from the information available in the visualization. The following paragraphs focus on two of the most common types of chartjunk: Grids and Ducks.

Grids are chartjunk since they *"carry no information, clutter up the graphic and generate graphic activity unrelated to data information"* [Tuf97, p. 113]. A grid may help the viewer if the visualization is a look-up table. However if the grid and the data are not easy to distinguish it can be difficult for the user to understand the visualization. Figure 3.3 shows a train schedule. In the left train schedule the data and the grid are difficult to distinguish due to the dark grid lines whereas in the left the lighter grid makes it easier for the viewer to understand the visualization and locate the required information.

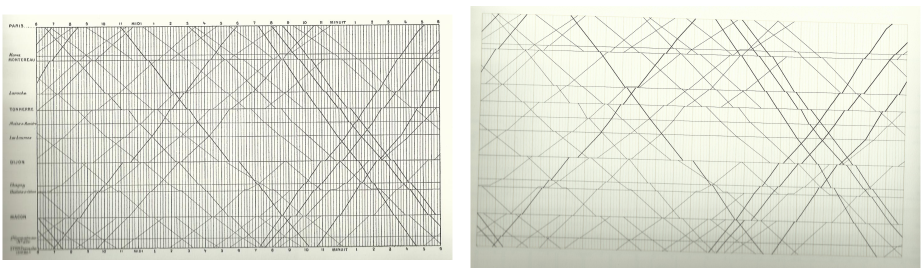


Figure 3.3: The effect of grid lines

Source: [Tuf97, p. 115-116]

Ducks are elements of a visualization which are not related to the data. Figure 3.4 contains elements which are ducks.

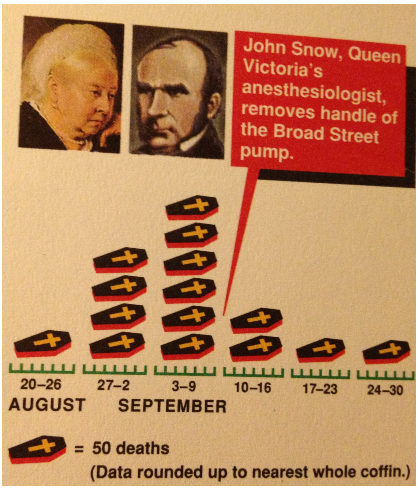


Figure 3.4: Visualization of the cholera outbreak in London with chartjunk

Source: [Tuf97, p. 37]

This visualization or "infographic" shows the deaths during the cholera outbreak in London. It differs from the previous visualizations of this event with its use of ducks. The bars have been transformed to **coffins** and **pictures of John Snow and Queen Victoria** and **a text describing their relationship** has been added to entertain the viewer. It adds no value to the visualization and patronizes the viewer.

An example of chartjunk and ducks having a catastrophic effect is the explosion of the 1986 space shuttle Challenger [1]. The visualization the manufacturing engineers presented NASA with in order to convince them of cancelling the launch can be seen in Figure 3.5.

[1] See https://en.wikipedia.org/wiki/Space_Shuttle_Challenger_disaster for more information

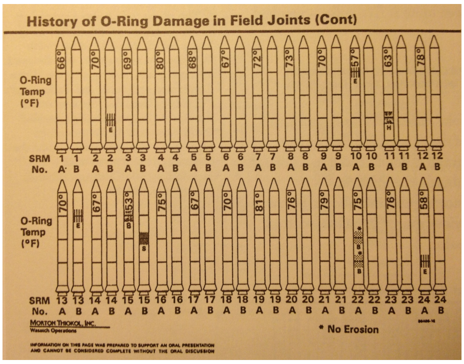


Figure 3.5: Visualization used to try to convince NASA of cancelling the launch

Source: [Tuf97, p. 48]

This visualization contains **48 ducks**: the rockets. The integrity of this visualization is diminished further when one notices the missing legend and the text in the bottom left corner: *Information on the page was prepared to support an oral presentation and cannot be considered complete without the oral discussion.* Therefore the viewer has no possibility of understanding this visualization without being present. If the viewer was present it could be possible that he still would not understand it since the legend explaining the symbols on the rockets was on a previous slide.

This approach should be avoided since it is both *"ludicrous and corrupt"* [Tuf97, p. 49]. Furthermore **a visualization should be able to stand alone**. The viewer should be able to answer his questions about the data from the visualization alone. Labels and legends should be included if they are needed to explain the visualization.

The smallest effective difference

One definition of Occam's razor is: *What can be done with fewer is done in vain with more.* This also applies to the field of visualization where *"minimal differences allow more differences"* which makes it possible to **make large data sets coherent** [Tuf97, p. 77]. Cartography is the field where this is demonstrated best as shown in Figure 3.6. The small differences in colour allow the viewer to determine the height of the land and the depth of the seas. The depth of the sea is made easier to determine by using numbering along the contours.

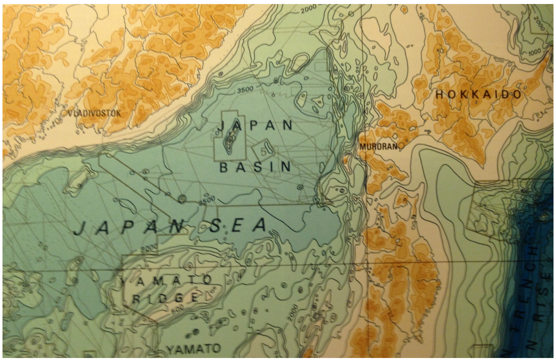


Figure 3.6: Example of the smallest effective difference

Source: [Tuf97, p. 76]

Figure 3.7 shows the effects of using the opposite approach. It has become difficult to distinguish land from sea due to the difference in colours. The data has been lost in the clutter of colours whereas the *"minimal distinctions reduce visual clutter"* in Figure 3.6 [Tuf97, p. 77].

Figure 3.7: Example of the opposite of the smallest effective difference

Source: [Tuf97, p. 77]

The seven stages of visualizing data

Technology has made it easier to get data to answer one's questions but *"it's not being used to its greatest potential because it's not being visualized as well as it could be"* [Fry08, p. 2]. [Fry08] has devised **the seven stages of visualizing data** in order to create the greatest possible visualizations based on the data. The stages are briefly described in Figure 3.8.

Acquire
Obtain the data, whether from a file on a disk or a source over a network.

Parse
Provide some structure for the data's meaning, and order it into categories.

Filter
Remove all but the data of interest.

Mine
Apply methods from statistics or data mining as a way to discern patterns or place the data in mathematical context.

Represent
Choose a basic visual model, such as a bar graph, list, or tree.

Refine
Improve the basic representation to make it clearer and more visually engaging.

Interact
Add methods for manipulating the data or controlling what features are visible.

Figure 3.8: The seven stages of visualization

Source: [Fry08, p. 5]

The 3 first stages **Acquire**, **Parse** and **Filter** are similar to the Preparation and Collection stage described in *Stage-based model*. In these stages one gets motivated to collect data to answer a question, choose the data and collect it.

The last stages is equivalent to the Integration stage described in *Stage-based model*: The data is transformed to a data representation:

- The data is represented in a properly scaled visualization in the **Represent** stage
- The scale could be chosen in the **Mine** stage by determining *"the minimum and maximum values so that it can be presented on a screen at a proper scale"* [Fry08, p. 9]
- The visualization's readability is refined *"by calling more attention to particular data (establishing hierarchy) or by changing attributes (such*

as color)" in the **Refine** stage [Fry08, p. 11].

- The visualization is made interactive to let the viewer explore the data with ease in the **Interact** stage

The seven stages are accompanied by 2 principles:

The *avoid the all-you-can-eat buffet* principle states *"that less is more. Too much information will cause the viewer to miss what's most important or disregard the image entirely because it's too complex"* [Fry08, p. 17].

The *know your audience* principle states visualizations should be aimed at a specific audience. The viewers' goals and needs have to be known in order to create the best possible visualization. Additionally it becomes possible to create simple visualizations without *"assuming that your users are idiots and "dumbing down" the interface for them"* if their goals and needs are known [Fry08, p. 17].

CHAPTER 4

Results

This chapter describes the early validation test, the *3 steps for developing a quantified self app* approach, the development of a prototype app and the results of the test of it.

4.1 Early validation test

It was necessary to complete an early validation test in order to determine the problems with quantified self apps' data representations.
The first section describes the approach of designing and conducting the early validation test.
The following sections describes the results of the test. Each of these sections focuses on the results of tests of a specific approach to data representation. Finally the problems with today's data representations in quantified self apps are determined based on the results of the tests.

Format

The most common data representations in quantified self apps were determined in *Apps* to be **lists**, **bar charts**, **line charts** and **scatter plots**. In addition to these approaches many also used **icons** and **colours** to show the data.
It was chosen to use screenshots of quantified self apps using these data representations in order to determine the problems with them. In addition to the already available data representations 3 additional data representations were created: bar charts, scatter plots and bubble charts. The bubble chart differs from the other data representations since it uses the size of the bubbles to inform the viewer of the difference between the bubbles. This data representation was included despite it going against the *graphical excellence* principle. If the data is divided into series a bubble's size will only be to scale to other bubbles in its series e.g. if [2,3] and [1] are two series the **1 bubble** might be larger than the **2 bubble**. It was still included to determine if size would effect the participants' observations.

In order to determine the problems the participants were asked write down what the data representations told about the data. The participants were also asked to write down if they liked the data representation and its user-friendliness e.g. one might prefer a list to a scatter plot but still dislike the list in question.

The participants were asked to do the test on a computer instead of a smartphone. This was to ensure that only the data representations were tested and not the smartphone's screen size.

It is possible to download the

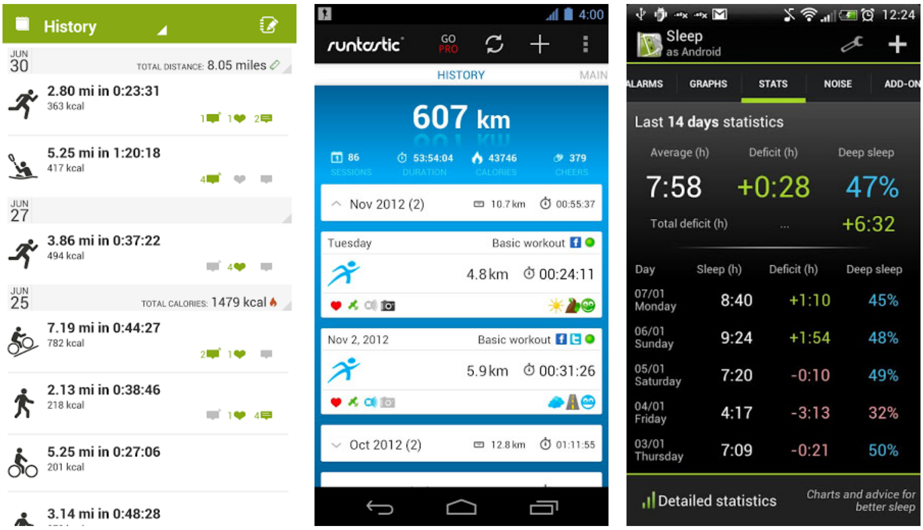
Empty questionnaire on
`https://dl.dropboxusercontent.com/u/14781368/EarlyValidationTest/Questionnaire.pdf`

The participants' filled-out questionnaires on
`https://dl.dropboxusercontent.com/u/14781368/EarlyValidationTest/Participants.zip`

Both of the above on
`https://dl.dropboxusercontent.com/u/14781368/EarlyValidationTest/Both.zip`

Tests of the list

All participants were able to understand the lists in Figure 4.1. However none were able to determine if the amount of exercise or sleep was sufficient or excessive. This was expressed by p1 who **guessed** what was enough, too little or too much.



Endomondo Runtastic Sleep As Android

Figure 4.1: Lists tested in the early validation test

Since both Endomondo and Runtastic are exercise apps p2, p3 and p4 focused on calories:

- p2 wrote that it is *"cool that Runtastic shows the aggregated calories and time"*
- p3 wrote that *"calories was the most important information"* and it is was displayed as it was not important in Endomondo
- p4 wrote that the *"calories should be given"* in Runtastic

The reason as to why they focused on calories could be that it **makes it easier to compare the different activities to each other**. p4 wrote that the comparison could be done if the different activities in Endomondo showed speed instead of distance and duration. Since it was not easy to compare p2 stated that the data *"did not motivate him"* and p3 stated that he *"did not learn anything"*. This is an example of the effect of cascading barriers described in *Stage-based model*. The choice of data representation in the Integration stage had prevented both from being motivated to take action in the Action stage.

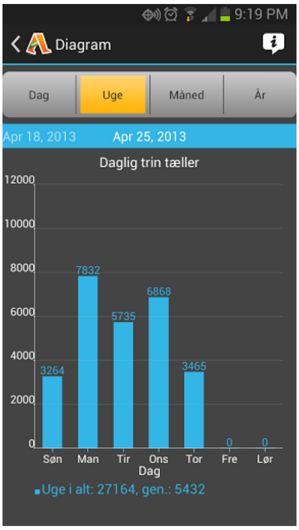
p2, p3 and p4 had difficulties interpreting the icons in Endomondo and Runtastic. As stated in *Apps* some of the icons' meaning are not clear. p4 ***"did not understand the meaning of the icons"*** in the Endomondo app. p3 experienced the same difficulties with Runtastic and thought the ***"use of icons was excessive"***. Also p2 and p4 had difficulties with the icons in Runtastic. p4 *"assumed that the type of each activity was running"* due to the blue running man whereas p2 had to guess what the smiley and surface icon represented.
The difficulties the participants experienced with icons are in accordance with the principles of avoiding *chartjunk* and the *all-you-can-eat buffet* described in *Visualizing data*. Icons does not provide the user with more information and a high number of icons can distract the user from the data. This is caused by the developers wanting to show the users all the available information in one list.

Despite the difficulties with the use of icons all participants preferred Runtastic's approach to Endomondo's due to the design and the **additional data which made it easier to make comparisons**.

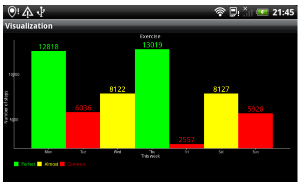
In contrast to both Endomondo and Runtastic Sleep as Android uses colours in addition to a list to allow its users to *"quickly get an overview"* (p1, p2). The purpose of applying colours is to make the representation clearer according to the *Refine stage* in *The seven stages of visualization*. However both p1 and p4 **questioned the use of colours**. p4 asked *"what the blue colour means"* and p1 assumed the green was good, the red was bad and *"the blue was between the two"*.
The result of refining the representation was confusion instead of clarity because of **the use of colours without an explanation**. Despite the confusion all participants liked the approach used by Sleep as Android. Only p3 had difficulties understanding the data representation.

Tests of the bar chart

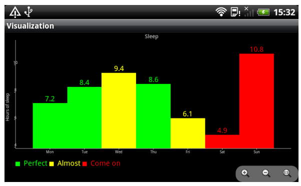
p1 and p2 liked all bar charts in Figure 4.2 whereas p4 disliked the sleep bar chart. However all participants was able to understand 2 of the data representations due to the *"good overview"* (p1, p3). The exception was the created sleep bar chart (p1, p3, p4). In accordance with the results of the colour usage in Sleep as Android in the previous section **colours confuse users if their meaning is not probably explained**. Both p1 (*"confusing with colour usage and legend"*) and p4 (*"what do the colours mean?"*) had troubles discerning the difference between 2 bad days. On the other hand they had no troubles distinguishing between good and bad days due to the colours.



Accupedo



Created exercise bar chart



Created sleep bar chart

Figure 4.2: Bar charts tested in the early validation test

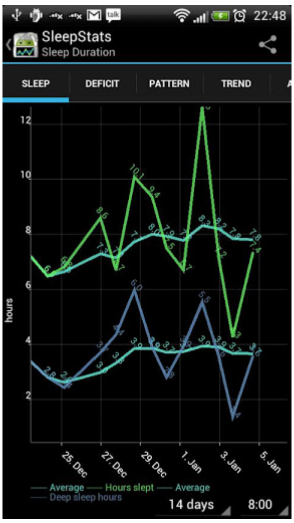
The exercise bar chart was understood by the participant more easily since the recommendations for daily steps only has a lower limit whereas sleep is recommended to be between 7-9 hours. The colours helped the participants to get *"a better overview"* (p1) and to see if *"the number of steps is good or bad"* (p1, p2, p4). Both p1 and p4 liked the use of colours but p4 noted that **"the colours are to garish"** in both created bar charts. This is in accordance with principle of the smallest effective difference described in *Visualizing data*.
p4 noted that *"a line indicating how much you need to move"* is needed to

determine the recommended daily steps. Furthermore p4 suggested that the names of the series (Perfect, Almost and Come on) should be better. This is in accordance with the *know your audience* principle described in *Visualizing data*. The series' names go against this principle since they are "dumbed down" and have become chartjunk since they add no extra value to the visualization.

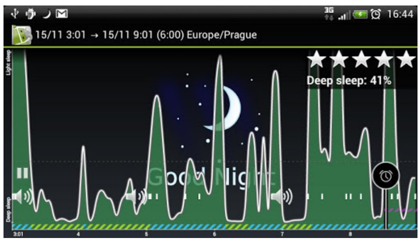
Finally Accupedo's data representation did not answer any of the participants' questions. p3 asked *"what it could be used for?"* and p2 asked *"is 10000 steps good or bad?"*. Furthermore p2 felt that it *"could be nice to compare the steps to something"* whereas p4 felt the *"week total and average should be more clear"*.

Tests of the line chart

The Sleepstats line chart in Figure 4.3 divided the participants. Only p4 liked and understood it whereas the 3 others disliked it and did not understand it.
The reason as to why p1 and p3 had difficulties with this visualization is that is goes against the *avoid the all-you-can-eat buffet* principle. Both p1 and p3 was prevented from reflecting on the visualization due to the **overload of information** which made it hard to *"distinguish data"* (p1). Another cause of the confusion is **small text** (p2, p3). This also caused problems for p2 who had troubles determining the meaning of the lines due to the small text in the legend and the y-axis.
Furthermore p4 felt that *"an indicator of how much you should sleep"* should have been included. This would tell the user if the hours slept is too much or too little.



Sleepstats



Sleep as Android

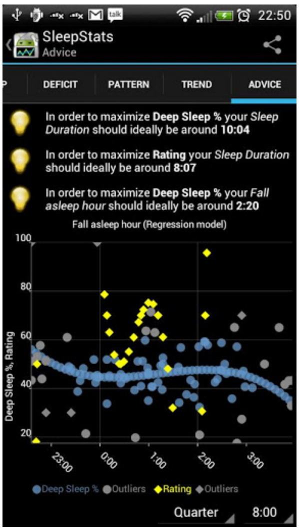
Figure 4.3: Line charts tested in the early validation test

Also the Sleep as Android line chart divided the participants. p1, p2 and p4 liked it and p1 and p4 understood it whereas p3 disliked it and did not understand it.
The reason as to why the participants preferred the Sleep as Android line chart is that it *"easier to understand"* (p2). Only p3 had difficulties understanding it. The reason could be that the ***axis labels are too small*** (p4).
p1 complimented the graphics of the design despite of the use of chartjunk (the

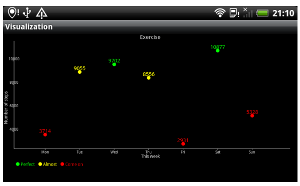
moon and Good Night in the background) whereas p4 found the white stars to be *"misleading"*: *"They look like empty stars meaning 0/5 stars. They should be coloured."*

Tests of the scatter plot

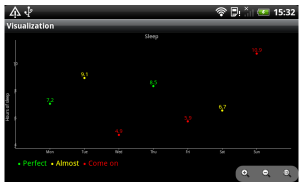
All participants disliked and did not understand the Sleepstats scatter plot in Figure 4.4. The participants found the line chart *"too confusing"* (p1 ,p2) and *"difficult to understand"* (p3, p4). p1 stated that the reason was **the amount of data**. Similar to the Sleepstats line chart *the avoid the all-you-can-eat buffet* principle was not followed and the data is lost in the clutter.



Sleepstats



Created exercise scatter plot



Created sleep scatter plot

Figure 4.4: Scatter plots tested in the early validation test

p1, p3 and p4 disliked and p3 did not understand the created exercise scatter plot. p1 and p4 thought that the values were ***too small*** which made it difficult to read. p2 and p4 found the dots confusing and p4 thought the small size of them made it **difficult to compare them**.

All participants disliked and p1 and p4 did not understand the created sleep scatter plot. The difficulties the participants had with this visualization is the same as the problems with the created exercise scatter plot and bar charts.

p4 had the same difficulties with the created scatter plots as with the bar charts: **garish colours**, **missing a line indicating how much you need to move** and **better names of the series**.

Tests of the bubble chart

The participants' feelings towards the bubble charts in Figure 4.5 were:

- p1, p2 and p4 participants disliked both visualizations whereas p3 liked them
- p2, p3 and p4 did not understand the exercise visualization whereas p1 did
- All participants did not understand the sleep visualization

Created exercise bubble chart

Created sleep bubble chart

Figure 4.5: Bubble charts tested in the early validation test

These charts confused the participants because they contain *"no data"* (p1) besides the bubbles. It is **not indicated what the size of the bubbles mean** (p2, p3, p4). Additionally the participants had the same difficulties with these visualizations as they had with the other created visualizations: garish colours, missing a line indicating how much is recommended and better names of the series.

Problems with quantified self apps' data representations

The early validation test revealed 6 categories of problems with today's data representations in quantified self apps. The categories are summarized in the 6 paragraphs below.

Icons The Windows phone app in Figure 2.2 is an example of successfully using icons in a data representation. The meaning of the icon is written next to it in order to eliminate confusion. When this is not the case the user can be **confused by an icon if its meaning is not self-explanatory**.
Another problem is **excessive use of icons** in apps such as Runtastic in Figure 4.1. The data is lost in the clutter because the user is trying to decipher the meaning of the icons.

Colours Colours can be used to highlight data if they are used correctly. However colours can also prevent users from reflecting on the data if **the colours are not explained** such as Sleep as Android in Figure 4.1. Colour usage without explanations will cause the users to guess their meaning and to have no answers to their question.
Another problem is the **difference in colours**. If the contrast between colours are high it can irritate the user like the created visualisations irritated p4. This can force users to focus on the colours instead of the data.

Information overload Users can be prevented from reflecting on the data if a data presentation **goes against the *avoid the all-you-can-eat buffet* principle** and contains **too much data** such as the Sleepstats data representations in Figure 4.3 and Figure 4.4. The data will be lost in its own clutter and the user will not be able to understand it.

Size If the **size of the text and/or numbers** is too small to read the user will be prevented from understanding the data representation. In Figure 4.3 the small text size of the axis-values in Sleep as Android prevented p3 from understanding what the data representation was showing.
Another problem was the **use of size to indicate the relationship** between data. If this is not done in accordance with the graphical excellence principle and the meaning of the size is not explained the users will be confused like the participants did in the tests of the bubble charts in Figure 4.5.

Comparisons Fitbit allows its users to **customize their dashboard** so they can see the data they want to see. This is not a approach used by all quantified self apps. The early validation test showed that some of the participants **wanted to compare different types of activities**. However they were prevented from doing this since the **data representations did not provide the participants with a unit of measure making this possible** like in Runtastic in Figure 4.1.
A recurring problem the participants had throughout the test was the **lack of recommendations**. They needed a number or a line representing what they should at least do which they could compare to what they have done. This need was expressed in all bar charts in Figure 4.2, Sleepstats in Figure 4.3 and the bubble charts in Figure 4.5.

Explanations Another recurring problem was **missing explanations**. If the user is confused by an element in a data representation and no explanation is available this will prevent the user from understanding it as mentioned earlier in *Icons*, *Colours* and *Size*.

These problems correspond to the determined problems in *Comparison of apps*. The only difference is the missing white space problem. This is due to no data representations in the test had white spaces.

4.2 3 steps for developing a quantified self app

An approach for developing quantified self apps was derived based on the results of the early validation test. The approach is also based on the principles of data visualization described in *Visualizing data*, the seven stages of visualization described in *The seven stages of visualizing data* and the stage-based model described in *Stage-based model*. The goal of this approach is to eliminate the problems of today's quantified self apps' data representations.

In order to achieve this the approach consists of 3 steps:

Step 1 Get the data

Step 2 Create the best possible data representation

Step 3 Encourage the user to take action

The approach also consist of 1 principle: **Let the user control the app**.
Each step and the principle are explained in more detail in the following sections.

Get the data

This step corresponds to the *Collection* stage in the *stage-based model* and the *Acquire* and *Parse* stages in the *seven stages to visualization*. In this step the app tracks 1 or multiple facets of the user's life. This data allows the user to answer *Status*, *History*, *Goal* and *Discrepancies* questions.
The app must allow the app to track the context of his situation in addition to tracking facets of the user's life in order for the user to answer his *Context* and *Factors* questions. This can be achieved by using the smartphone's built-in sensors (GPS, bluetooth), web services (weather) or APIs (weather, banking). An alternative to this system-driven context tracking could be a user-driven context tracking approach such as notes used in the Sleep Cycle alarm clock app. This makes it possible for the user track the most important context data according to him.

Create the best possible data representation

This step corresponds to the *Integration* stage in the *stage-based model* and the *Represent*, *Refine* and *Interact* stages in the *seven stages to visualization*. In this step the app creates the data representation. The primary focus of this step is to show the data in accordance with the *avoid the all-you-can-eat buffet* principle. This means the app must avoid *junk such as ducks, grids and icons*, *high-contrast colours* and *information overload*.
Another focus of this step is to minimize white space in data representations. White space adds no value to the data representation and can be minimized by adjusting the scale to fit the data.
The third focus of this step is to explain the data representation clearly. Elements which can confuse a user must be explained with either labels, legends, a help-button or a guide. This makes it possible for developers to create data representations without "dumbing it down".
Finally it is necessary to make the visualizations interactive in accordance with the *Interact* stage of *seven stages of visualization*. This allows users to both explore and control the visualization.
It is possible for developers to create the best possible data representation based on the data if the developer follow these guidelines.

Encourage the user to take action

This step corresponds to the *Action* stage in the *stage-based model*. In this step the app has to overcome the fact that numbers are brutally honest. If the user's data is discouraging the app must encourage the user to continue with the app and self tracking. This can be achieved by giving the system human-like characteristics, using recommendations and/or allowing sharing on social networking sites such as Facebook or Twitter.
A system can be given **human-like characteristics** by recording messages from friends and family and playing them when the user's data is both positive and negative. Another approach is to have the app send the user messages of when to do something if he is not reaching his set goal.
Sharing on social networking sites can help motivate the users since the comments on the user's data will be from friends, family or people in the same situation. This was proven by the founder of Moodscape described in *The life within the quantified self movement*.
Recommendations can help people in the Discovery phase to quickly transition to the Maintenance phase. If the app provides its users with a goal from the start the users can quickly set their program-level goals in order for them to achieve their principle-level goal.

Let the user control the app

This principle is based on the users' different needs. It should be possible for the user to change the following areas of the app:

Text All users will not have the same eyesight. Therefore some users need a **larger font** in order for them to read the text in the app as proven in the early validation test.
Another barrier is **language**. Some users will prefer to have apps with text in their native language compared to English text.

Colours Some people are colour blind. Therefore it should be possible for users to choose between different **colour schemes** which takes colour blindness and preferences into account. Some people do not mind high-contrast colours as shown in the early validation test where p1, p2 and p3 did not comment on it whereas p4 disliked it.

Data representation The early validation test showed that people likes and dislikes differ. Some like lists and others like visualizations. Therefore it should be possible for the user to select his preferred type of data representation.

Estimates and conversions The early validation test showed that the participants focused on different types of data: p2, p3 and p4 preferred calories as data instead of duration and distance in the test of the lists. The app should allow users to see a converted estimate of the tracked data in order for the users to see the data they need.

Goals The early validation test showed that people ask Discrepancies questions. They needed to compare their data to a goal. Therefore it should be possible for users to add goals to the data representations which the data can be compared to.

Personal information Adjust the app automatically to fulfil the users' needs. This can be achieved by having the user type in his personal information which can be used to change the app i.e. the font can be made larger and the recommendations and/or the colours can be changed.

According to the Reflection stage of the stage-based model described in *Stage-based model* a data representation will be great if it is based on the users' questions instead of the data. If the areas above are customizable the created data representation can be transformed from being based on the data to being based on the users' questions.

4.3 Prototype

This section describes the result of developing a pedometer app in accordance with the *3 steps for developing a quantified self app* approach. However the primary focus of the development was on Step 2 and 3.

Get the data

A pedometer app has 2 ways of getting the step count data: the **built-in accelerometer** or **external devices** such as Fitbit. However none of these approaches were used to get the data. It would require a lot of time to implement and improve a step counter algorithm in order to use the built-in accelerometer. The external devices were also not chosen since none were available at the time of development and some of the devices would require developing an interface between the smartphone and the device.
openSNP could be used as an alternative to developing an pedometer app which tracks the data itself. openSNP is *"a community-driven platform for people who*

are willing to share phenotypic and genetic information for the public"[Pla12]. A section of openSNP is dedicated to users sharing their Fitbit data. An example of the downloadable Fitbit data can be seen in Figure 4.6.

date	steps	floors	minutes asleep	minutes awake	times awaken	minutes until fell asleep
17-02-2012	17006	53	226	1	1	3
18-02-2012	16162	34	423	22	7	5
19-02-2012	15735	64	420	0	0	0
20-02-2012	10025	27	395	18	9	7
21-02-2012	12656	35	311	34	9	21
22-02-2012	18310	58	387	14	5	6
23-02-2012	13276	37	431	39	12	13
24-02-2012	13470	24	332	23	7	8
25-02-2012	20898	82	421	25	10	10
26-02-2012	11638	47	443	0	0	12
27-02-2012	9599	20	357	29	8	15
28-02-2012	16754	58	412	35	13	9
29-02-2012	11764	31	424	27	8	6
01-03-2012	15535	49	324	8	2	12
02-03-2012	11652	43	360	20	5	11
03-03-2012	17777	64	493	22	6	8
04-03-2012	17402	38	392	24	8	3
05-03-2012	9989	21	309	15	4	9
06-03-2012	12638	42	349	48	10	17
07-03-2012	11555	16	374	58	16	8
08-03-2012	306	0	443	28	14	14
09-03-2012	14425	66	387	24	11	21
10-03-2012	20730	60	329	31	14	15
11-03-2012	8426	42	384	35	15	7
12-03-2012	10418	24	434	36	8	23

Figure 4.6: Example of data available at openSNP

The data can only be downloaded as a CSV file. This would require going through the *Parse*, *Filter* and *Mine* stages of the *seven stages of visualization* in order to create a data representation. Therefore it was chosen to use random numbers since this approach was less time consuming and the focus of development was on step 2 and 3.[1]
The numbers are randomly chosen between 2,000 and 15,000 in order to simulate the steps an average user will take every day.

Create the best possible data representation

The data representation was chosen after getting the data. Since the visualizations proved more difficult to the participants during the early validation test it was chosen to create a visualization instead of a list in order to determine if the 3

[1]For more information see http://docs.oracle.com/javase/6/docs/api/java/util/Random.html

steps approach works. The chosen type of visualization was a bar chart since both of the examined pedometer apps in *Apps*, Fitbit and Accupedo, also used this type of visualization.
It was chosen to implement the bar chart with the AChartEngine charting tool. Other libraries such as AndroidPlot was also considered but they did not fulfil the requirements stated by Step 2. AndroidPlot only creates static visualizations which prevents the user from interacting with it.
AChartEngine gives the developer complete control of visualizations with Renderers. They give the developer control of colours, background, scale, labels, legends and the interactivity. Therefore AChartengine fulfil all of the requirements of Step 2.

It was chosen to create 3 different long-term data bar charts: 1 week, 1 month and 2 months. This required the use of a colour scheme. It was chosen to use a low-contrast colour scheme in accordance with Step 2. An example of a low-contrast colour scheme can be seen in Figure 4.7.

Figure 4.7: Example of a low-contrast colour scheme

The influence of the low-contrast colour scheme can be seen in Figure 4.8 which shows the 1 week bar chart. As it can be seen the bar chart shows the steps for each day, the average is both written beneath the bar chart and represented in the bar chart with a line and the chart's title explains that the bar chart shows the last week's steps. The total for the week was excluded in accordance with Step 2 since it provides the user with no additional value and can be classified as junk.

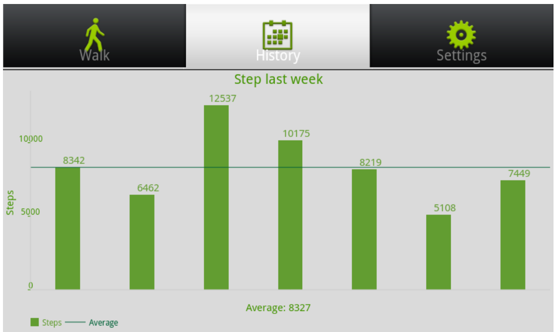


Figure 4.8: 1 week bar chart

The only difference between the 1 week and 1 month bar chart in Figure 4.9 is the removal of the chart values. They were removed to prevent information overload.

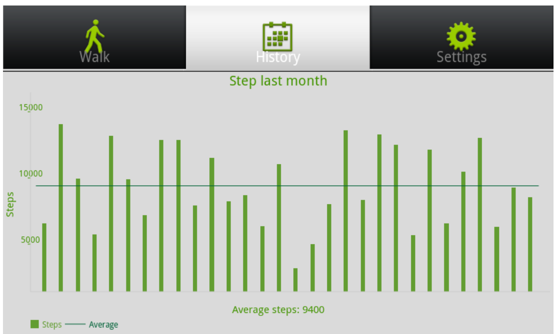


Figure 4.9: 1 month bar chart

The only differences between the 1 month bar chart and the 2 months bar chart in Figure 4.10 are that the months have been given different colours and each month's average has been calculated separately in order for the user to be able to quickly compare the 2 months.

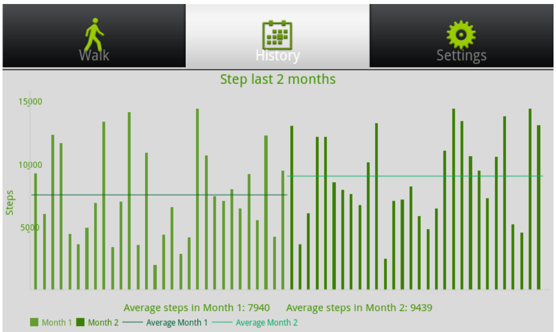


Figure 4.10: 2 months bar chart

Encourage the user to take action

The participants of the early validation test requested a number or a line indicating how much they at least need to do in several sleep and exercise apps. Therefore it was chosen to use the recommended number of daily steps in order to encourage the user and ease the transition from the Discovery phase to the Maintenance phase. According to [Ham11] the recommended number of daily steps for

Girls are 12,000 steps

Boys are 15,000 steps

Adults are 10,000 steps

Older adults are as many as possible

It was chosen to use 10,000 steps as a recommendation since the majority of the possible users will be adults. The recommendation was added to the bar charts as a horizontal line in order for the users to have a goal to compare to in accordance with Step 3.
The recommendation line can be seen in Figure 4.11. Besides the addition of this line there are no differences between this 1 week bar chart and the previous one in Figure 4.8.

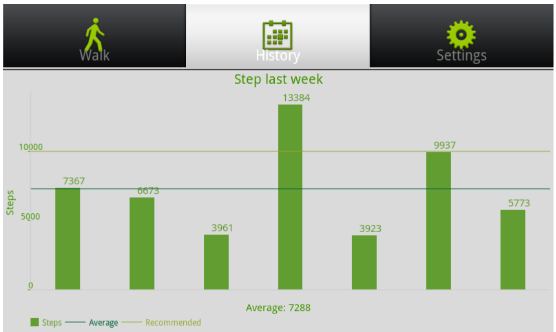


Figure 4.11: 1 week bar chart with recommendation

There are however differences between the revised versions of the 1 month and 2 months bar charts in Figure 4.12 and Figure 4.13 and the previous ones. In these the number of days the user has been below the recommended number of daily steps is written. This was not included in the 1 week bar chart since the user can see the step count of each day and compare this to the recommendation line.

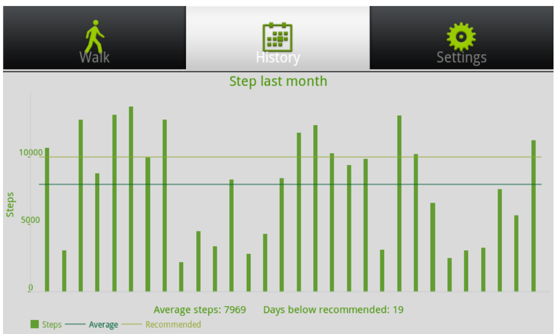


Figure 4.12: 1 month bar chart with recommendation

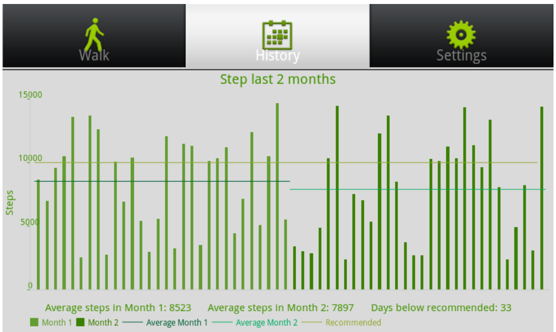


Figure 4.13: 2 months bar chart with recommendation

4.4 Prototype test

A test of the prototype pedometer app was conducted in order to determine whether or not the 3 steps approach are a solution to the problems with today's data representations in quantified self apps. This section describes the process of designing and conducting the test and the results of it.

Format

The participants were asked to write down what the bar charts in *Prototype* told about the data. This makes it possible to determine whether or not the 3 steps approach is a solution to the problems since the results can be compared to the results of the early validation test.

The participants were asked to do the test on a smartphone.

It is possible to download the

Empty questionnaire on
https://dl.dropboxusercontent.com/u/14781368/PrototypeTest/Questionnaire.pdf

The participants' filled-out questionnaires on
https://dl.dropboxusercontent.com/u/14781368/PrototypeTest/

Participants.zip

Both of the above on
https://dl.dropboxusercontent.com/u/14781368/PrototypeTest/Both.zip

Results

The results of the prototype test can be found in Table 4.1. It shows that p1 and p2 were able to see all available information whereas p3 and p4 did not see the days above and below the daily recommended steps.

Available information	p1	p2	p3	p4
Daily steps	✓	✓	✓	✓
Average daily steps	✓	✓	✓	✓
Recommended daily steps	✓	✓	✓	✓
Days above daily recommendation	✓	✓		
Days below daily recommendation	✓	✓		
2 month bar chart - The month with more steps	✓	✓	✓	✓

Table 4.1: Results of the prototype test

Both p1 and p2 noted that they would have preferred a larger font. p2 further noted that it would be easier to understand the bar chart if the language had been Danish. These 2 problems were caused since the app was developed without following the principle of letting the user control the app.

CHAPTER 5

Conclusion

The first objective of this thesis was to determine the problems with today's data representations in quantified self apps. It was determined with the early validation test that there 6 different categories of problems:

- icons
- information overload
- comparisons
- colours
- size
- explanations

The second objective was to examine the literature and studies related to visualization, quantified self, sleep and exercise. This made it possible to understand the problems:

Icons Icons are ducks according to the *chartjunk* principle. They are not related to the data and can confuse the user if they are not explained. This were the cause of some of the participants' confusion during the early validation test.

Colours The *smallest effective difference* principle can help developers to create a data representation where colours highlight data. However if the colour usage is not explained or the contrast between colours is too high the colours will confuse the users as it were the case for some data representations in the early validation test.

Information overload Less is more. A series of simple data representations are better than 1 data representation trying to explain everything according to the *avoid the all-you-can-eat buffet* principle. Some data representations in the early validation test confused the participants because they went againt this principle.

Size It is important that developers follow the *know your audience* principle since the users will have different needs. Some need a larger font to be able to read the text. Apps which did not follow this principle were difficult to understand for the participants in the early validation test.

Comparisons The questions people ask about their data vary. Some prefer to find out their current status and some want to compare their current status to a goal. According to the *know your audience* the goals of the users should be known to the developer. The apps in the early validation test which did not account for this where the apps which needed a recommendation to compare the data to in order for the participants to become motivated.

Explanations If a developer do not follow the *know your audience* principle and include elements in data representations without explaining them the data representations will confuse the users. This was also the case for several apps in the early validation test which used either colours or icons that confused the participants because their meaning was not explained clearly.

The third objective was to derive an approach which helps developers to avoid today's problems with data representations in quantified self apps. The approach consist of 3 steps and 1 principle which were based on the findings in the early validation test and the examination of the literature and studies related to data visualization and quantified self. The 3 steps are:

Step 1 Get the data

Step 2 Create the best possible data representation

Step 3 Encourage the user to take action

The principle states that developers should let the user control the app.

The fourth objective was to develop a prototype of a quantified self app with the 3 steps approach. The prototype was chosen to be a pedometer app.

The app uses random numbers as data in order to create data representations. The data representations focus on long-term data representations and consist of 3 bar charts: 1 week, 1 month and 2 months.
The bar charts were created in accordance with Step 2 of the approach:

- The charts are simple
- The colours are low-contrast
- No chartjunk is present
- The data representation is scaled
- The contents are explained
- The charts are interactive

Step 3 was followed by including a recommendation line in the bar charts for the user to have a goal to compare his data with. This line represents the users' daily goal of walking the 10,000 steps recommended to adults.

The final objective was to test the prototype. The problems the participants had with the app was the text size and the language. This could have been avoided if the app had been developed in accordance with the principle of letting the user control the app.
If the results of the prototype test and the early validation test are compared it shows that the participants in the early validation test requested either more or fewer elements in the data representations whereas the participants in the prototype test did not express these needs. This indicates that the 3 steps approach could be a possible solution to the problems with today's data representations in quantified self apps.

5.1 Future work

If the work on this thesis are to be continued the early validation test would be re-designed. The focus of the completed test is too narrow in terms of types of apps and data to state that quantified apps in general have a problem with data representations. The completed test focuses only on long-term sleep and exercise data representations. In order to confirm the hypothesis it is also necessary to test other types of apps and short-term data representations.

The prototype test would also be re-designed in order to conclude that the 3 steps approach is a valid solution to the problem. The test would have to include available apps which can be compared to an app developed with the 3 steps approach.

Both tests have a common problem: too few participants. The reason as to why the number of participants in both tests are 4 is due to their expertise. If beginners were able to tell more about the data in an app developed with the 3 steps approach compared to what experts could tell about the data in available quantified self apps the 3 steps approach could be a possible solution. However the number of participants has to be higher to conclude this.

If the 3 steps approach is concluded to be a solution, this would be re-written in order for it to help developers to create quantified self apps with data representations that benefit the users instead of preventing them from answering their questions.

Bibliography

[Fry08] Ben Fry. *Visualizing data: exploring and explaining data with the Processing environment*. O'Reilly Media, 2008.

[Ham11] Cindy Hamilton. The recommended steps per day on a pedometer. http://www.livestrong.com/article/242970-the-recommended-steps-per-day-on-a-pedometer/, 2011. Accessed: 2013-07-09.

[LDF10] Ian Li, Anind Dey, and Jodi Forlizzi. A stage-based model of personal informatics systems. In *Proceedings of the 28th international conference on Human factors in computing systems*, pages 557–566. ACM, 2010.

[LDF11] Ian Li, Anind K Dey, and Jodi Forlizzi. Understanding my data, myself: supporting self-reflection with ubicomp technologies. In *Proceedings of the 13th international conference on Ubiquitous computing*, pages 405–414. ACM, 2011.

[Pla12] Vincent Plagnol. A review of opensnp, a platform to share genetic data. http://www.genomesunzipped.org/2012/05/a-review-of-opensnp-a-platform-to-share-genetic-data.php, 2012. Accessed: 2013-07-03.

[Tuf83] Edward R Tufte. *The visual display of quantitative information*. Graphics press Cheshire, CT, 1983.

[Tuf97] Edward R Tufte. *Visual explanations: images and quantities, evidence and narrative*. Graphics Press Cheshire, CT, 6th edition, 1997.

[Vel12] Chris Velazco. Social fitness tracker endomondo hits 10 million users, tracks 250k workouts each day. http://techcrunch.com/2012/06/27/social-fitness-tracker-endomondo-hits-10-million-users-tracks-250k-workouts-each-day/, 2012. Accessed: 2013-07-01.

[Web13] Harrison Weber. Runtastic races past 30m app downloads and 25m mobile users as it launches its hardware in the us. http://thenextweb.com/insider/2013/05/13/runtastic-passes-30-million-downloads-as-it-launches-its-hardware-in-the-us/, 2013. Accessed: 2013-07-01.

[Wol09] Gary Wolf. Know thyself: Tracking every facet of life, from sleep to mood to pain, 24/7/365. *Wired. com*, 2009.

[Wol10] Gary Wolf. The data-driven life. *The New York Times*, 28, 2010.